IDENTIFICATION OF THE TARGETS OF THE HUMAN TRANSCRIPTION FACTOR PATZ1

by
MELİKE GEZEN

Submitted to the
Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Sabancı University
July 2022

ABSTRACT

IDENTIFICATION OF THE TARGETS OF THE HUMAN TRANSCRIPTION FACTOR PATZ1

MELİKE GEZEN

Molecular Biology, Genetics and Bioengineering, PhD Thesis, 2022

Thesis supervisor: Prof. Dr. Selim Çetiner

Keywords: PATZ1, CTH, Hydrogen Sulfide, CRISPR/Cas9, RNA-Seq

POZ-AT hook-zinc finger protein 1 (PATZ1) is a ubiquitously expressed N-terminal BTB domain containing transcription factor which has a C2H2 type zinc finger DNA binding domain at its C-terminal. According to the recent findings PATZ1 acts both as a suppressor and an activator in a context dependent manner. Even though regulatory role of Patz1 is associated with various cancers, such as Ewing's Sarcoma and prostate cancer, the target genes of PATZ1 are not known. As a transcription factor remains controversial due to lack of genomic studies. We identified some of the genes that are bound and regulated by the PATZ1 transcription factor in the HCT116 colon cancer cell line by using techniques such as CRISPR/Cas9 genome editing tool, RNA-seq and CHIP-seq data analysis. We showed that PATZ1 binds to the promoters of its target genes and regulates their expression by both activating and suppressing these targets. For the first time in literature, we identified transcriptional targets of PATZ1 in HCT116 human colon cancer cell line.

ÖZET

INSAN TRANSKRİPSİYON FAKTÖRÜ PATZ1 PROTEİNİN HEDEFLERİNİN BELİRLENMESİ

Melike Gezen

Moleküler Biyoloji, Genetik ve Biyomühendislik, Doktora Tezi, 2022

Tez Danışmanı: Prof. Dr. Selim Çetiner

Anahtar Kelimeler: PATZ1, CTH, H2S, DNA Hasarı, Transkripsiyon, RNA-Dizileme

POZ-çinko parmak proteini 1(PATZ1) ifadesi tüm hücrelerde görülen, N-ucunda BTB bölgesi ve C-ucunda çinko parmak bölgesi bulunduran bir transkripsiyon faktörüdür. Son yapılan çalışmalara bakacak olursak PATZ1 hem transkripsiyonel aktivatör hem de baskılayıcı olarak hücre içeriğine göre çalışır. Ewing's Sarkoma ve prostat kanseri gibi bir çok kanser türü ile ilişkilendirilmesine rağmen PATZ1'ın hangi genleri kontrol ettiği henüz genomik çalışmaların eksikliği sebebi ile bilinmemektedir. CRISPR/Cas9 genom düzenleme, RNA dizileme ve kromatin immün çökeltme ve dizileme analizi gibi yöntemler kullanarak, HCT116 insan kolon kanseri hücre hattında, PATZ1'ın bağlanarak kontrol ettiği genlerden bazılarını belirledik. PATZ1'in hedef genlerinin promoter bölgelerine bağlanarak bazılarının ifade seviyelerini arttırdığını bazılarının ifadesini ise baskılayarak kontrol ettiğini gösterdik. Literatüre ilk defa PATZ1'ın HCT116 insan kolon kanseri hücre hattında kontrol ettiği genleri kazandırdık.

Ailem’e…

ACKNOWLEDGEMENTS

I would like to thank to my dear advisor Prof. Dr. Batu Erman, who accepted me as his Ph.D. student to work in his research laboratory which I called “home” for last 6 years. During my Ph.D., he supported me to improve myself as a scientist, always guiding me, encouraging me, and helping me to discover my passion; science, while giving me the motivation to reach my goals. From the first day I started my journey at the laboratory until the last day of my Ph.D., I felt his kindness and endless support.

I am grateful to my “lab-mates” from Erman laboratory, who supported me with their friendship and stayed by my side whenever I needed them. I would like to thank all the students, interns and friends who took part in this project with me. I appreciate all the help and support from Liyne Noğay, Sanem Sarıyar, Sofia Piepoli, Sarah Barakat, Pegah Zahedimaram, and other members of our laboratorty. I was blessed for having Liyne and Sanem as my “partners in crime”, and amazing lab partners. I was lucky to have all the helpful scientific discussions about experiments we struggled with fun breaks in between. It would be extremely hard to cope with all the obstacles came to my way without their calming hugs.

I must acknowledge my friends, Ghazaleh Gharib, Zaeema Khan, Selin Acar and Hülya Yılmaz for their presence in my life. I feel lucky to have them as my allies. I would like to thank them for their endless patience, love, support, understanding and encouragement. Finally, I want to thank to my family and especially to my best friend, my mother Zühal for their never-ending love and support. They always believe in me.

I would like to thank to the Scientific and Technological Research Council of Turkey (TÜBİTAK) for supporting my doctoral studies. This study was supported by The Scientific and Technological Research Council of Turkey with the grant number 118Z015 and 20AG007.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

α	Alpha
β	Beta
µ	Micro
Amp	Ampicillin
A	Amper
AP	Alkaline Phosphatase
bp	Base pair
BTB	Broad-complex, Tramtrack, and Bric-à-brac
rSAP	Shrimp Alkaline Phosphatase
gRNA	Guide RNA
Col	Colony
Da	Dalton
DBD	DNA Binding Domain
DEG	Differentially Expressed Gene
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethylsulfoxide
DN	Double Negative
DNA	Deoxyribonucleic Acid
DP	Double Positive
DSB	Double Strand Break
ds-cDNA	Double Strand Complementary DNA
E. Coli	Escherichia coli
FACS	Flourescence Activated Cell Sorting
FBS	Fetal Bovine Serum
GEO	Gene Expression Omnibus
GFP	Green Flourescent Protein
GO	Gene Ontology
GOF	Gain of Function
GSEA	Gene Set Enrichment Analysis
h	Hour
HBS	HEPES-Buffered Saline

Indel	Insertion Deletion
Kan	Kanamycin
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knock Out
LB	Luria Broth
M	Molar
mRNA	Messengar RNA
NCBI	National Center for Biotechnology
Neo	Neomycin
NES	Nuclear Export Signal
NHEJ	Non-Homologous End Joining
NLS	Nuclear Localization Signal
OD	Optical Density
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEI	Polyethylenamine
POZ	Poxviruses and Zinc-finger
PRD	proline rich domain
Pu	Purine
Py	Pyrimidine
RE	Response Element
RFLP	Restriction Fragment Lenght Polymorphism
RNA	Ribonucleic Acid
RNA-Seq	RNA Sequencing
RPKM	Reads Per Kilobase Per Million
rpm	Rotation per minute
RT-PCR	Real Time Polymerase Chain Reaction
SDM	Site-Directed Mutagenesis
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrilamide Gel Electrophoresis
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
SV40	Simian Virus 40
TALEN	Transcription Activator-Like Effector Nuclease
TBE	Tris Borate EDTA

TD	Tetramerization Domain
UV	Ultravioloet Light
V	Volt
WT	Wild Type
ZF	Zinc Finger

1. INTRODUCTION

1.1. Transcription Factors

Changes in gene expression are mainly regulated by transcription factors (TFs), which are proteins that bind to DNA usually localized in the 5' -upstream region of target genes, alone or in complex with other proteins that regulate the transcription machinery[1]. While there is a lot of information about transcription factors in the literature, the mechanism of their regulation of gene expression is still not known completely. Previous work uncovered both the conserved and unique mechanisms underlying the regulation of genetic expression by TFs. Understanding the mechanisms by which TFs evolved and how they impact the evolution of other regulatory molecules can give insight into organism evolution and phenotypic diversity. TF function involves two core features: 1) The ability to recognize and bind to short, specific DNA sequences within regulatory elements; and 2) the ability to recruit or interact with proteins that participate in transcriptional regulation. Most transcription factors function by using their DNA-binding domains (DBDs) to recognize regulatory regions in promoters and enhancers in a sequence-specific manner [2]. TFs are known to interact in a variety of ways, including assisting each other in binding DNA (cooperative binding) or changing the chromatin state or transcription through various processes (synergistic regulation). Transcription factors can be classified according to their regulatory function and can be classified based on their structural properties such as the structure of their DNA-binding domains (DBDs). In some cases, the structural categorization of TFs can also reveal their function. TFs with a homeodomain, for example, are typically linked to developmental processes, whereas those with a "winged" helix-turn-helix (HTH) motif are linked to the interferon regulatory factor family and the induction of immunological responses to viral infections [3]. A single

transcription factor (TF) can recognize many different DNA binding sequences with varying binding affinities. The identification of DNA binding motifs is a crucial step in understanding of the regulatory roles of transcription factors (TFs) and how they affect gene regulatory networks.

1.1.1. Zinc Finger Containing Transcription Factors

Transcription Factor IIIa (TFIIIa) from Xenopus *laevis* was the first ZNF discovered in the late 1980s. As a result, a novel class of transcriptional activator proteins containing a 30-amino-acid repeating region was discovered. This new family of proteins had the ability to bind to particular DNA sequences [4]. The classical C2H2-type zinc finger motif is the most common form of zinc finger motif. The C2H2-type ZNF family has 5,926 members, according to the InterPro database (updated on April 14th, 2016). The common C2H2-type ZF is a short protein motif made up of two histidine and two cysteine residues that keep a zinc ion in place by coordination bonds. The C2H2 motif (Fig.1), which is commonly seen in tandem arrays, is composed of the sequence Y/F-X-C-X2–4–C-X3–F-X5–L-X2-H-X3–5–H, where X stands for any amino acid residue and F represents hydrophobic amino acids. The structure of such ZF consists of helix and antiparallel beta structures [5]. Cysteine/histidine combinations such as C2–H2, C2–CH, and C2–C2 distinguish non-classical zinc-finger types from classical zinc-finger types. By interacting with hydrogen donors and acceptors exposed in the major groove of DNA, four amino acid residues near the tip of the finger in the N terminal region of the helix engage in DNA recognition. The DNA sequence recognition specificity of zinc fingers is determined by the identity of the amino acids at the contact site. By altering these amino acids, a high degree of selectivity for a certain three-base-pair DNA sequence may be accomplished [6]. Protein modules containing several zinc-finger motifs, each of which recognizes a distinct three-base-pair DNA sequence, have been created to bind to specific DNA sequences via this recognition process. The first effective technique for introducing breaks at specified places of genomic DNA was to combine this recognition module with a sequence-independent endonuclease [7]. ZF proteins have a highly conserved seven-amino-acid inter-finger spacer, TGEKP(Y/F)X, in addition to the conserved sequence inside the zinc finger region. This region, known as a H/C link, has recently been shown

to be phosphorylated on its threonine residue during the cell cycle's mitotic stage, rendering the protein inactive [8]. Several ZF domains can bind to RNA [9] and mediate protein–protein interactions [10]. ZF-containing transcription factors have extra conservative amino acid motifs in addition to zinc fingers. These structural domains are thought to have a role in controlling intracellular localization, DNA binding, and gene expression through regulating how these transcription factors interact with one another and other cellular targets.

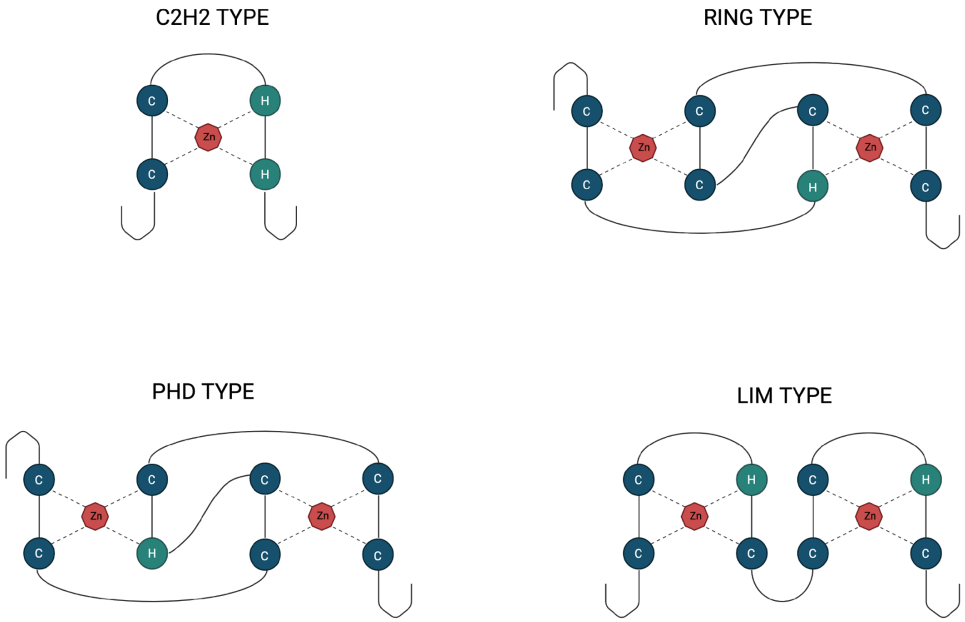


Figure 1. A schematic representation of the structure of C2H2, RING, PHD, and LIM zinc-finger domains.

ZFs localize in different cell compartments such as the nucleus (GATA2), cytoplasm (MDM2), cytoskeleton (ABLIM1) and membrane (TRAF4)[11]. ZF proteins contribute to variety of physiological processes, namely cell proliferation, differentiation, and death, through regulating gene expression, therefore maintaining tissue homeostasis. The role of ZFs in cancer start and progression has recently been underlined. Many tumor suppressor genes and oncogenes belong to the zinc-finger family. ZFs are implicated in all the major cancer development pathways, from carcinogenesis to metastasis. Furthermore, ZFs have a role in cancer via acting as transcription factors[12]. In cancer

patients, ZF expression is increased or downregulated, indicating that ZFs can operate as tumor suppressors or oncogenes. Furthermore, the actions of several ZFs appear to be tumor specific. As a result, an attractive target for drug development is specific ZFs in order to prevent or repair aberrant protein expression. Furthermore, given the great specificity of several ZFs in terms of function and expression for certain tumors, this protein class might be valuable as a prognostic indicator.

1.1.1.1. The BTB-ZF transcription factors

The developmental mutants tramtrack [13], broad-complex [14], and bric-à-brac [15] were used to clone some of the first C2H2 ZF genes in Drosophila melanogaster. The N-terminal homology of the Broad-complex and Tramtrack proteins was discovered, and this was further extended to the mammalian ZF proteins ZBTB14 and ZBTB25 (respectively known as ZF5 and KUP at the time) [16]. The adoption of the abbreviations BTB (for Broad-complex, Tramtrack, and Bric-à-brac) or POZ (for Poxvirus and Zinc finger) for this newly discovered domain was prompted by the discovery of a longer region of homology among other Drosophila proteins [15] and several poxvirus proteins [17]. The BTB domain is 120 amino acids long, usually found in a single copy, and is linked to a number of other domains, including the C2H2 zinc finger and Kelch domains [18].

The BTB domain's primary molecular function is to facilitate protein-protein interactions, albeit the implications of these interactions may vary depending on the protein's overall context [18]. 156 BTB domain-containing proteins are encoded by human genome and 49 of these are commonly known as BTB-ZFs which contain C2H2 ZF domains. These TFs are involved in key biological processes such as development, malignant transformation, and lymphocyte differentiation. Proteins generally use their BTB domains for oligomerization and the recruitment of other transcriptional repressors and activators. BTB-ZFs can also engage other histone modifying enzymes like HDACs, causing differences in the epigenome. The BTB domain may be found in a variety of proteins, including ion transporters, signal transducers, and proteases, and its role in TFs is a relatively late evolutionary adaptation in higher organisms [19]. Dimerization appears to be critical for controlling BTB-ZF TF activity by increasing the affinity and specificity

of DNA recognition or rendering dimers inactive below a particular binding threshold. As a result, BTB-domain-mediated heterodimerization is an important step in gene regulation. The replacement of a dimer component can shift the complex from one that represses transcription to one that promotes it [19]. The BTB domain also recruits transcriptional co-regulators, which are critical for mediating fast epigenetic changes in chromatin structure via histone methylation or acetylation. BTB-ZF TFs have also been demonstrated to have structural and organizational roles, particularly in mediating the enhancer–promoter interactions by encouraging DNA loop formations (Fig.2) [20].

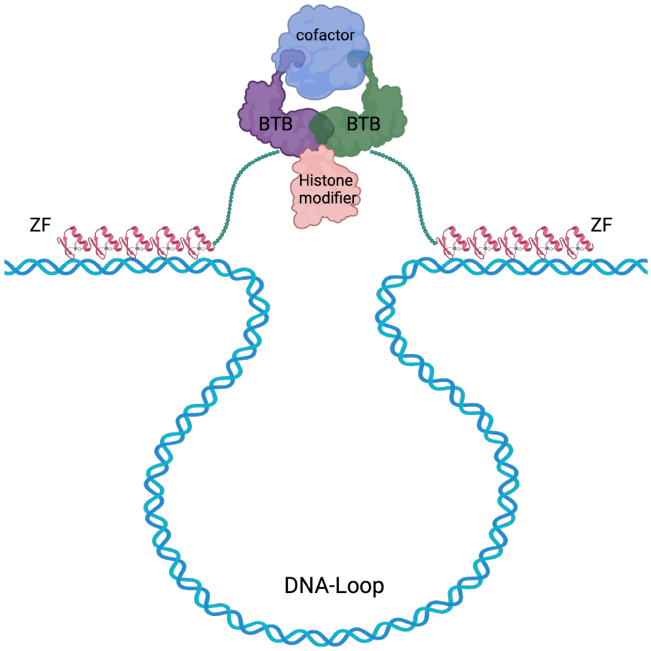


Figure 2. Representative figure of BTB-ZF TF dimerization, recruitment of transcriptional cofactors, histone modifying enzymes and DNA-Loop formation. BTB domains of BTB-ZF proteins (green and purple) can homo- or heterodimerize, creating a scaffold onto which cofactor proteins can dock (blue). BTB domains are found in the N termini of BTB-ZF proteins, which bind to DNA using different numbers of zinc finger motifs (red).

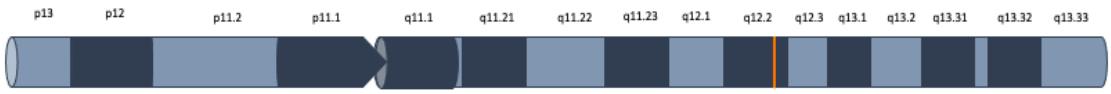
BTB-ZF transcription factors control cellular proliferation by inhibiting or inducing cyclin kinase inhibitors. Some of those TFs have also been demonstrated to promote or inhibit cell survival through regulating members of the Bcl2 pro-survival family. As a

result, their dysregulation is frequently linked to cancer. Bcl6 is most likely the BTB-ZF TF whose role as a proto-oncogene is better understood in B cells [21]. 40 percent of diffuse large B-cell lymphoma (DLBCL) and 5–10 percent of follicular B-cell lymphoma have Bcl6 chromosomal translocations.

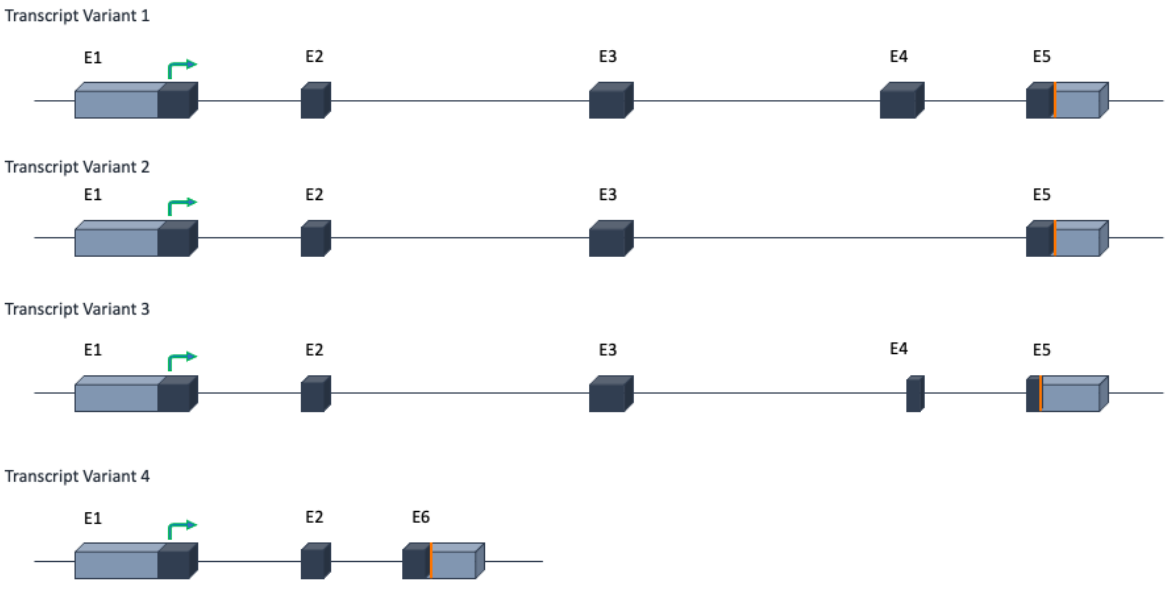
1.2. PATZ1

The POZ/BTB and AT-hook-containing Zinc finger protein 1 (PATZ1) also known as Zfp278 or MAZ-related factor (MAZR) gene encodes four protein isoforms, and it is a member of the POZ and Zinc Finger (POK) transcriptional factor family, which is involved in a variety of biological processes including embryogenesis [22], senescence [23], T-cell development [24], and cancer [22]. It may behave as an activator or a repressor depending on the cellular context. The human Patz1 gene is located on chromosome 22q12.2 and has four alternative spliced variants that produce four structurally similar isoforms (Fig.3). PATZ1 was discovered as one of the interaction partners of BACH2 [25], a BTB-ZF protein, and RNF4, a mediator of androgen receptor activity [26].

a.



b.



c

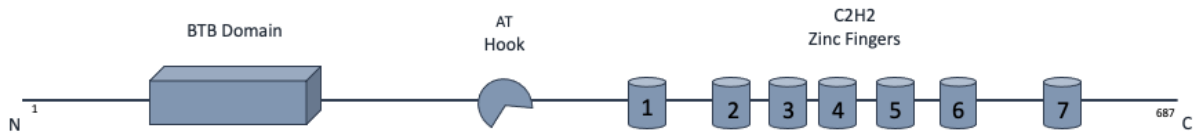


Figure 3. The Patz1 gene and Patz1 alternative splice variants a. Chromosomal location and b. Alternative Splicing Variants of the human Patz1 gene. c. Schematic representation of hPATZ1 protein structure.

PATZ1 can bind to proteins that are involved in chromatin stability and remodeling such as High Mobility Group AT-Hook 1 (HMGA1) [26]. PATZ1's functional relationship with the pluripotency master genes Nanog and Pou5f1, which are both targets and regulators of Patz1 gene, has been characterized as crucial in the preservation of embryonic stemness [27]. Newly born PATZ1 deficient mice are smaller in size and at a non-Mendelian frequency which shows the importance of PATZ1 in embryonic development [28]. PATZ1 transcription factor may regulate gene expression in a variety of ways, depending on if it operates alone or in interaction with other co-factors. For example, PATZ1 inhibits the expression of Cd8 gene during the Double Negative stage of T lymphocyte development by attracting chromatin modifying enzymes to the Cd8 enhancer, such as N-CoR and it represses the Thpok gene via binding to its enhancer during the Double Positive stage, changing these cells into the CD8+CD4- lineage [28,29]. PATZ1 is expressed in various areas throughout early embryogenesis, including the central nervous system, where it is mostly found in proliferating neural progenitors (NPCs) of the periventricular neocortical neuroepithelium, suggesting a function for Patz1 in neurogenesis [30]. In addition, Patz1 inactivation causes a severe impairment in the number and regeneration capacity of neural stem/progenitor cells (NS/PCs) that points out a key role of PATZ1 in maintenance and proliferation of NS/PCs [31]. The fact that high Patz1 expression occurs in the adult hippocampus region, which is important for many aspects of learning, memory, and mood regulation, implies that abnormal PATZ1 expression may have a role in neurological and psychiatric illnesses [31]. Recently it has been shown that in Parkinson's Disease related DYRK1A locus, an SNP; rs11088398 strongly disrupts a PATZ1 binding site within a neuron-specific enhancer which has direct interactions with an upstream DYRK1A promoter [32]. In one of the recent studies,

PATZ1 expression in human gliomas was studied, and it was shown to be overexpressed relative to normal brain and to be highly enriched in the proneural glioblastoma multiforme (GBM) subtype, where it acts as a useful prognostic marker. They also discovered that the level of PATZ1 protein in GBM which co-localizes with stem cell markers is higher in GBM-derived stem cells than tumor differentiated cells and corresponds with their ability to develop as spheres [33]. It has been shown that during myogenic differentiation MyoD acts as a repressor on a subset of H3Kme2/3 marked genes, which include the pluripotency transcription factor PATZ1 [34]. PATZ1 was discovered in a chromosomal translocation event, like many other BTB-ZF proteins involved in a condition named Ewing's sarcoma that is a very serious bone and soft tissue tumor that generally affects children and is caused by the BTB domain of PATZ1 being translocated into the zinc finger domain of EWS [35]. In another study, PATZ1 expression was discovered to be an independent predictive factor for serous ovarian carcinoma (SOC) and PATZ1 has been shown to suppress tumor growth by inhibiting tumor cell proliferation and invasion [36]. PATZ1 reduces cell proliferation by increasing CDKN1B expression, implying that PATZ1 acts as a tumor suppressor in liver cancer as well [37]. A different study shows that in the colon cancer cell line SW1116, promotes cell growth, and its knockdown suppresses cell proliferation which indicates that PATZ1 acts as a protooncogene in this cell line [38]. These data suggest that PATZ1 is involved in both tumor formation and suppression in a context dependent manner. Yet this remains to be clarified.

1.3. Hydrogen Sulfide (H2S) and Cystathionine Gamma-Lyase (CTH/CSE)

Hydrogen sulfide is the third endogenous gaseous transmitter following carbon monoxide and nitric oxide [39]. It is involved in many physiological processes such as angiogenesis, vasorelaxation [40], inflammatory response [41] and neuronal activity [42]. Therefore, it is also associated with a number of diseases such as heart failure [43], atherosclerosis [44] and neurodegenerative diseases [45].

Endogenous production of H_2S in mammalian cells is mainly from three enzymes: cystathionine β-synthase (CBS), cystathionine γ-lyase (CTH/CSE) and 3-

mercaptopyruvate sulfotransferase (3-MST). CBS is mainly expressed in central nervous system where as CTH (CSE) is expressed in peripheral tissues [46]. Both enzymes are pyridoxal-5′-phosphate-dependent which use L-cysteine as substrate. L-cysteine, the substrate of CBS and CTH, can be obtained from dietary sources or released from endogenous proteins. It may also be produced endogenously from l-methionine via the transsulfuration pathway as shown in figure k [47], which involves homocysteine as an intermediary [46]. Transitional presence of H_2S in low concentrations of 30uM to 100uM intracellularly creates a challenge of accurate real-time measurement of H_2S in biological preparations which represents a major impediment to H_2S investigations [48]. There are several methods to measure intracellular H_2S levels such as methylene blue method, MBB method, S^{2-}-sensitive electrodes method and Polarographic H_2S sensor method. One of the most commonly used methods is using Fluorimetric Probes such as sulfidefluor-1/2 (SF-1/ 2), dansylazide azide, sulfide-selective fluorescent probe-1/2, and 7-Azido-4-methylcoumarin [49]

Persulfidation of specific cysteine residues has been implicated in endogenous H2S-mediated posttranslational modification of numerous proteins. Persulfidation, in a context-specific way, can result in both increase and attenuation of protein activity [50]. In a recent study it has been shown that H_2S is involved in the disposition of mitochondrial DNA (mtDNA). Because of decreased expression of mitochondrial transcription factor A (TFAM), CTH deficient mice have lower mtDNA content and copy number. In normal conditions H_2S sulfhydrates the transcriptional repressor interferon regulatory factor 1 (IRF-1) and enhances its binding to the promoter of DNA methyltransferase 3a (Dnmt-3a) to inhibit its expression. In contrast, under low H_2S conditions, this binding does not occur and Dnmt3a expression increases resulting methylation of target promoters such as TFAM, whose expression is thereby diminished [51]. H_2S controls several cytoprotective actions. H_2S controls inflammatory responses by sulfhydrating the p65 subunit of nuclear factor kappa B (NF-kB), in addition to its role(s) in oxidative stress. The transcription factor specificity protein 1 (SP1) increases Cth gene expression when macrophages are activated by TNF. This results in the formation of H_2S, which sulfhydrates the NF-kB p65 subunit. The ribosomal protein S3 (rps3) binds to sulfhydrated NF-kB, resulting in enhanced interaction with the promoters of cytoprotective target genes [52]. In another study, scientist proved that diminished sulfhydration of Keap1 protein (which sequesters Nrf2 and target it for proteasomal degradation) in CTH knockout mice, and MEFs derived

from the CTH knockout mice show higher levels of oxidative stress and signs of senescence [39]. Cysteine is a component of the primary antioxidant GSH and a strong antioxidant itself, in addition to its involvement in protein synthesis. Aberrant redox homeostasis and neurodegeneration have both been related to disruption of cysteine and GSH metabolism [47]. CTH and CBS both play significant roles in redox balance management. In hepatic cells, almost half of the cysteine produced by the transsulfuration process is used for GSH biosynthesis, according to studies [53].

CTH coded by the Cth gene is the only enzyme that can directly generate cysteine de novo in mammals [54]. Genetic loss of CTH causes cystathioninemia (or cystathioninuria: MIM 219500) which is an autosomal recessive inborn disease with increased plasma/urinary levels of cystathionine with an estimated incidence of 1 in 73,000–333,000 [55]. CTH is an inducible gene found in a variety of cells and tissues, and it plays a role in the development of a variety of diseases. In these situations, the signaling and regulatory mechanisms for aberrant expression or functional control of the CTH/H2S system are dynamic and complicated. CTH expression and activity at several levels are controlled by different mechanisms including posttranscriptional and epigenetic mechanism. For example, the direct binding of Sp1 transcription factor to human CTH promoter was demonstrated by CHIP and Dual-Luciferase Assay [56].

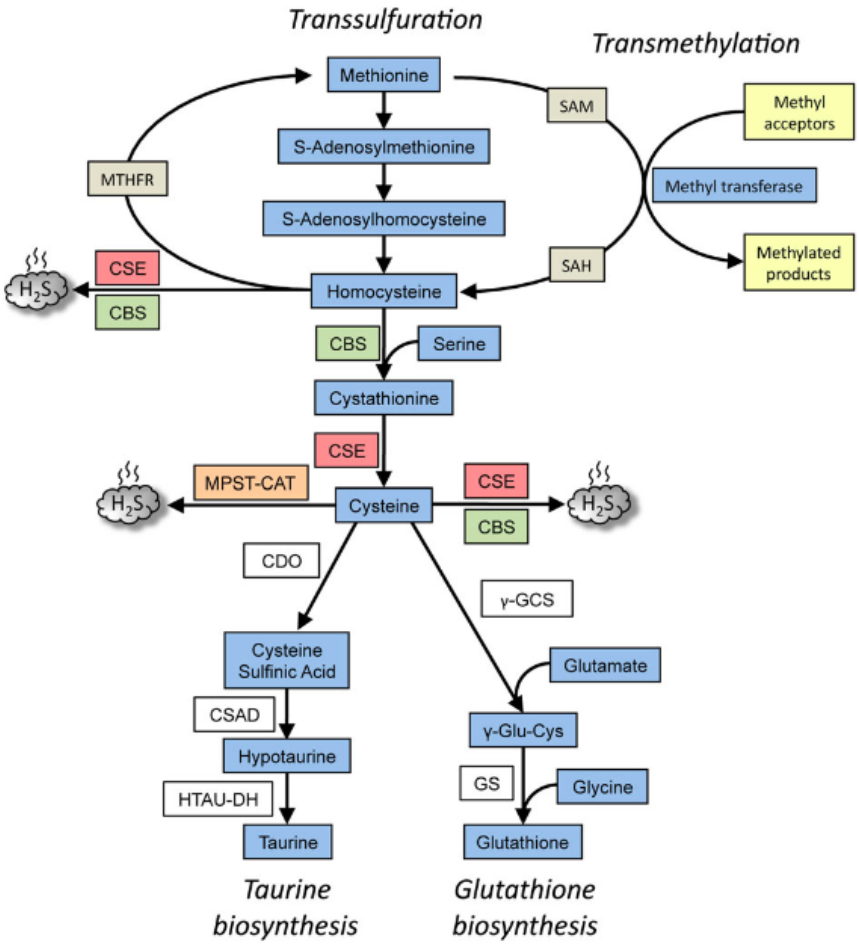


Figure 4. Overview of the transsulfuration pathway[47]. CTH (CSE) enzyme produces H_2S from homocysteine, cystathionine, and cysteine in transsulfuration pathway.

2. AIM OF THE STUDY

The aim of this project is to gain knowledge about the transcriptional role of the BTB-ZF transcription factor PATZ1 protein and identify its transcriptional targets in human colon cancer cell line HCT116. First, for loss of function studies we used CRISPR/Cas9 genome engineering technique to generate stable knockout cell lines. DEGs analysis was performed following the completion of RNA-Sequencing. We compared our RNA-Seq data with publicly available CHIP-Seq data (GSE92195) in order to obtain the DEGs that are directly bound by the PATZ1 TF. Expression analysis of the DEGs was performed via qPCR and Western Blotting analysis. According to our results we chose Cth gene to focus on and for the first time in literature we showed that the human transcription factor PATZ1 binds to and negatively regulates Cth gene promoter in HCT116 human colon cancer cell line.

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Chemicals

All the chemicals used in this study are listed in Appendix A.

3.1.2. Equipment

All the equipment used in this study is listed in Appendix B.

3.1.3. Buffers and Solutions

Agarose Gel: For 100 ml 1% w/v gel, we dissolved 1 g of agarose powder in 100 ml 0.5X TBE buffer via heating. 0.01% (v/v) ethidium bromide was added to the solution.

Calcium Chloride (CaCl2) Solution: 60 mM CaCl2 (0,5M stock), 15% Glycerol, and 10 mM PIPES with the pH of 7.00 were mixed in distilled water up to 500 mL. We filter-sterilized the solution and stored at +4°C.

Blocking Buffer: For 20 ml, 1 g skim milk powder was dissolved in 20 ml of PBS.

FACS Buffer: For 500 ml 1X solution, we mixed 2.5 g bovine serum albumin (BSA) and 0.5 g sodium azide in 500 ml sterile 1X PBS and stored at -20°C.

HEPES-buffered saline (HBS): For 100 ml 2X solution 0.8 g NaCl, 0.027 g Na2HPO4.2H2O and 1.2 g HEPES were mixed in 90 ml of distilled water. pH was

adjusted as 7.05 by NaOH pellets and the solution was completed up to 100 ml via distilled water. We sterilized the buffer with filter and kept at -20°C.

1X PBS-T: 0.5 ml of Tween 20 was dissolved in 1 L of 1X PBS.

Polyethylenamine (PEI) Solution: For 1 mg/ml solution, 100 mg polyethylenamine powder was dissolved in 100 ml distilled water by heating to 80°C and the pH was adjusted to 7.0 with HCl. The solution was filter sterilized, aliquoted and stored at -20°C.

Protein Loading Buffer: To prepare 4X protein loading buffer (Laemmli), 4 ml 100% glycerol, 2.4 ml Tris (1 M pH 6.8), 0.8 g SDS, 0.01% bromophenol blue, and 2 ml β- mercaptoethanol were mixed and then the volume was completed to 10 ml.

SDS Separating Gel: To prepare 10 ml 10% separation gel, 3.34 ml Acrylamide: Bis-acrylamide (37.5:1), 2.5 ml Tris (1.5 M pH 8.8), 100 µl 10% (w/v) APS, 100 µl 10% (w/v) SDS, and 10 µl TEMED were mixed, and the volume was completed to 10 ml with water.

SDS Stacking Gel: To prepare 5 ml 4% stacking gel, 1.25 ml Tris (0.5M pH 6.8), 1 ml Acrylamide: Bis-acrylamide (37.5:1), 50µl 10% SDS (w/v), 15µl 10% APS (w/v), and 7.5µl TEMED were mixed and with distilled water the volume was completed to 5ml.

10X Tris-Glycine Buffer: 25 mM Tris Base and 200 mM Glycine were mixed, and the pH is adjusted to 8.3 with HCl in distilled water.

SDS Running Buffer: For 1X SDS running buffer, we mixed 100 ml Tris-Glycine with 5 ml 20% SDS solution and completed the volume to 1 L with 895 ml distilled water.

1X Transfer Buffer (1 L): 100 ml of 10X Tris Glicine (pH 8.3), 200 ml methanol and 1.88 ml of SDS 20% mixed and volume completed to 1 L with distilled water. Stored at 4°C.

Enhanced Chemiluminescence (ECL) Reaction Buffer: 25 µl Luminol, 12.5 µl Cumaric acid, 234 µl Tris 1.5 and 1.5 µl of H2O2 30% were mixed in 4.728 ml of distilled water.

3.1.4. Growth Media

Luria Broth (LB): To make 1 L of 1X LB medium, 20 g LB powder was mixed with ddH2O to make 1 L and sterilized via autoclave for 15 min at 121°C. Just before using

the liquid medium, for antibiotic selection, kanamycin at a final concentration of 50 g/ml or ampicillin at a final concentration of 100 g/ml, were added.

LB-Agar: For LB-Agar bacterial medium, 15 g of Agar powder was mixed with 20 g of LB powder in 1 L of distilled water and autoclaved at 120°C for 15 min. Antibiotics were added after the liquid medium was cool but not solidified. The medium was poured into sterile Petri dishes which were stored at 4°C until use.

DMEM: HEK293T, HEK293FT, HCT116, U87MG, cell lines were cultured in DMEM growth medium which includes 10% heat-inactivated fetal bovine serum (FBS) and 1% Pen-Strep (100 U/ml Penicillium and 100 μg/mL Streptomycin).

Freezing medium: To freeze cell lines we used heat-inactivated fetal bovine serum containing 10% DMSO (v/v).

3.1.5. Molecular Biology Kits

All the commercial molecular biology kits used in this study are given in Appendix C.

3.1.6. Enzymes

All the restriction enzymes, DNA modifying enzymes and polymerases and their working buffers used in this thesis were from either Fermentas or NEB.

3.1.7. Bacterial Strains

Escherichia coli DH-5α We used (F- endA1 glnV44 thi-1 relA1 gyrA96 deoR nupG lacZdeltaM15 hsdR17) competent cells for bacterial transformation of general plasmid DNAs.

3.1.8. Mammalian Cell Lines

HEK293T: Geneticly engineered human embryonic kidney 293 (HEK293) cell line which stably expresses the large T antigen of SV40 virus (ATCC: CRL-1573™).

HEK293FT: The 293FT Cell line is generated for high-titer lentivirus production using the ViraPower™ Lentiviral Expression System. Presence of the SV40 large T antigen allows very high levels of protein expression.

HCT116: Human colorectal carcinoma cell line (ATCC CCL-247™)

U87MG: Derived from malignant gliomas (see also ATCC® HTB-15™ and HTB-16™) by J. Ponten and associates from 1966 to 1969.

3.1.9. Plasmids and Primers

The plasmids and the primers used in this study are listed in Table 3-1 and Table 3-2, respectively.

PLASMID NAME	**PURPOSE OF USE**	**SOURCE**
PCMV-HA	Cloning	Clontech
pCMV-HA-PATZ1	Mammalian Expression	Lab Construct
pCMV-HA-PATZ1Alt	Mammalian Expression	Lab Construct
pcDNA-GFP	Mammalian Expression	Lab Construct
pspCas9(2A) BB Puro	Mammalian expression plasmid for the CRISPR/Cas9 system with GFP expression gene	Addgene (#48138)
pLentiCRISPR v2	Lentiviral construct for CRISPR/Cas9	Addgene (#52961)

	expression with Puromycin resistance gene	
pLegoIG2PURO	Lentiviral Expression	A kind gift from Prof. Boris Fehse (University Medical CenterHamburgEppendorf, Hamburg, Germany)
pLegoIG2PURO-hPATZ1 v1	Lentiviral Expression	Lab Construct
pLegoIG2PURO-hPATZ1 v2	Lentiviral Expression	Lab Construct
pLegoIG2PURO-hPATZ1 v3	Lentiviral Expression	Lab Construct
pLegoIT2PURO	Lentiviral Expression	A kind gift from Prof. Boris Fehse (University Medical CenterHamburgEppendorf, Hamburg, Germany)
pGL14.4	Luciferase Plasmid	Promega
pMDLg/pRRE	Virus production/packaging plasmid (Gag/Pol)	Addgene (#12251)
pRSV-REV	Virus production/packaging plasmid (Rev)	Addgene (#12253)
phCMV-VSV-G	Virus production/packaging plasmid (Env)	Addgene (#8454)
pGL14.4-CTH Promoter	Luciferase Assay	Lab Construct

Table 3-1. The list of the plasmids used in this project. Plasmid names, their purpose of uses and sources are given.

PRIMER NAME	SEQUENCE	PURPOSE OF USE
Trnp1 Forward	GAGGCTGAGCTCCTTCCAACT	qPCR
Trnp1 Reverse	GCGCTGGGATTGTGTAAGGT	qPCR
AmacR Forward	GCTGATCAAAGGACTTGGACTAAAG	qPCR
AmacR Reverse	CCTTCGTCTCTGCAAATACATC	qPCR
Nkx2-5 Forward	CAAGTGTGCGTCTGCCTTTC	qPCR
Nkx2-5 Reverse	CGCGCACAGCTCTTTCTTTT	qPCR
Acan Forward	GACACCCCATGCAATTTGAGAAC	qPCR
Acan Reverse	GTAATTGCAGGGAACATCATTCCA	qPCR
SLC1A5 Forward	TGGACTGGCTAGTCGACCG	qPCR
SLC1A5 Reverse	CTGTGCTTCTCGACTCCGT	qPCR
Artn Forward	GCAGACTGGACCCTTACCG	qPCR
Artn Reverse	CTGGGGCTGCTAGTCAGTTG	qPCR
Tspan13 Forward	ACAGAGATCCTGGGTGTTTGG	qPCR
Tspan13 Reverse	CCTTCCTTAAACTGGCAAATGGA	qPCR
GAPDH Forward	TCCTGCACCACCAACTG	qPCR
GAPDH Reverse	TCTGGGTGGCAGTGATG	qPCR
MEX3B Forward	ACCCAGTTCTGAGCATGTCG	qPCR
MEX3B Reverse	TGTTCTTATTCCGGGAGGCG	qPCR
NR3C2 Forward	TTGCCTTGAGCTGGAGATCG	qPCR
NR3C2 Reverse	GTGCATCCCCTGGCATAGTT	qPCR
MIR1915HG Frw.	CACCTTTCCCCTACTACCGC	qPCR
MIR1915HG Rev.	GGGTGGGGACGCCTCT	qPCR
NR2F1 Forward	TTTCTTCAAGAGGAGCGTCCG	qPCR
NR2F1 Reverse	TTGGAGGCATTCTTCCTCGC	qPCR
HSP90AA1 Frw.	AGCACGTCGTGGGACAAATA	qPCR

HSP90AA1 Rev.	CTGGACCAGAAATCCCGACG	qPCR
EFNA5 Forward	TCGTGTTTTCGATGTTAACGACA	qPCR
EFNA5 Reverse	GGCTGACTCATGTACGGTGT	qPCR
ETV1 Forward	TCACTTCAGCTCTGGCAGTT	qPCR
ETV1 Reverse	GCGGAGTGAACGGCTAAGT	qPCR
FGF2 Forward	GCTGTACTGCAAAAACGGGG	qPCR
FGF2 Reverse	AGCCAGGTAACGGTTAGCAC	qPCR
BHLHE41 Frw.	GCAGTTGAACATGGACGAAGG	qPCR
BHLHE41 Rev.	AGGTATCCTTGGTGTCGTCTC	qPCR
PDLIM7 Forward	GCCGATCCCAACGAGGC	qPCR
PDLIM7 Reverse	GATGTGTGTGAGGCTACCCG	qPCR
RAB27B Forward	CTCGGGAACTGGCTGACAAA	qPCR
RAB27B Reverse	CACACACTGTTCCATTCGCT	qPCR
TGFB2 Forward	AGTGGGAAGACCCCACATCT	qPCR
TGFB2 Reverse	AAAGTGGACGTAGGCAGCAA	qPCR
TMEM200A Frw.	GGGGTCCCTCCACTCTAACT	qPCR
TMEM200A Rev.	TTGGGGCACAAGCAACCTAT	qPCR
VSNL1 Forward	TTAGTAGCACTGAGACCCCCA	qPCR
VSNL1 Reverse	TTCTGTGGTGTGTGGACCAG	qPCR
CTH Forward	GCACTCGGGTTTTGAATATAG	qPCR
CTH Reverse	CAGATGCCACTTGCCTGAAG	qPCR
CTH Promoter Frw.	GCTAGCAGGGGGAGTTTCTCTCTGTACT	Cloning
CTH Promoter Rev.	AAGCTTCCCTGGAGGTCCATTGCTC	Cloning
hPatz1-EcoRI Frw	GAATTCATGGAGCGGGTGAACGACGC	Cloning
hPatz1-NotI Rev	GCGGCCGCTCATTTCCCTTCAGGCCCCAT	Cloning
hU6_forward	GAGGGCCTATTTCCCATGATTC C	Sequencing
hPatz1 e2f	TCAAGCAGGTGCACACTTCTGAG	qPCR
hPatz1 e3r	TGGGACGACCTCCACAAAGC	qPCR

hPatz1 e4f	GGCCCAGCAACTTCTGCAGTATC	qPCR
hPatz1 e6r	TGAGAGGTCACCATAGGAGTCAGAG	qPCR

Table 3-2 . The list of the primers used in this project. Primer names, their sequences, and their purpose of uses are given.

3.1.10. DNA and Protein Molecular Weight Markers

DNA ladders and protein molecular weight markers are listed in Appendix C.

3.1.11. DNA Sequencing

Sequencing service was provided by McLab, CA, USA (http://www.mclab.com/).

3.1.12. Software, Computer-Based Programs, and Websites

The software and computer-based programs used in this project are listed in Table 3-3.

SOFTWARE, PROGRAM, WEBSITE NAME	**COMPANY/WEBSITE ADDRESS**	**PURPOSE OF USE**
NCBI BLAST	http://blast.ncbi.nlm.nih.gov/Blast.cgi	Basic local alignment search tool
FlowJo V10	Tree Star Inc.	Viewing and analyzing flow cytometry data
ImageJ	http://imagej.nih.gov/ij/	Counting IPSC colonies

CLC Genomic Workbench	CLC bio	Constructing vector maps, restriction analysis, DNA sequencing analysis, DNA alignments, etc
Ensembl Genome Browser	http://www.ensembl.org/index.html	Human and mouse genome
LightCycler 480 SW 1.5	ROCHE	Analysing qPCR results
Biorender	https://biorender.com/	Drawing Figures

Table 3-3. The list of the software, programs and websites. Name of the software, programs, and websites, their producers/websites and purpose of uses are given.

3.2. Methods

3.2.1. Bacterial Cell Culture

Bacterial Culture Growth: We grow Escherichia coli (E. coli) DH5α strain in Luria Broth (LB) at 37 ºC overnight, at 221 rpm on shaker. For the preparation of glycerol stocks, we added 10% (v/v) glycerol to the bacterial culture to get 1 ml final volume in cryovials and stored them at -80 °C. To obtain single colonies, we spread the bacterial culture on LB/agar containing petri dishes and incubated overnight at 37 ºC. All growth media were supplemented with or without selective antibiotic prior to any application.

Competent Bacteria: First we inoculated previously prepared competent E.coli DH5α in 40 ml of LB in a 250 ml flask without antibiotics and incubated at 37 ºC overnight, shaking at 221 rpm. The next day, we took 4 mL of this overnight culture into 400 mL fresh LB medium in a 2 L flask and incubated at 37 ºC overnight, shaking at 221 rpm until it reached the optical density (at 590 nm) 0.375. Next, 400 ml culture was divided into 50 ml tubes which were sterile on ice and incubated for 10 min followed by

a centrifugation at 1600 g for 10 min at 4 ºC. After centrifugation, we resuspended each pellet in 10 mL ice-cold CaCl2 solution and centrifuged at 1100 g for 5 min at 4 ºC. The pellets were again resuspended in 10 mL of ice-cold CaCl2 solution and incubated on ice for 30 min. After the final centrifugation at 1100 g for 10 min at 4 ºC, we mixed the suspensions in one 50 ml tube and aliquoted as 200 µl volume for each ice-cold microcentrifuge tube. As last step we froze aliquots in liquid nitrogen and stored at -80 °C. We calculated the transformation efficiency of cells via transforming different concentrations of pUC19 plasmid (Efficiency should be between 107-108 cfu/µg).

Transformation of Competent Bacteria: We used the Heat-shock transformation method in our transformation experiments. Frozen chemically competent bacteria were thawed on ice until they became viscous. After that, competent bacteria were given the needed concentration of DNA (for pure plasmid transformation, 1 pg-1 ng; for ligated product transformation, 5-20 µl ligation mix). For 30 minutes, the mixture was kept on ice. The cells were then heat-shocked for 90 sec at 42 °C before being put on ice for 1 minute. The heat-shocked cells were inoculated with 800 ml LB (without antibiotics) and the final culture was incubated at 37 °C for 45 min. We centrifuged the cells at 13,200 rpm for 30 seconds after this incubation and resuspended the bacterial pellet in 100 ml of LB remaining from the supernatant. Finally, suspension was placed over an LB-agar plate with the appropriate antibiotic for selection, and the plates were incubated overnight at 37 °C.

Plasmid DNA Isolation: Plasmid DNA isolation was performed using either the alkaline lysis protocol from Molecular Cloning: A Laboratory Manual (Sambrook et al) or Macherey Nagel Midiprep Kit according to the manufacturer protocols. The concentration and purity of the DNA isolated were determined by using a NanoDrop-spectrophotometer.

3.2.2. Plasmid Construction

Polymerase Chain Reaction (PCR): You can find the optimized PCR conditions in Table 3.4. The thermal cycler conditions are: initial denaturation step at 95 ºC for 5 min which is followed by the denaturation step (at 95 ºC, for 30 seconds) with 30 (or 35) cycles, annealing step (specific temperature, for 30 seconds) and the extension step (at 72 ºC, for 1 min for every 1 kilobase of DNA). These cycles were then followed by a final extension step at 72 ºC for 10 min.

PCR Reaction	**Volume Used**
Template DNA	1-10 µL
Q5 or Taq Polymerase Buffer (with $MgCl_2$)	2.5 µL
dNTP mix (10 mM each)	0.5 µL
Forward Primer (10 µM)	2 µL
Reverse Primer (10 µM)	2 µL
Q5 or Taq Polymerase (2.5U/µL)	0.125 µL
ddH_2O	Up to 25µL
Total	25µL

Table 3-4. Optimized PCR conditions

Restriction Enzyme Digestion: We performed the restriction digestions via mixing the required amount of DNA with the appropriate enzymes and their working buffers in 1.5-2 mL centrifuge tube, followed by incubation at a temperature optimum for the enzymes to work for a duration of 1,5-2 hours. For diagnostic digestions 1 µg of DNA was used. Generally, we used 10 µg or more DNA for gel extraction and cloning. If we performed the digestion of the plasmid DNA with a single restriction enzyme then the linear plasmid was dephosphorylated by shrimp alkaline phosphatase (rSAP) enzyme for an additional 30-40 min at 37 ºC for ligation reaction.

Agarose Gel Electrophoresis: We loaded the digested sequences on 1% Agarose gel and ran until band separation were clearly visible. 20 µl of each sample mixed with loading dye were loaded to each well. 1,5 µl of a DNA ladder was also loaded as reference. All DNA molecules which are negatively charged run by electrophoresis in the direction of negative to positive. We used 100-110 volts for 45-60 min. Once the run was complete the gels were visualized under UV light and desired DNA bands were cut from the gel and extracted using a Qiagen Gel Extraction Kit according to the manufacturer protocol.

Ligation: We performed ligation reactions via T4 DNA Ligase, in 1:3 vector to insert ratio, using 100 ng of the plasmid DNA. Ligation mixtures were incubated at 16 ºC overnight. The next day, we transformed the mixture into competent DH5α bacteria that we prepared and spread them onto LB/agar petri dishes with desired antibiotics and incubated at 37 ºC.

3.2.3. Mammalian Cell Culture

Maintenance of Cell Lines: HCT116, HEK293T, HEK293FT, U87MG, cell lines were cultured in DMEM medium in sterile tissue culture dishes in an incubator set to 37 ºC with 5% CO_2. When the confluency of the cells reached over 80%, we split the cells into warm, fresh medium. In order to freeze cells, we resuspended the cells at the exponential growth phase in freezing medium and transferred to cryovials. Tubes were stored at -80 °C for at least 24 h in isopropanol carrying boxes and then transferred to liquid nitrogen tank for long- term storage. After thawing the cells, we washed the cells with appropriate growth medium to remove any residual DMSO which can be highly toxic.

Transient Transfection: One day prior to the transfection, we split 3 x 10^5 cells into 6-well plates. On the day of the transfection, 3-5 µg of plasmid DNA was mixed in 200 µL serum-free DMEM in a sterile 1,5 ml eppendorf tube. PEI (1µg/µL) was added to the DNA-DMEM mix at a 3:1 ratio of PEI (µg) to total plasmid DNA (µg) and mixed

immediately by vortexing. After 20 min of incubation at room temperature, the mixture was added drop wise onto the cells.

Production of Lentivirus: To produce lentiviral particles carrying desired DNA sequences, we used HEK293FT cell line. Transient transfection of this cell line is performed by using $CaCl_2$ transfection method. 3rd generation lentiviral plasmids are used as seen in the plasmids list. After the transfection of HEK293FT cells with plasmids at 24 h and 36 h supernatants carrying viral particles collected and stored at -80°C for long term usage.

Lentiviral Infection: To determine infectious particles per ml, $0.5x10^5$ HCT116 cells/well was seeded into 24-well plates and incubated at 37 °C, 5% CO2 for 5-6 h. Serial dilutions of viral supernatant were added to each well with additional proteamine sulfate (8 µg/ml). Culture media was changed with DMEM following 16 h of incubation. Flow cytometry analysis was performed after 48 h of incubation. Multiplicity of infection (MOI) calculations were done according to the percentages of fluorochrome expressing cell populations.

Lentiviral transductions of HCT116 and U87MG cells were carried out in 6-well plates and 10 cm tissue culture dishes. For 10 cm dishes $4,5.10^6$ cells and for 6-well plates 5.10^5 cells were used in the presence of 8 µg/ml proteamine sulfate. Proper MOI was decided according to the previous description. After addition of the viral supernatants to the plates, cells were incubated for overnight (16 hours) at 37 °C, 5% CO2. After 16 h of incubations done, we changed the media with fresh ones and incubated the cells for 72 h more. For the selection of positive cell populations 1 mg/ml puromycin was used for additional 4 days and selected cells were expanded or FACS performed.

Flow Cytometry and Intracellular Staining: Staining solution including 0.03% saponin, %1 BSA, 0.1% NaN_3, 10% 1x PBS completed with 459 ml of dh_2O up to 500 ml of solution is prepared. Before starting, cells were collected, counted, and washed two times with 1 ml staining solution. We added 100 ul of staining solution onto each sample and incubated for 15 min at 4 °C. After incubation centrifuged at 300 g for 5 min and supernatant discarded. 30 ul of antibody diluted in staining solution in appropriate concentration added on to the cells and incubated for 30 min at 4 °C. Washed three times

with 1 ml staining solution. 30 ul of secondary antibody diluted in staining solution in appropriate concentration added on to the cells and incubated for 30 min at 4 °C. Washed three times, centrifuged and supernatant discarded. 1 ml of FACS buffer or 1x PBS added for Flow cytometry analysis on proper channels.

SDS-PAGE and Western-Blot: For cell lysis, we harvested the cells using trypsin. Cells were centrifuged at 300 g for 5 min and the supernatant was discarded. The cells were washed once with 1X PBS and centrifuged again. Next, we resuspended the pellet in an appropriate amount of protein loading buffer and the mixture was boiled at 95 ºC for 10 min. The lysates were used immediately or stored at -20 ºC for later use. We used 8-10% separating part and a 4-6 % stacking part for the gels. After the samples were loaded into the gels, the gels in 1X running buffer were run at 100 V (constant voltage) for 1.5 h using a BIORAD Mini Protean Tetra Cell. After the run was complete, the gels were transferred to 0.45 µm PVDF membranes (which were activated by using methanol) in 1X transfer buffer at constant current of 250 mA for 1.5 h at 4 ºC using BIORAD Mini Trans-blot. Next, we blocked the membranes using blocking buffer made of 10 mL PBST – skim milk powder at room temperature for 1 hour on a shaker. Primary antibody incubations were done overnight at 4 ºC or 2 h at room temperature and secondary antibody incubations were done for 1 hour at room temperature. After every incubation, we washed the membranes with 15 ml of PBST 3 times for 10 min each. After the final washing step, the membranes were incubated with lab made ECL for 1-5 min at room temperature in the dark room and exposed to X-Ray films (Fuji).

3.2.4. RNA-Sequencing

RNA-Sequencing: We prepared PATZ1 knockout and WT HCT116 cells for RNA isolation. Total RNA was extracted, polyA selected and triplicate samples were sequenced at Beijing Genomics Institute via Illumina HiSeq 2000 sequencing machine. We performed the analysis using CLC Genomics Workbench. To obtain differentially expressed genes (DEGs) we used following criteria: fold-change > 2 and FDR value < 0.005. In order to identify biological processes regulated by the PATZ1 transcription factor, we performed DAVID analysis for gene ontology (GO) terms. P-values were

calculated by DAVID and GO terms and the ones with P-values smaller than 0.01 were listed as significant.

4. RESULTS

4.1. Generation of PATZ1 Knockout HCT116 Human Colon Cancer Cell Line by Transient Transfection of CRISPR/Cas9 Genome Engineering Tool

To generate a PATZ1 knockout HCT116 cell line we decided to use the CRISPR/Cas9 system which is a two-component efficient genome editing tool. The first component is the Cas9 protein which has two endonuclease domains called RuvC and HNH that cleave opposite DNA strands and as a result create double strand breaks in the target DNA locus (DSBs). The second component is a single guide RNA carrying a scaffold sequence and a 20-base pair spacer sequence complementary to the target DNA and adjacent to the PAM sequence. This sgRNA guides the complex to the target locus to create a DSB that will be followed by either non-homologous end-joining (NHEJ) or homology-directed repair (HDR) depending on the availability of exogenous donor templates. In our case we chose to continue with NHEJ to create Indels in the target genomic location.

The chosen target sequence as gRNA preceding a 5'-NGG PAM sequence including a cut site which was in the Exon 1 of human PATZ1 gene was cloned into a CRISPR plasmid pspCas9(BB)2A-PURO and ligation was approved by sequencing (Fig. 5).

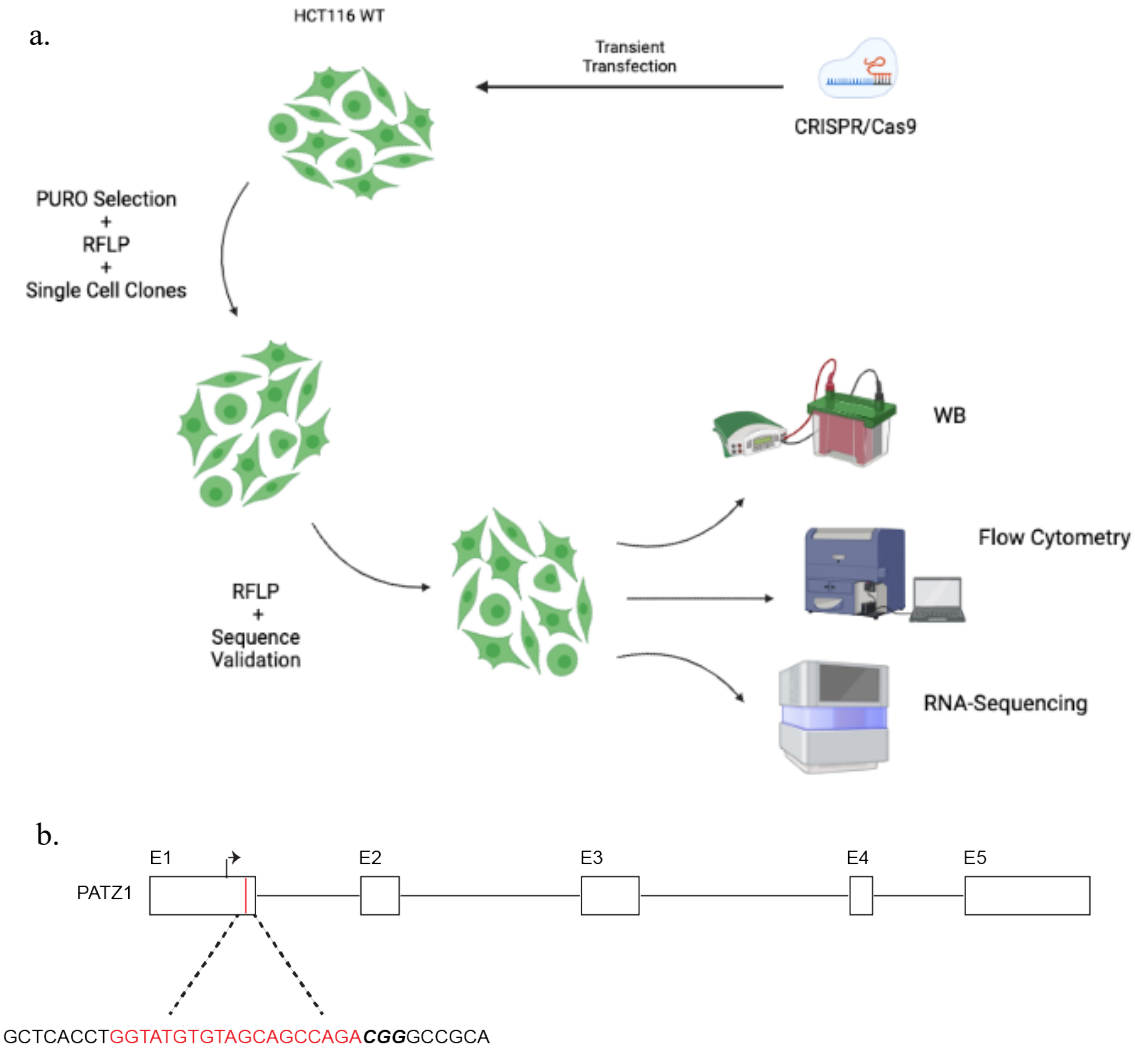


Figure 5. Experimental strategy to target the Patz1 gene. a. Experimental design of generation of the PATZ1 knockout HCT116 cell line via CRISPR/Cas9 system. b. CRISPR/Cas9 gRNA sequence targeting Patz1 gene Exon 1

To validate knockout cell pools, first we decided to perform RFLP (Restriction fragment length polymorphism) and designed a primer pair that gives a PCR product including target site with a diagnostic digestion option by BglI enzyme (Fig 6a). WT HCT116 human colon cancer cells were transiently transfected with generated CRISPR plasmid and after 72h of puromycin selection at 37 °C, genomic DNAs were collected. Using the collected genomic DNA, we performed PCR and performed diagnostic digestion using BglI enzyme. RFLP results controlled for mutant DNA bands (Fig.6b).

Pool-1 with the mutant DNA bands selected to grow single cell clones. After expansion period, single cell clones were scanned by using the same RFLP method to find a knockout cell line. Clone A12-2 with two different mutant bands was chosen for further analysis. (Fig.6c).

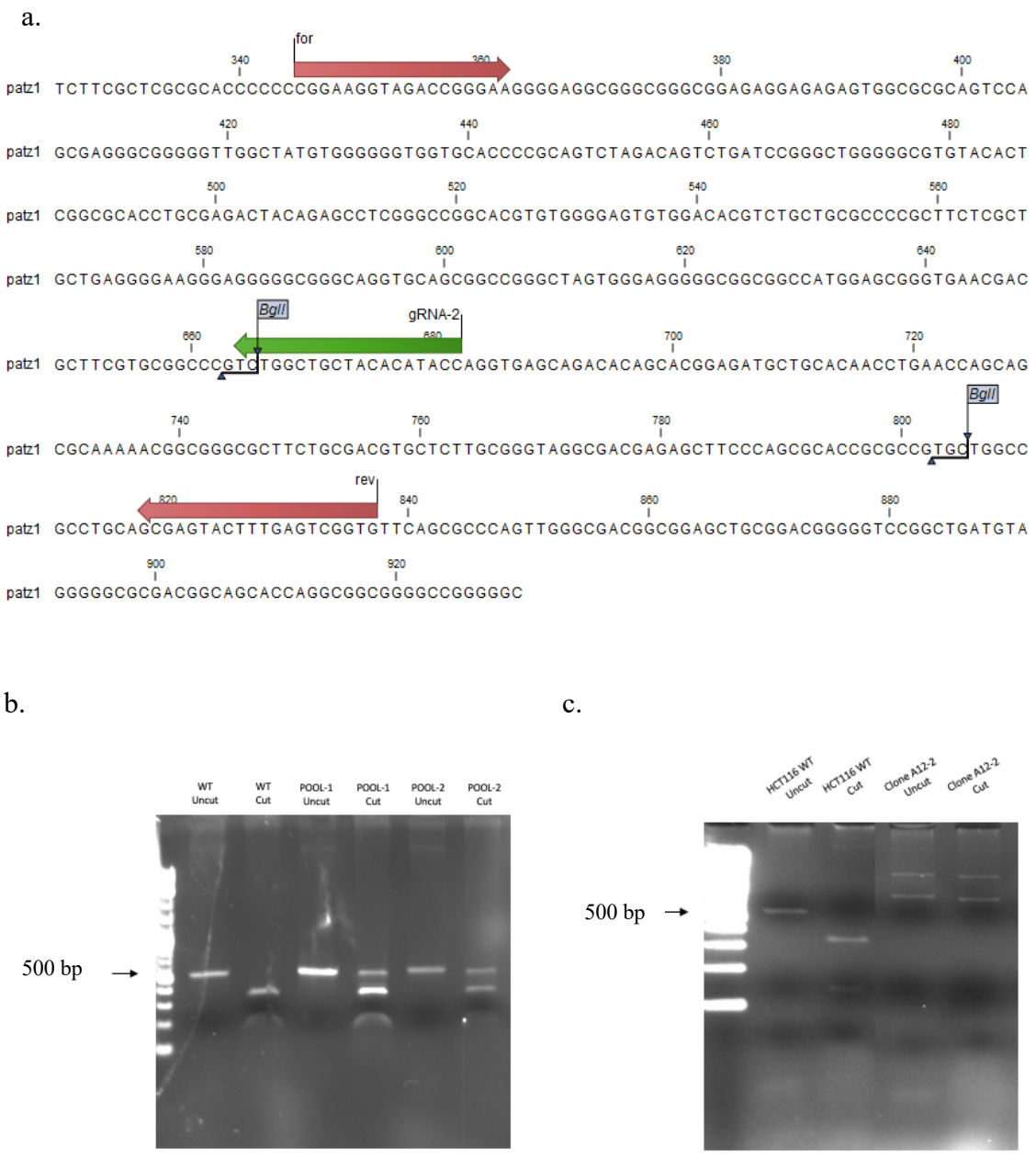
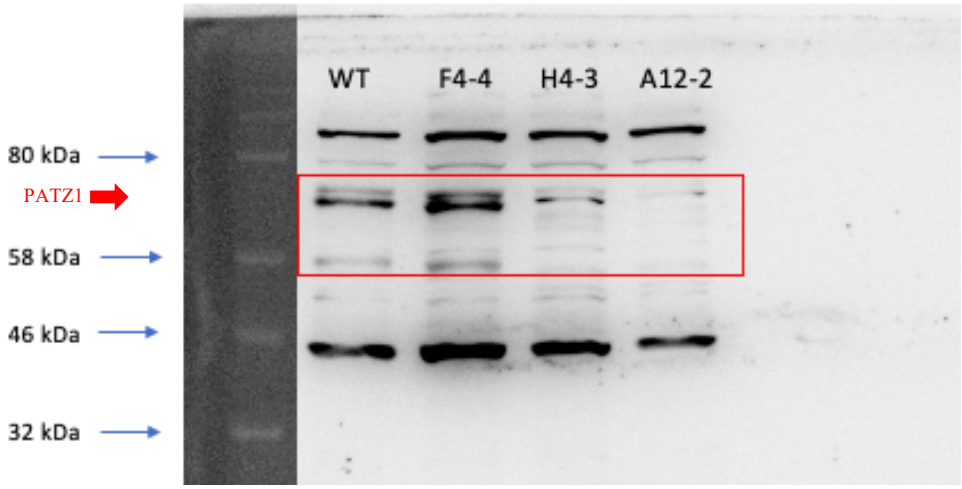


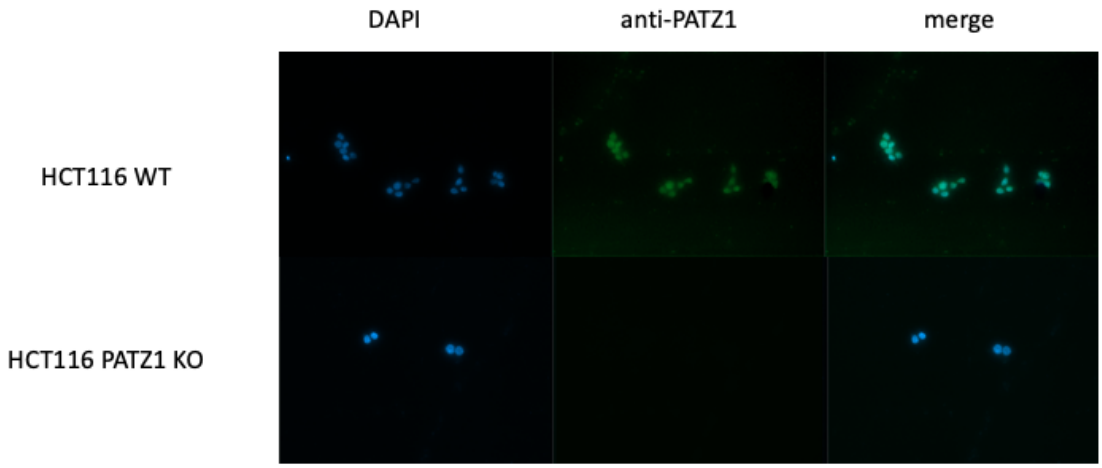
Figure 6. Confirmation of the gene editing a. Binding sites of PCR oligos and the cut sites of BglI enzyme b. RFLP results of WT HCT116 cells and HCT116 CRISPR pools. c. RFLP results of WT HCT116 cells and single cell clone A12-2.

To confirm the absence of the human PATZ1 protein in the selected single cell clones, first we prepared cell lysates from both WT and PATZ1 KO HCT116 cell lines which we used for SDS-PAGE followed by anti-PATZ1 blotting (Fig.7a). In addition to western blot, intracellular and immunofluorescence stainings of WT and PATZ1 KO HCT116 single cell clone A12-2 was performed. We used primary anti-PATZ1 mouse antibody (sc-30577, D-5) and secondary Alexafluor 488 anti mouse IgG Goat antibody (were for staining and we saw that the single cell clone A12-2 did not express the hPATZ1 protein (Fig.7b-7c). Next, we gel-purified PCR products by cutting the correct bands from agarose gel and sent these two different mutant DNA bands of clone A12-2 to sequencing and confirmed the mutations by sequence analysis (Fig.8a-b).

a.



b.



c.

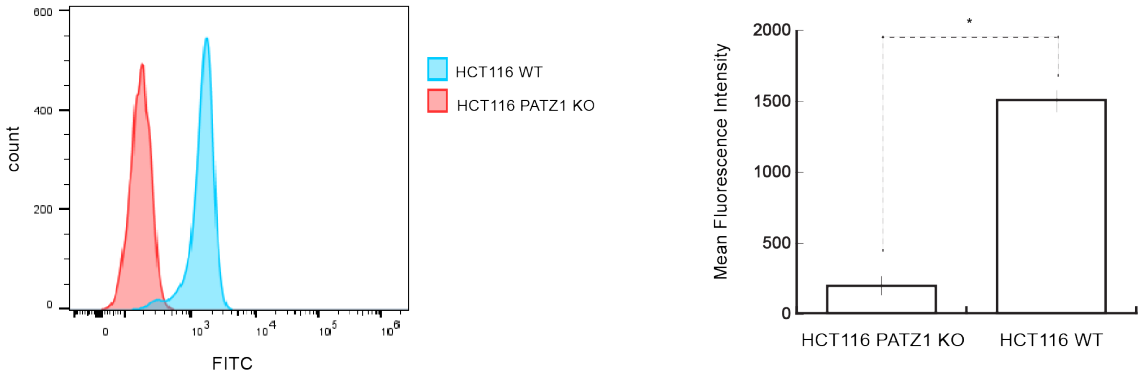
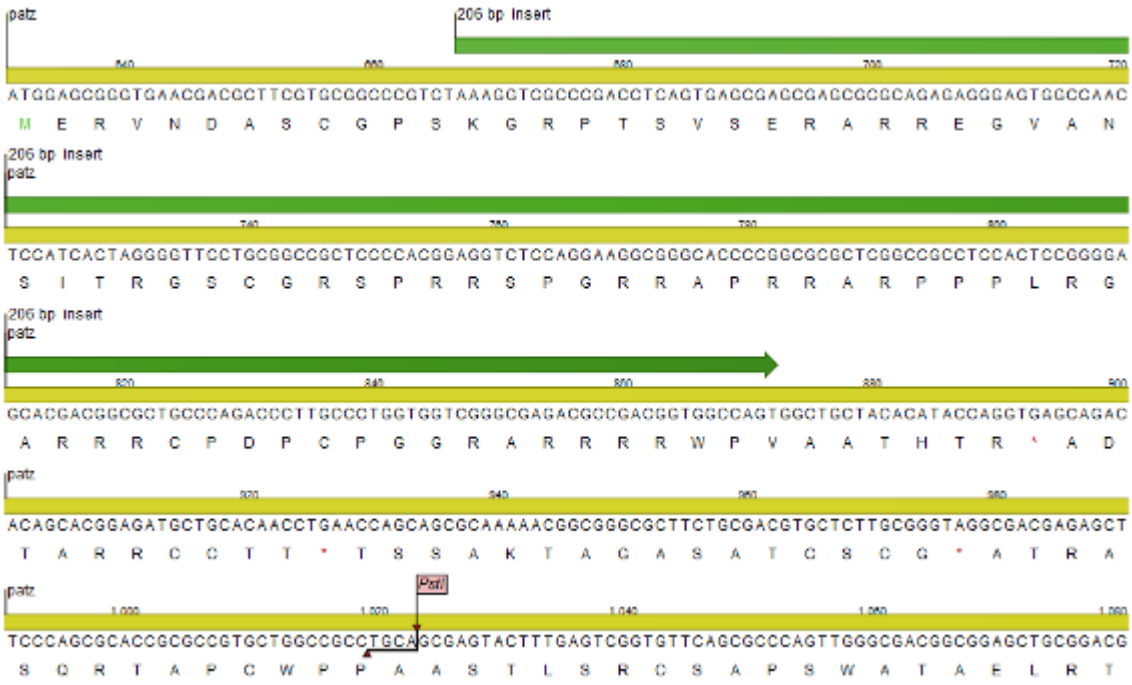


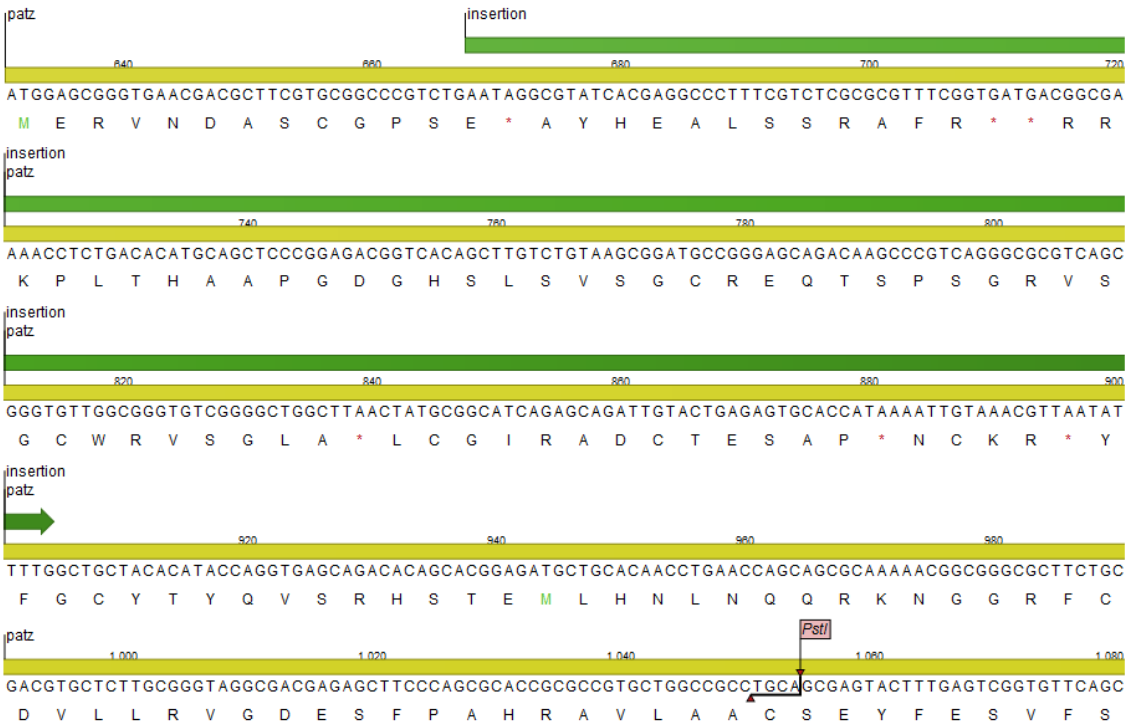
Figure 7. Validation of the absence of the PATZ1 protein. a. Anti-PATZ1 Western blotting results of WT control cells and single cell clones F4-4, H4-3 and A12-2 b. Immunofluorescence staining c. Intracellular Anti-PATZ1 staining of WT HCT116 cells and single cell clone A12-2.

a.



MERVNDASCGPSKGRPTSVSERARREGVANSITRGSCGRSPRRSPGRRAPRRARPPP
LRGARRRCPDPCPGGRARRRRWPVAATHTR*ADTARRCCTT*TSSAKTAGASATCS
CG*ATRASQRTAPCWPPAASTLSRCSAPSWATAELRTGVRLM*GARRQHQAAGPG
AAGSWRCTLSAPRYLGTFWTSPTLPASWCAWRAFPNS*RPPSSC**GRLSRSARKSS
NSPTYRSWYPLPAPI*CSFAPLGPRTWASLWT*PTGQPWQPTAMASPAACSQRRRQ
LGRLVQPLQAKPLCLCYLGWTACPWWLDPYPPNC*LPHSPVWHPVPLP*LASEAG
AAQGRPTCWTQCLGPQGA*GRQASFHAVYVVRCSLMPTGSGSTRPSTVSPASSWA
TSTFLLRGWVRMGYPSLKTPTAPERGAGPGSRWLVRSAARSSVMCIILTGTSCPTLG
RSPTPALCVGCGSREKTACPTMCGPMMGPWASLTSARAVGKASPGLIT*TDISSRCT
LLSGLTSVRPAMLLLPPETVCAPTWPVMKTRCPARCVGSTCGQHTWQTT*RSTAR
GPATSAVSVTEVSPLPPT*RSMLKPTTVFPFPRSPGTRSPS*MGEQRSTAPGPMATKK
ARNAHIRIRLRALTPMVTSQMPAT*RRQRSRVPMALSPATWQSPKTKWSLMGRRS
THALNVGASSALSPT*TNTSRRCMSGLSGAPWGTWALPLAHLSLLSRTCLSSSPLGF
RLFSRHLRHL**ILRLTSSPWGLKGN

b.



MERVNDASCGPSE*AYHEALSSRAFR**RRKPLTHAAPGDGHSLSVSGCREQTSPSG
RVSGCWRVSGLA*LCGIRADCTESAP*NCKR*YFGCYTYQVSRHSTEMLHNLNQQR
KNGGRFCDVLLRVGDESFPAHRAVLAACTHYFESVFSAQLGDGGAADGGPADVG
GATAAPGGGAGGSRELEMHTISSKVFGDILDFAYTSRIVVRLESFPELMTAAKFLLM
RSVIEICQEVIKQSNVQILVPPARADIMLFRPPGTSDLGFPLDMTNGAALAANSNGIA
GSMQPEEEAARAAGAAIAGQASLPVLPGVDRLPMVAGPLSPQLLTSPFPSVASSAPP
LTGKRGRGRPRKANLLDSMFGSPGGLREAGILPCGLCGKVFTDANRLRQHEAQHG
VTSLQLGYIDLPPPRLGENGLPISEDPDGPRKRSRTRKQVACEICGKIFRDVYHLNRH
KLSHSGEKPYSCPVCGLRFKRKDRMSYHVRSHDGSVGKPYICQSCGKGFSRPDHLN
GHIKQVHTSERPHKCQTCNASFATRDRLRSHLACHEDKVPCQVCGKYLRAAYMA
DHLKKHSEGPSNFCSICNRGFSSASYLKVHVKTHHGVPLPQVSRHQEPILNGGAAF
HCARTYGNKEGQKCSHQDPIESSDSYGDLSDASDLKTPEKQSANGSFSCDMAVPK
NKMESDGEKKYPCPECGSFFRSKSYLNKHIQKVHVRALGGPLGDLGPALGSPFSPQ
QNMSLLESFGFQIVQSAFASSLVDPEVDQQPMGPEGK*

Figure 8. Sequencing results of the single cell clone A12-2. a. Sequencing result of allele with the short insertion from single cell clone A12-2 and protein sequence coded by this sequence showing premature stop codons. b. Sequencing result of allele with the long insertion from single cell clone A12-2 and protein sequence coded by this sequence showing premature stop codons.

4.2. Generation of PATZ1 Knockout U87MG Human Glioblastoma Cell Line by Lentiviral Delivery of CRISPR/Cas9 Genome Engineering Tool

To generalize our findings in PATZ1 deficient HCT116 cells, we wanted to create Patz1 gene knockout cell lines from different developmental lineages. As the Patz1 gene is known to be important for the development of glioblastomas, where it can undergo translocations, we chose a common glioblastoma cell line, U87MG, which is not known to have a translocation involving the Patz1 gene but does express this gene. To create PATZ1 KO U87MG human glioblastoma cell line we used lentiviral CRISPR/Cas9 Delivery System (Fig. 9). We cloned the same gRNA which was selected to target exon

1 of the human Patz1 gene in HCT116 cells into the lentiviral CRISPR vector; pLentiCRISPR v2 to use in lentivirus production.

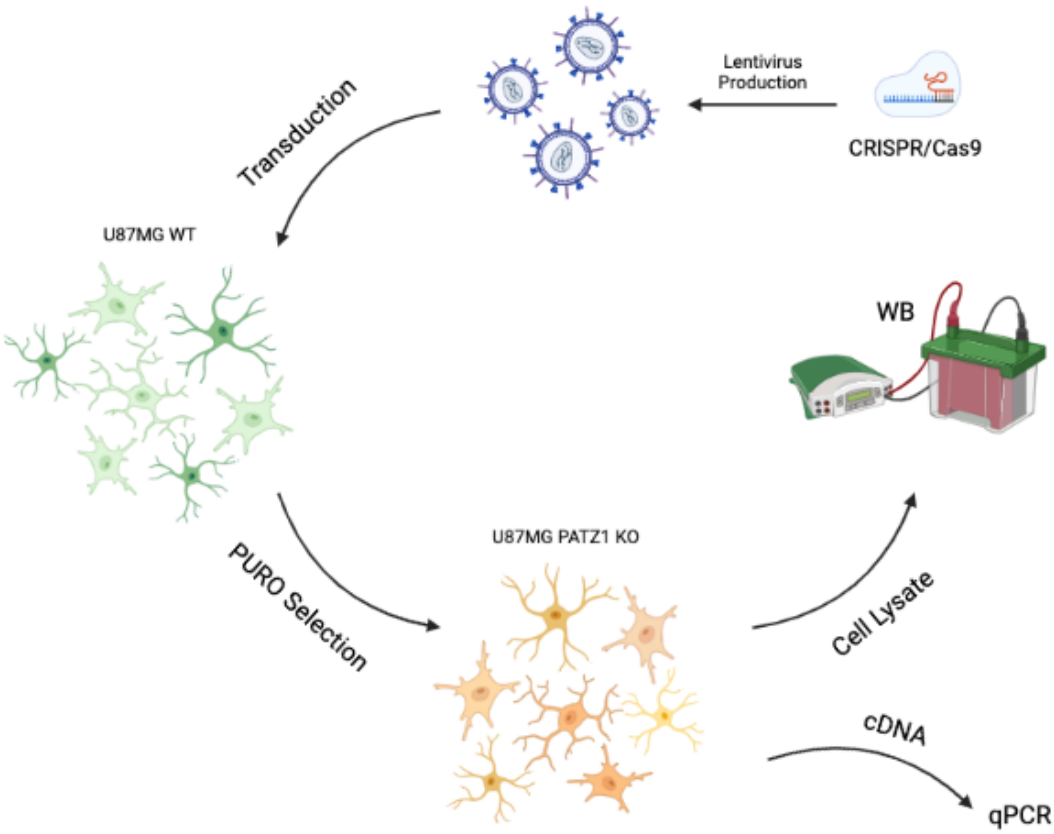


Figure 9. Experimental design of lentiviral CRISPR/Cas9 delivery to generate PATZ1 KO U87MG cell line.

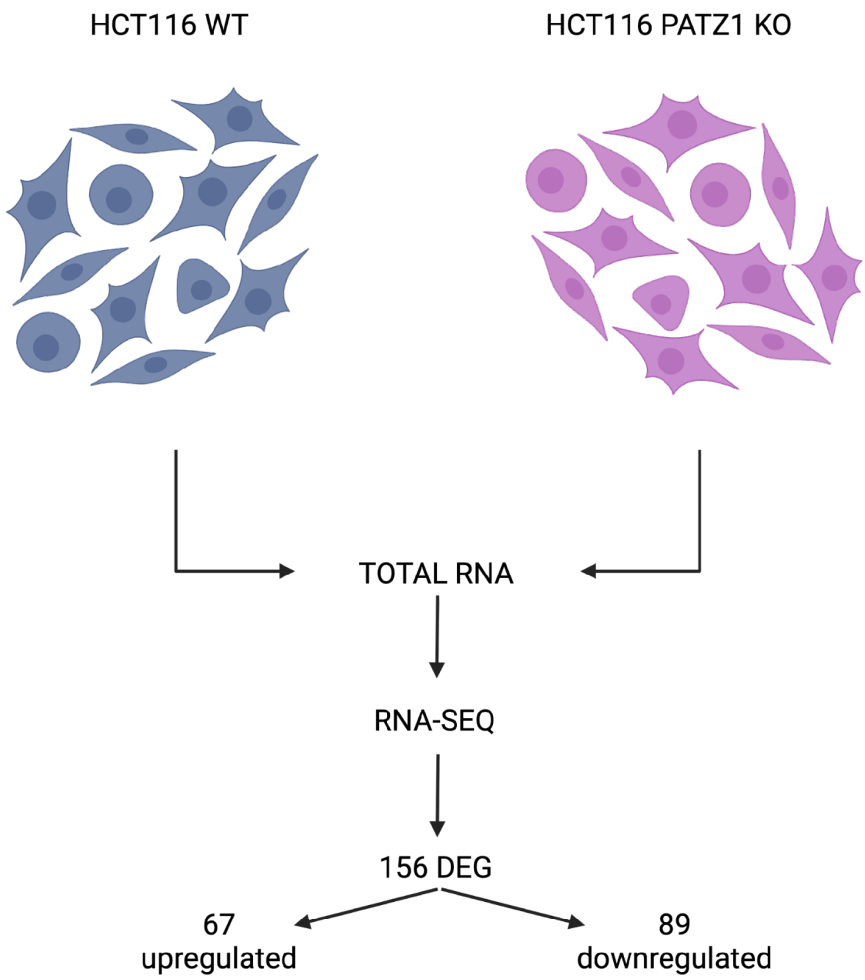
HEK293FT cells were transiently transfected with 3rd generation lentiviral packaging plasmids and plentiCRISPRv2 plasmid carrying our gRNA. 16 h after transfection media changed and fresh media added. After addition of fresh media at 24 h and 36 h points media containing lentiviral particles collected from the dishes and stored at -80 for longterm usage. Following the collection of viruses and titration experiment as mentioned in the methods we transduced the cells and selected with puromycin for 72 h at 37 °C. After the selection and expansion of puromycin resistant cells we prepared cell lysates and performed anti-Patz1 western blotting to check the intracellular PATZ1 protein levels (Fig. 10).

Figure 10 Anti-PATZ1 Western blotting of HCT116 PATZ1 KO, U87MG WT and U87MG PATZ1 KO cell Pool. PATZ1 shows the full-length protein isoform and PATZ1-Alt shows the short protein isoform.

4.3. RNA-Sequencing; WT vs. PATZ1 KO HCT116 Cells

To analyze the transcriptional expression difference and identify the differentially expressed genes between WT and PATZ1 knockout HCT116 cell lines we performed RNA-Seq analysis. Following total RNA isolation from WT and PATZ1 KO HCT116 cells using Trizol, RNA sequencing was performed on an Illumina HiSeq 2000 platform at the Beijing Genomics Institute. After the comparison of the expression values (RPKM) of genes in WT and PATZ1 KO HCT116 cells (KO vs WT) we identified 156 differentially expressed genes (DEGs) (Fig. 11a). 89 of these DEGs were downregulated and 67 were upregulated in the absence of PATZ1 protein (Appendix D). We used gene ontology (GO) analysis to get an overview of the biological processes that are highly enriched in the DEGs of our WT vs. PATZ1 KO HCT116 cells comparison and found out that in the absence of PATZ1 protein, biological processes like neurogenesis, cell adhesion, tissue development, neuron differentiation, regulation of muscle cell differentiation were affected (Fig. 11b).

a.


HCT116 WT
HCT116 PATZ1 KO
TOTAL RNA
RNA-SEQ
156 DEG
67
upregulated
89
downregulated

b.

GO Term Related to Biological Process	p-Value
neurogenesis (GO:0022008)	2.17E-07
cell adhesion (GO:0007155)	6.76E-07
animal organ morphogenesis (GO:0009887)	1.01E-06
chemotaxis (GO:0006935)	2.72E-06
tissue development (GO:0009888)	2.87E-06
anatomical structure morphogenesis (GO:0009653)	3.10E-06
regulation of muscle cell differentiation (GO:0051147)	5.55E-06
neuron differentiation (GO:0030182)	5.88E-06
axon guidance (GO:0007411)	6.00E-06
axonogenesis (GO:0007409)	6.13E-06
neuron projection guidance (GO:0097485)	6.22E-06
positive regulation of osteoblast differentiation (GO:0045669)	1.04E-05
nervous system development (GO:0007399)	1.32E-05
cranial skeletal system development (GO:1904888)	1.41E-05
generation of neurons (GO:0048699)	1.44E-05
cell differentiation (GO:0030154)	1.67E-05
axon development (GO:0061564)	1.71E-05
regulation of cell differentiation (GO:0045595)	1.92E-05

Figure 11. Experimental design of RNA sequencing a. Schematic representation of the experimental flow b. biological process GO terms that are enriched in our DEG set (WT vs PATZ1 KO HCT116 cells).

To find out the genes that were bound by the PATZ1 protein, we used CHIP-Seq data publicly available on NCBI-GEO (Gene Expression Omnibus) database [57] (GEO accession GSE92195). In this CHIP-seq experiment GFP fused the PATZ1 protein was overexpressed in HEK293 cells, crosslinked, and precipitated with CHIP grade anti-GFP antibody. After our analysis we found out 2263 genes that were bound by PATZ1 protein

in HEK293 cells. For further analysis we decided to select DEGs from our RNA-Seq data that were bound by PATZ1 protein in CHIP-Seq data and identified 15 genes common in between the datasets (Fig.12).

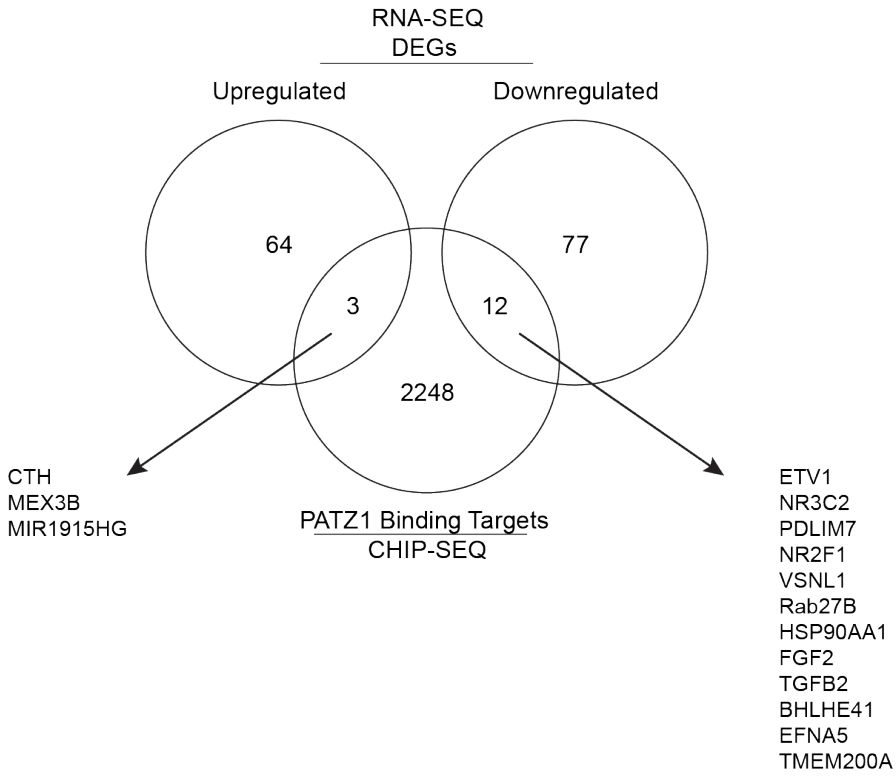


Figure 12. Common DEGs between RNA-Seq (HCT116) and CHIP-seq (HEK293) datasets showing the direct binding targets of hPATZ1 protein.

4.4. Expression Profiling of DEGs by qPCR

To compare the expression levels of 15 common DEGs that we selected from RNA-Seq and CHIP-Seq datasets, we designed specific qPCR primer pairs for each gene (table 3.2). Total RNA extracts were isolated from both WT and PATZ1 KO HCT116 cells by using TRIZOL and cDNA synthesis was performed by using reverse transcriptase. Using these cDNAs and primers we performed qPCR (Fig. 13).

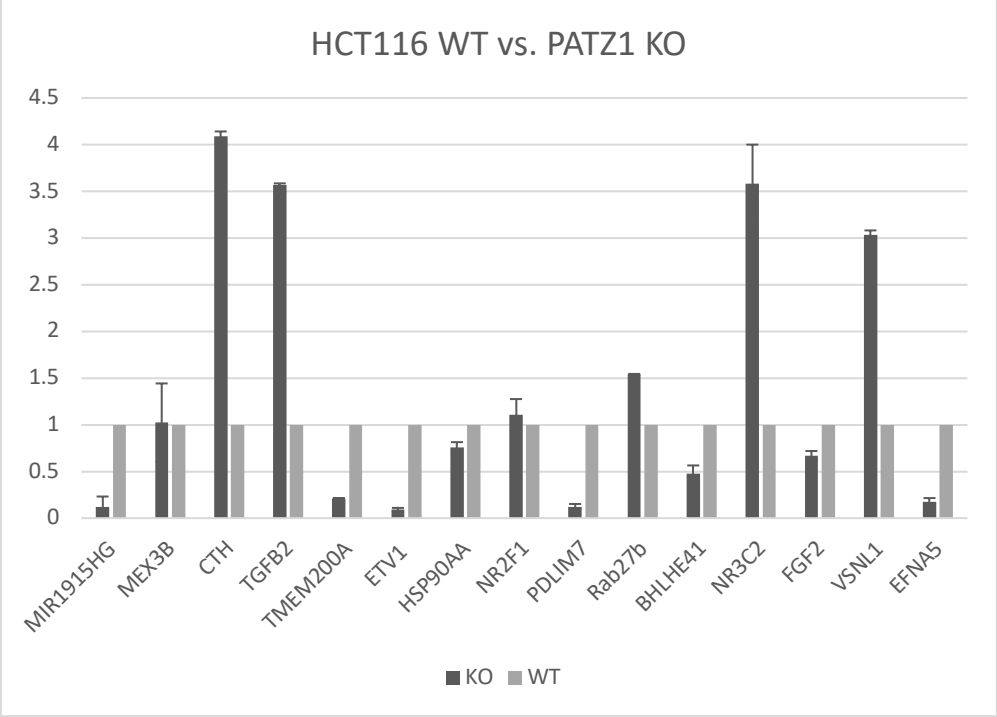


Figure 13. qPCR analysis of the DEGs directly targeted by PATZ1 protein in HCT116 cell line.

To determine if the gene expression controlled by PATZ1 was common among other cell types, we isolated total RNA and synthesized cDNAs from the WT and PATZ1 KO U87MG human glioblastoma cell lines that we generated by lentiviral CRISPR delivery system and performed qPCR (Fig.14).

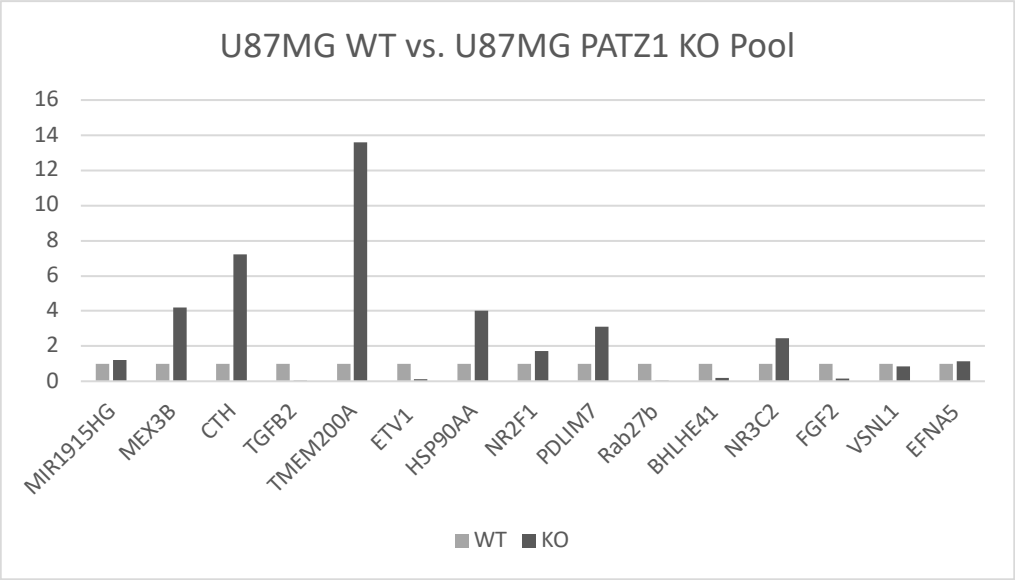


Figure 14. qPCR analysis of DEGs directly targeted by PATZ1 protein in U87MG cell line.

As can be seen in the qPCR results, expression level of the Cth gene was increased in the absence of the PATZ1 protein and CTH was one of the most upregulated DEG in both PATZ1 KO HCT116 and PATZ1 KO U87MG cell lines that likely is bound by and transcriptionally controlled by the PATZ1 transcription factor. To further characterize the mechanism of control of the Cth gene by PATZ1 and to elaborate these results we decided to focus on Cth gene for further detailed research.

4.5. Rescue Experiments

To further examine the effects of PATZ1 transcription factor on expression levels and confirm our hypothesis, we decided to perform transient and stable transfection of hPATZ1 cDNAs into WT and PATZ1 KO HCT116 cell lines. This would show us if expressing PATZ1 protein back in the knockout cells would recover the expressional phenotype caused by the absence of PATZ1 in cells.

4.5.1. Transient Transfection of PATZ1 cDNA into WT and PATZ1 KO HCT116 Cell Lines

First, to transiently express PATZ1 protein in HCT116 PATZ1 KO cells we performed transient transfection of PATZ1 cDNA which was cloned into eukaryotic expression vector carrying an HA-tag by using PEI transfection. 24 h after transfection we collected total RNAs and synthetize cDNAs to perform qPCR. Alongside the Cth gene that we selected to focus on, we also used some of other target genes as control for the experiment. As you can see in the figure transient transfection of HA-PATZ1 cDNA into the PATZ1 KO HCT116 cells did not create an opposite effect or any significant difference in the expression levels of the target genes (Fig.15). Which might be the destabilizing effect of N-terminal 3xHA tag that might cause decreased expression and altered function of the tagged protein [58].

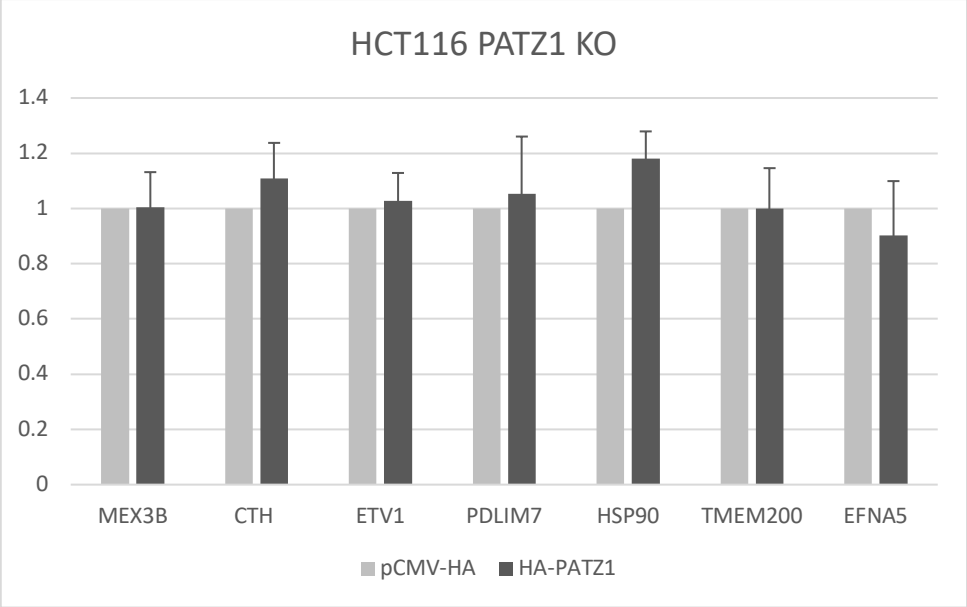


Figure 15. qPCR analysis of HCT116 cells that were transiently transfected with pCMV-HA-PATZ1 plasmid and pCMV-HA plasmid as control.

4.5.2. Lentiviral Transduction of hPATZ1 cDNAs into WT and PATZ1 KO HCT116 Cells

Considering the transient transfection results we decided to stably introduce the hPATZ1 cDNA and continue with lentiviral transduction of our PATZ1 KO HCT116 cells. Alongside the full length human PATZ1 cDNA we also cloned two other variants (variant 2 and 3) into the lentiviral expression vector LegoIG2Puro (explained detailed in methods section) as well. After the transduction of the WT and PATZ1 KO HCT116 cells with virus carrying PATZ1 var1, var2, var3 and empty vector as control, following 72 h of puromycin selection we proceeded with cell sorting using GFP expression of the transduced cells, since the expression plasmid had GFP and PURO selection markers. You can see the GFP ratios of the cells and sorted cell populations in the figure below (Fig.19). After sorting, cells were kept in puromycin for 72 h, and media replaced with fresh DMEM without puromycin.

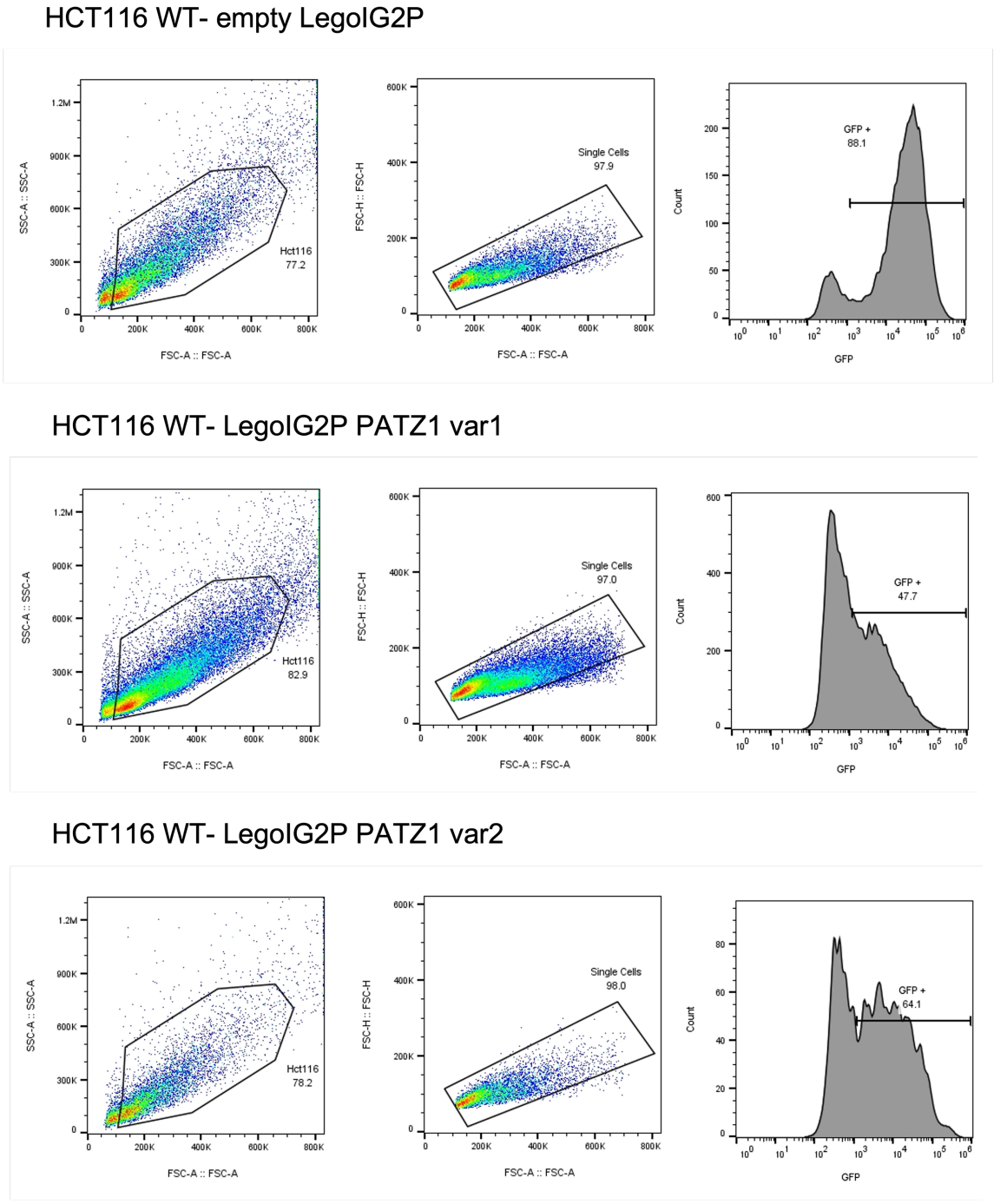


Figure 16. Cell sorting analysis of the lentivirus transduced GFP positive HCT116 WT cells.

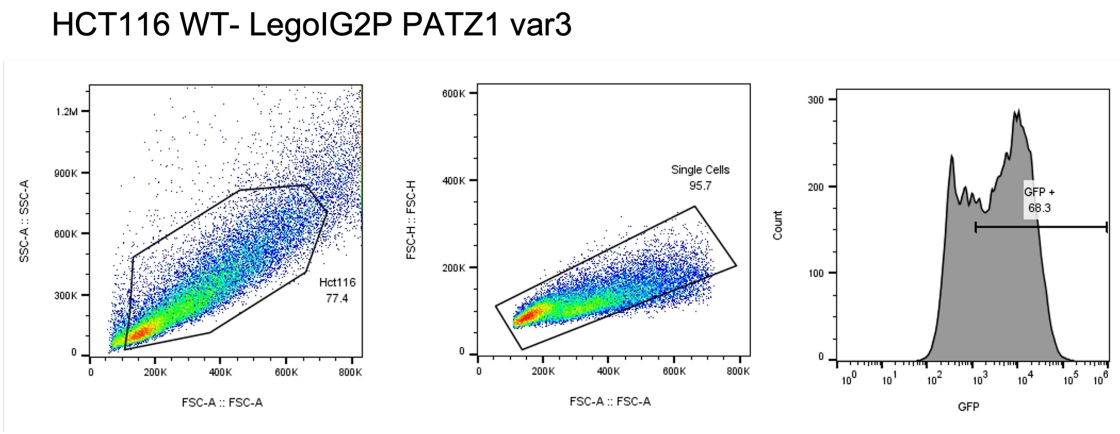


Figure 17. (Continued) Cell sorting analysis of the lentivirus transduced GFP positive HCT116 WT cells.

In parallel we also wanted to see if we can use integrating lentiviral vectors carrying human PATZ1 protein for transient transfection. For this purpose, we transfected PATZ1 KO HCT116 cells transiently with legoIG1Puro empty or PATZ1 carrying plasmids via PEI transfection and 24 h later checked both GFP and PATZ1 protein levels by Flow Cytometry (Fig. 17). In 24h we could not get any GFP and PATZ1 expression in transiently transfected cells.

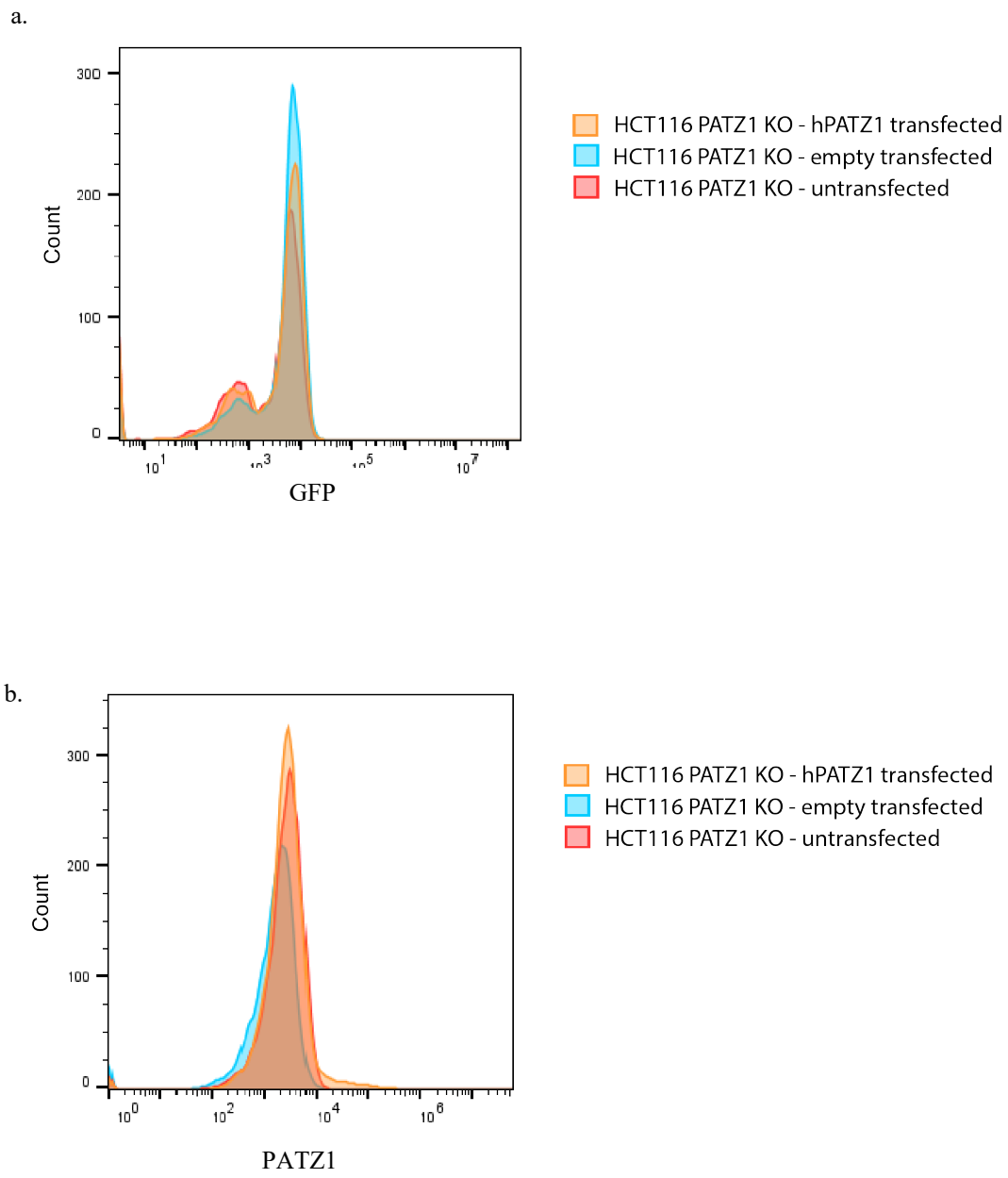


Figure 18. Flow Cytometry analysis of HCT116 PATZ1 KO cells transiently transfected with LegoIG2Puro empty or LegoIG2Puro carrying full-length hPATZ1 cDNA. a. GFP staining b. Anti-PATZ1 staining

After 72 h more in puromycin and expansion period, we wanted to confirm the hPATZ1 protein expression in GFP sorted- lentivirus transduced PATZ1 KO HCT116 cells. In order to do this, we performed intracellular anti-PATZ1 staining and performed flow cytometry analysis (Fig. 18).

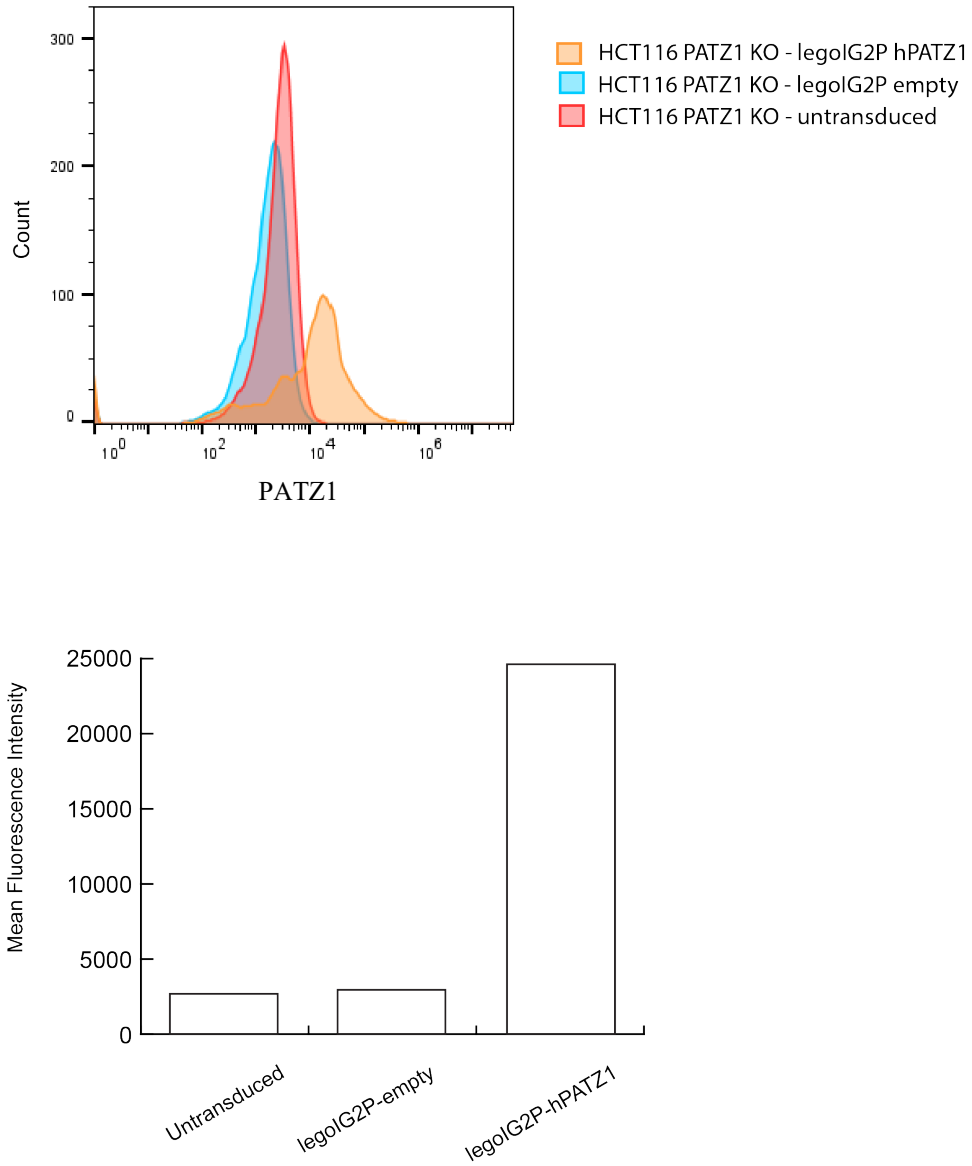


Figure 19. Flow Cytometry analysis of HCT116 PATZ1 KO cell line transduced with lentivirus carrying legoIG2Puro-empty or carrying full-length hPATZ1 cDNA. Intracellular staining performed by using primary anti-PATZ1(D-5) antibody and Alexa Fluor 647 anti-mouse Goat as secondary antibody.

To see the effect of expressing human PATZ1 protein stably back in the PATZ1 KO HCT116 cells on CTH protein expression, cell lysates for western blotting and cDNAs for qPCR were prepared. We performed anti-CTH blotting following SDS-PAGE and

qPCR following cDNA synthesis to see the expression levels of CTH in lentiviral transduced WT and PATZ1 KO HCT116 cells (Fig.19).

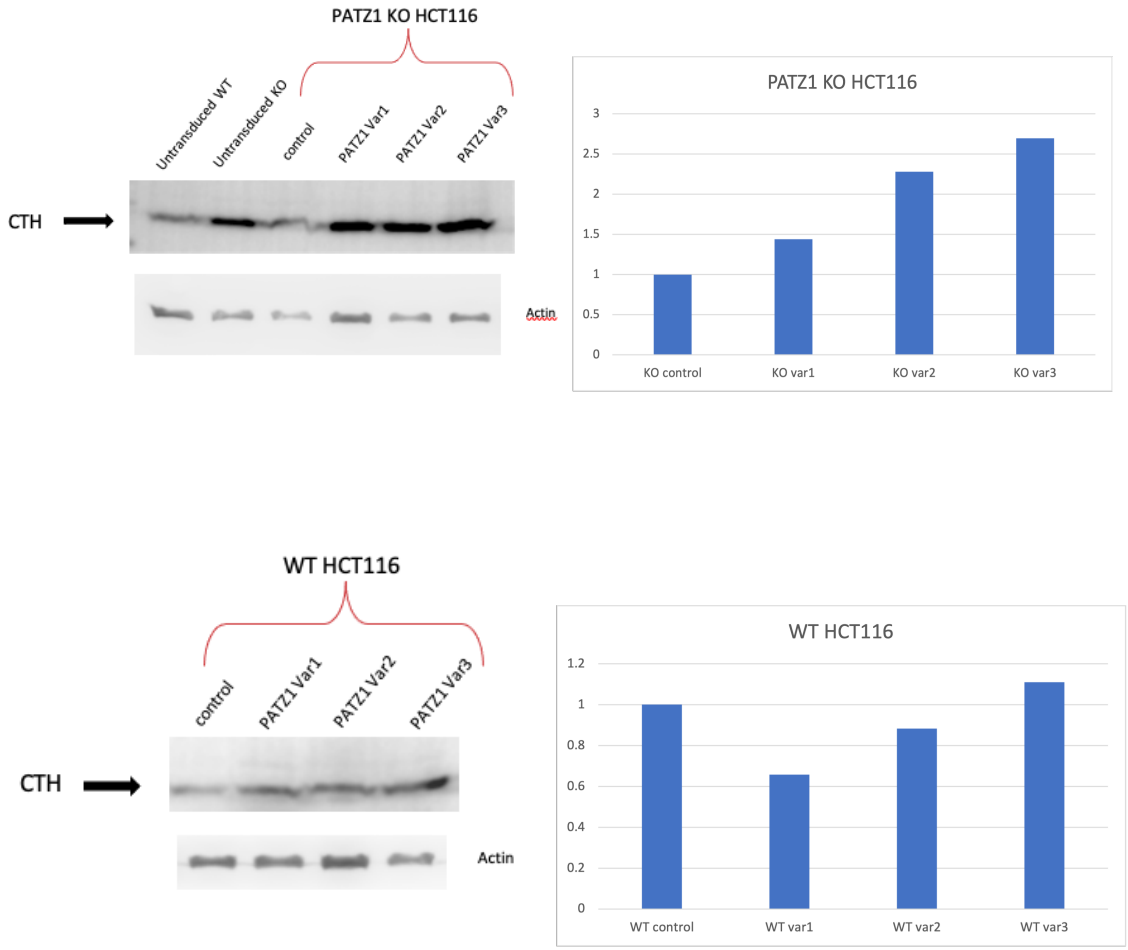


Figure 20. Anti-CTH western blotting and qPCR results of WT and PATZ1 KO HCT116 cells transduced with lentivirus carrying legoIG2Puro-empty or carrying full-length hPATZ1 cDNA, PATZ1 transcript variant 2 and variant 3.

As you can see in qPCR results introducing full length hPATZ1 cDNA back into the PATZ1 KO HCT116 cells did not recover the increased expression levels of CTH mRNA levels while we would expect decreased CTH mRNA levels compared to control. But interestingly, in WT HCT116 cells full length hPATZ1 cDNA decreased the CTH mRNA levels and created the expected effect. Anti-CTH blotting results were in parallel with qPCR results. Transcriptional roles of the PATZ1 variants are not known yet.

4.6. Intracellular Hydrogen Sulfide Levels of WT and PATZ1 KO HCT116 Cells

Cystathionine gamma-Lyase (CTH) protein is involved in the trans-sulfuration pathway from methionine to cysteine as an enzyme that produces hydrogen sulfide (H_2S). After we saw that the mRNA and protein levels of CTH in PATZ1 KO HCT116 cells are increased compared to WT cells, we decided to see if this has any effect on intracellular hydrogen sulfide levels.

4.6.1. Using a Chemical Compound for H_2S Measurement; 7AzMc

To see if there is any difference between WT and PATZ1 KO HCT116 cells we used a chemical compound 7AzMc (7-Azido-4-methylcoumarin) which is a fluorogenic probe that can be used for the detection of H_2S. In the presence of H_2S, aromatic azide moiety of the compound is selectively reduced to amino group which as a result creates the fluorescent 7-amino-4-methylcoumarin followed by an increase in the fluorescence (ex = 365 nm and em = 450 nm).

For this purpose, first we decided to try fluorescence microscopy. After the treatment of 50 uM of 7AzMc and 400 mM L-cysteine (substrate of the CTH enzyme), cells were washed with 1X PBS and visualized using florescence microscope with the correct channel (Fig. 20)

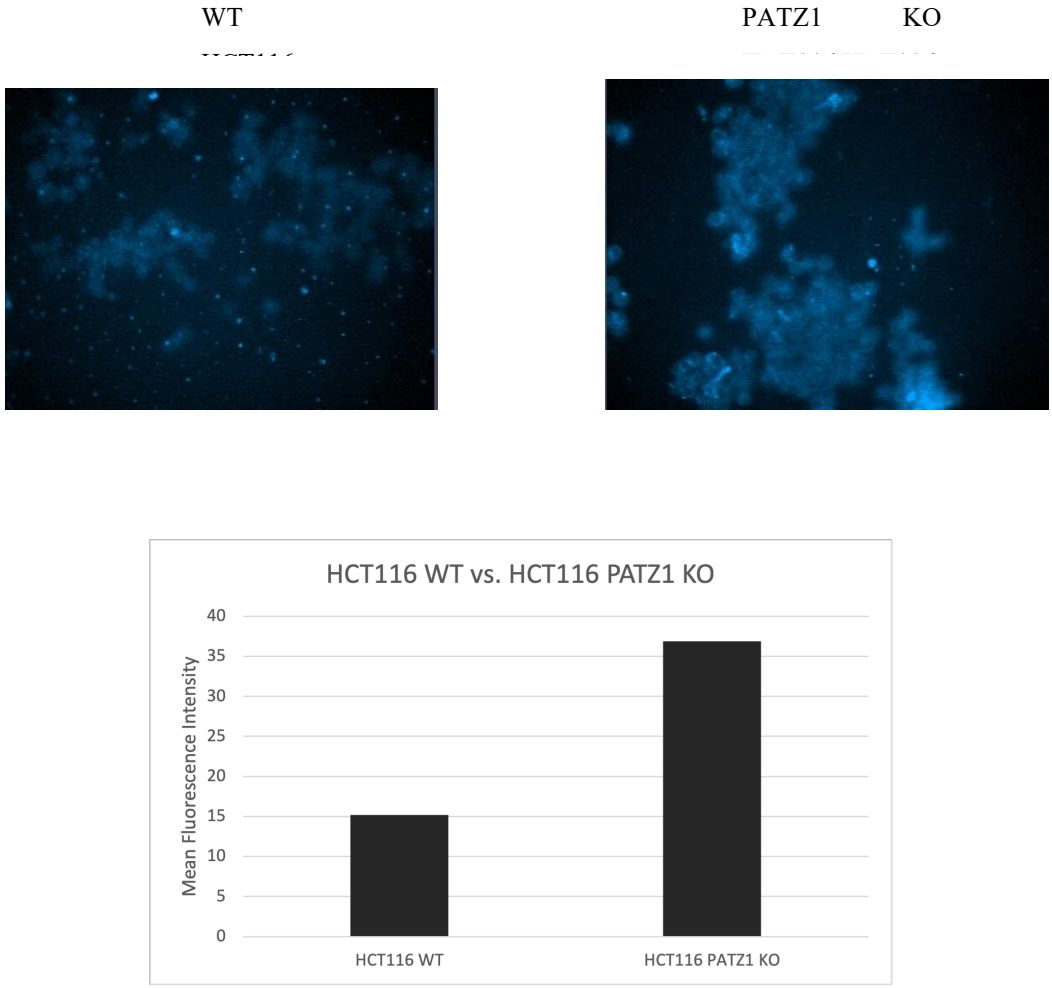


Figure 21. Fluorescence microscopy analysis of WT and PATZ1 KO HCT116 cells after the treatment with 7AzMc and L-cysteine for intracellular H_2S measurement.

In the images, due to the higher intracellular H_2S levels PATZ1 KO HCT116 cells have increased fluorescence intensity compared to the WT HCT116 cells.

As second approach we decided to use fluorescence spectrophotometer. To find the optimum conditions for treatment we used $2.5x10^4$, $5x10^4$ and $1x10^5$ cells per well in 96-well plate. In addition to the suggested concentration of 50 uM 7AzMc, we also tried 100 uM and 200 uM 7AzMc and incubated for 30 mins, 60 mins and 1 h at 37 ^oC (Fig. 21)

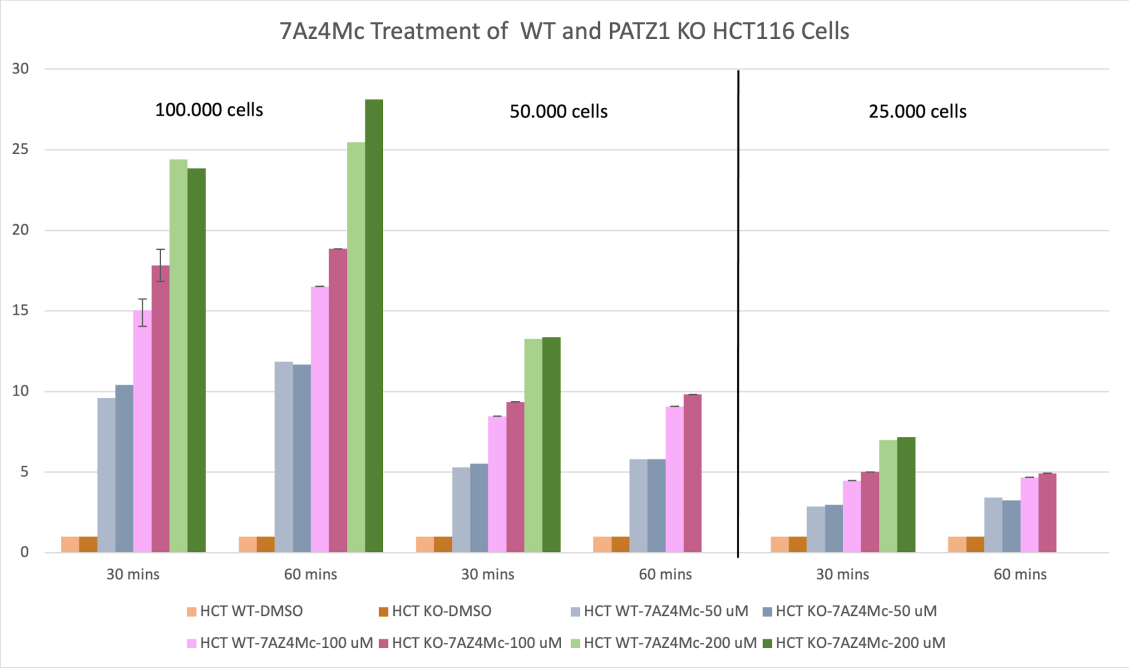


Figure 22. Optimization of the measurement of the mean fluorescence intensities of WT HCT116 and PATZ1 KO HCT116 cells treated with 7AzMc using fluorescence spectrophotometer.

Optimum conditions were 100 uM 7AzMc treatment of $1x10^5$ cells for 30 mins at 37 ºC. Our analysis shows that PATZ1 KO HCT116 cells have slightly higher levels of increased intracellular levels of hydrogen sulfide in comparison to the WT cells which is the result of increased H_2S producing CTH enzyme levels in the PATZ1 KO HCT116 cells.

4.7. PATZ1 Binds to Negatively Regulates CTH Promoter Region

In previous analysis we showed that CTH mRNA and protein levels increase in the absence of the PATZ1 protein in our CRISPR/Cas9 generated PATZ1 KO HCT116 cell line. As a result of this increase in CTH protein levels we also saw elevated levels of intracellular H2S in comparison to the WT HCT116 cell line. In CHIP-seq data coming from GFP-PATZ1 overexpressed HEK293 cells, peaks show us a 184 bp long region

immediately after CTH transcription start site which includes the promoter region. In order to study the effects of the PATZ1 protein on the transcriptional activity of the Cth gene we decided to PCR amplify a 1000 bp long DNA fragment starting upstream of the Cth gene including the possible PATZ1 binding sites and the promoter region and clone this fragment into a luciferase plasmid (Fig.22) to use Dual-Luciferase Assay system (Promega).

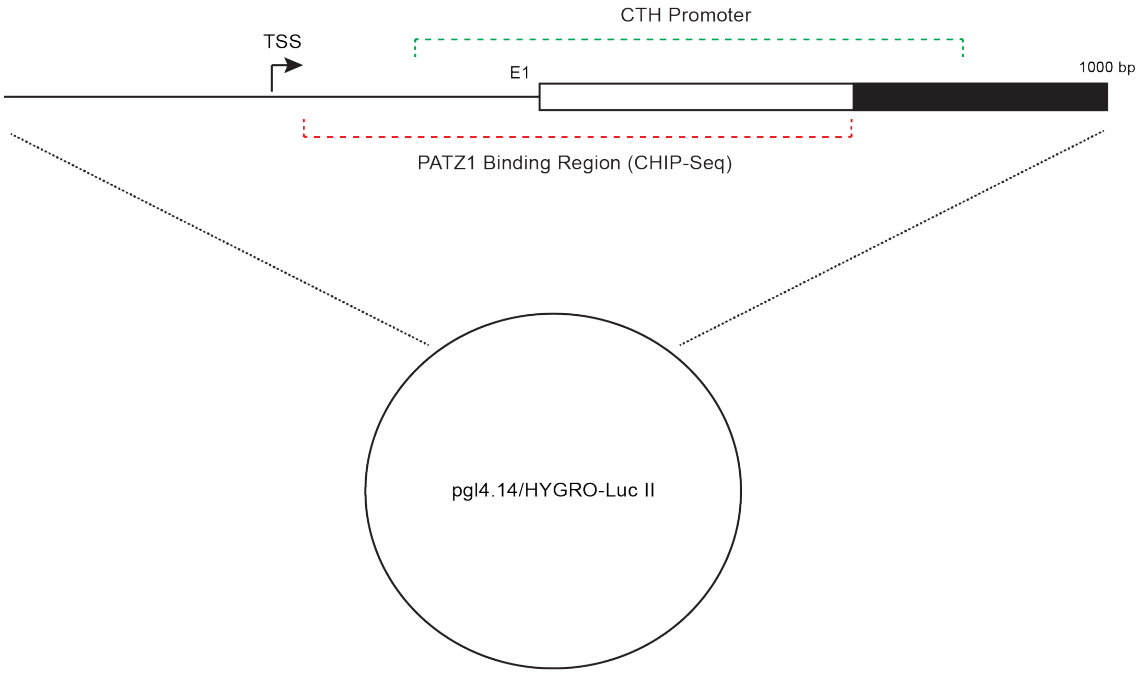


Figure 23. pgl4.14/Hygro plasmid construct ligated with PCR amplified Cth gene promoter region which includes the PATZ1 binding site from the CHIP-Seq data.

We used two different constructs for the assay. First plasmid is pgl4.14/Hygro which is a Firefly luciferase vector containing no promoter or no enhancer. These vectors, which contain a multiple cloning site immediately upstream of the LucII gene, allow us to clone target regulatory element to drive expression of the reporter gene. Second plasmid is the Renilla Luciferase vector to use as an internal normalization control.

We transfected PATZ1 KO and WT HCTT16 cells with empty pgl4.14 or pgl4.14-CTH promoter vectors with pRenilla internal control in each condition via using PEI transfection. 24 h after transfection according to the assay protocol we prepared the cell

lysates and as a first step checked the firefly luciferase activity. After stopping the first reaction we measured Renilla luciferase activity as second step. Firely luciferase activities were normalized to internal Renilla luciferase activities of each condition (Fig. 23).

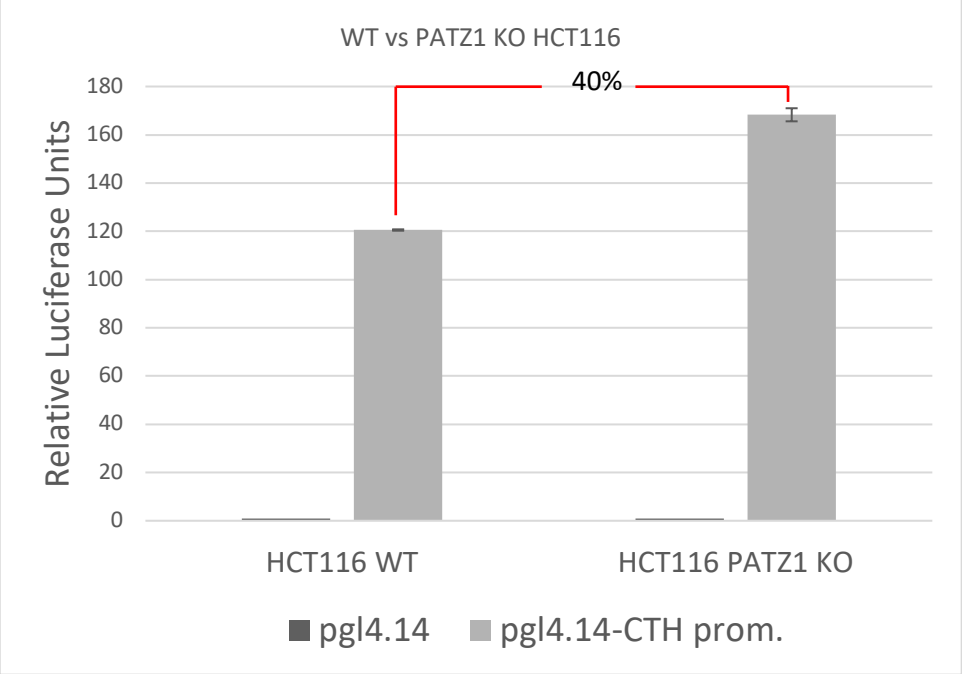


Figure 24. PATZ1 TF binds to the CTH promoter and suppresses the CTH promoter activity. Dual-Luciferase Assay results of WT and PATZ1 KO HCT116 cells transfected with empty and CTH-promoter ligated pgl4.14/Hygro luciferase plasmids and pRenilla plasmid as an internal control.

We performed the analysis of Dual-Luciferase Assay results of WT and PATZ1 KO HCT116 cells transfected with empty and CTH-promoter carrying pgl4.14/Hygro luciferase plasmids and pRenilla plasmid as an internal control. As a result, we saw that the luciferase activity of WT HCT116 cells was significantly lower in comparison to PATZ1 KO HCT116 cells. This shows us that PATZ1 transcription factor binds to the Cth gene promoter and suppresses its activity.

4.8. Comparison of Mouse RNA-Seq Data and Human CHIP-Seq Data

We decided to use the RNA-seq data of a previous study (GEO accession GSE65864) that was done using WT and PATZ1 KO MEF cells in our laboratory and the same CHIP-Seq data from HEK293 cells (GEO accession GSE92195) that we used for our previous analysis to select common genes from MEF DEGs and HEK293 CHIP results. There were 2263 genes that were bound by hPATZ1 transcription factor in HEK293 cells as we previously analyzed and 156 differentially expressed genes (DEGs), of which 83 were upregulated and 73 were downregulated, in PATZ1 KO MEFs as shown in our study [59]. We selected upregulated DEGs and downregulated DEGs in the absence of PATZ1 in MEFs and bound by PATZ1 in HEK293 cells (Fig.24) to design specific qPCR primers (Table 3.2).

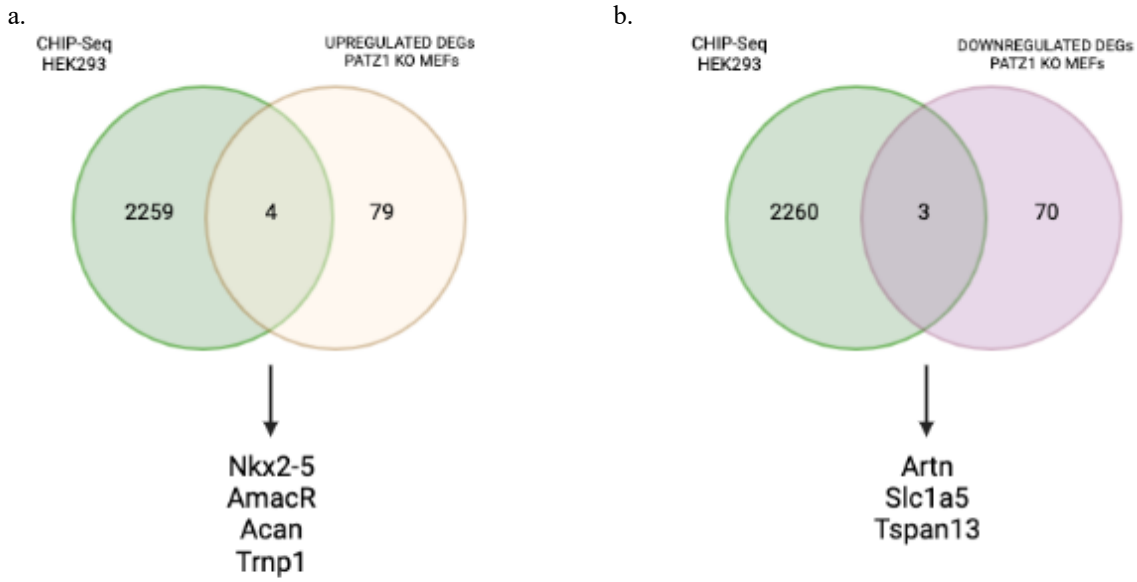


Figure 25. Schematic representation of comparison of RNA-Seq and CHIP-Seq data sets a. Upregulated common DEGs b. Downregulated common DEGs between MEF RNA-Seq and HEK293 CHIP-Seq data sets.

Following the isolation of total RNAs and the synthesis of cDNAs, we performed qPCR for AmacR, Acan, Trnp1, Artn, Slc1a5 and Tspan13 genes on our WT and PATZ1 KO HCT116 cell lines usisng specifically designed qPCR primers (Fig.25).

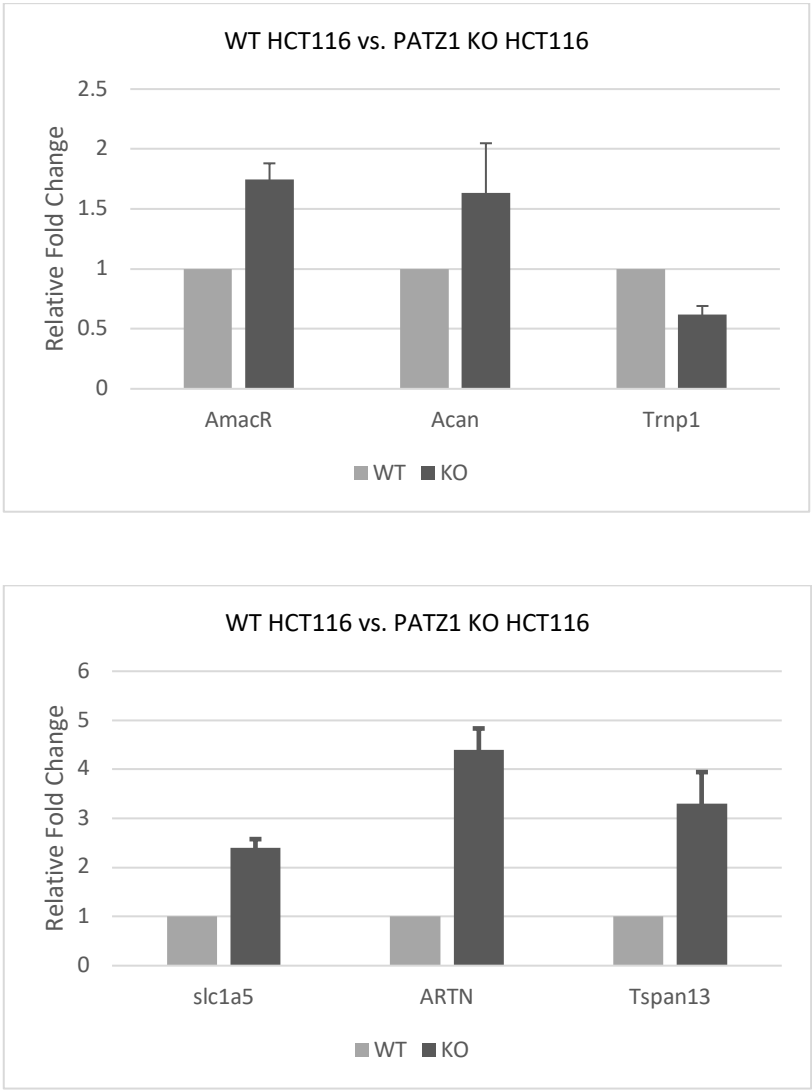


Figure 26. qPCR analysis of HCT116 cells for the upregulated and downregulated DEGs.

The qPCR experiments with the DEGs that we selected using MEF RNA-seq data did not give us the expected expression levels for the genes we chose on our PATZ1 KO HCT116 cell line. AmacR and Acan expression levels were upregulated as expected whereas Trnp1 gene was downregulated in PATZ1 KO HCT116 cells which was upregulated in PATZ1

KO MEF cells and Slc1a5, ARTN, Tspan13 genes were upregulated in PATZ1 KO HCT116 cells which were downregulated in PATZ1 KO MEF cells unexpectedly.
We also performed qPCR with the same primers using cDNAs from WT and PATZ1 KO U87MG cell line to check if would get the similar expression levels (Fig.26).

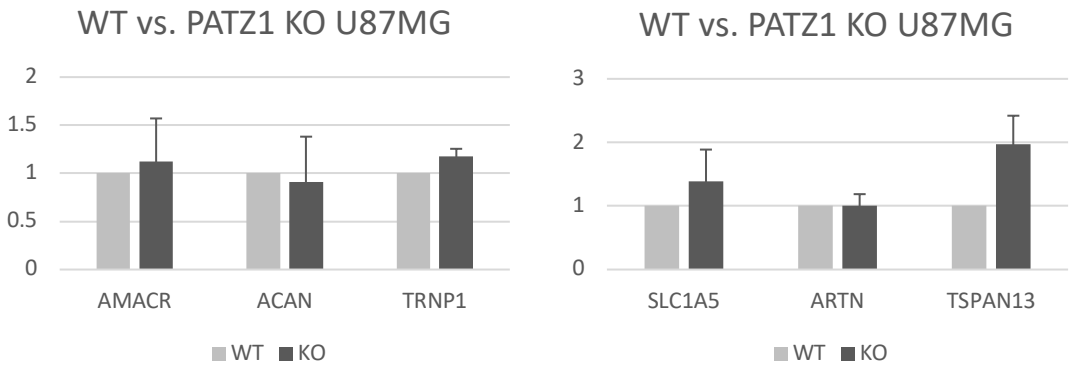


Figure 27. qPCR analysis of U87MG cells for the upregulated and downregulated DEGs.

The qPCR experiments with the DEGs that we selected using MEF RNA-seq data did not give us the expected results for all the genes on our PATZ1 KO U87MG cell line as well. We would expect to see significant increase in the expressions of AmacR, Acan, Trnp1genes whereas we would expect to see significant decrease in the expressions of Artn, Slc1a5 and Tspan13 genes according to RNA sequencing data of PATZ1 KO MEFs. These data support the fact that PATZ1 is a tissue specific, context dependent transcription factor that might act as both transcriptional repressor and activator.

4.9. Effect of PATZ1 Protein on Cell Proliferation

To determine if knocking out the patz1 gene in the HCT116 cell line has any effect on cell proliferation we performed the trypan blue cell viability assay. For this purpose, same number of WT and PATZ1 KO HCT116 cells were seeded into 6-well plates and after 24 h, 48 h and 72 h of incubation in normal growth conditions, cells were collected and

counted using trypan blue via cell counter. Since the HCT116 cells are adherent cells, to see the difference in proliferation rates we decided to count live cells. (Fig.27)

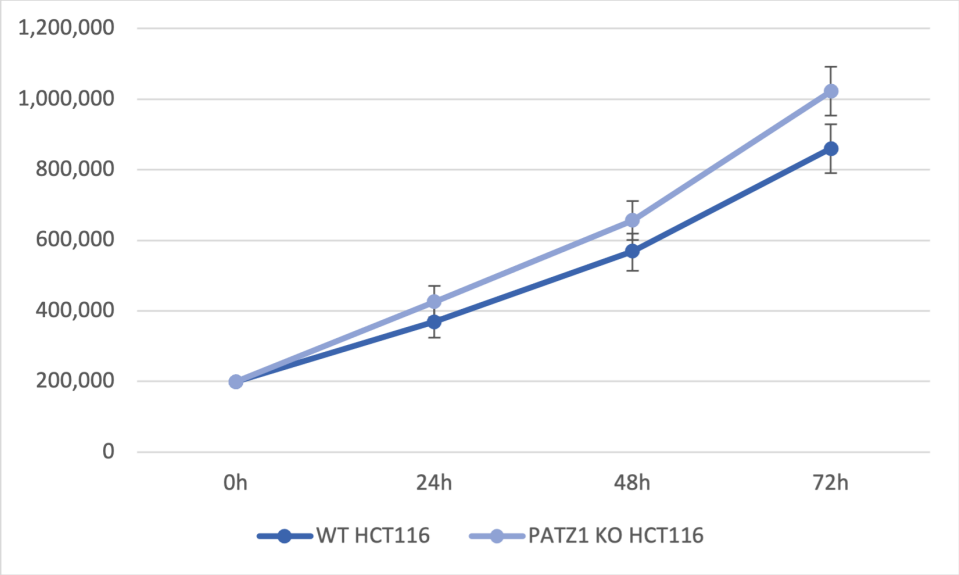


Figure 28. Cell proliferation measurement using manual cell counting via trypan blue staining.

As you can see in the figure above HCT116 PATZ1 KO cells showes slightly increased levels of cell proliferation in comparison to HCT116 WT cells.

4.10. mRNA Quantification of the Patz1 Gene

To determine the mRNA levels of the patz1 gene, we decided to perform qPCR by using specific primers binding to different regions on mRNAs which will allow us to quantify all alternative splicing variants of the patz1 gene (Fig. 28).

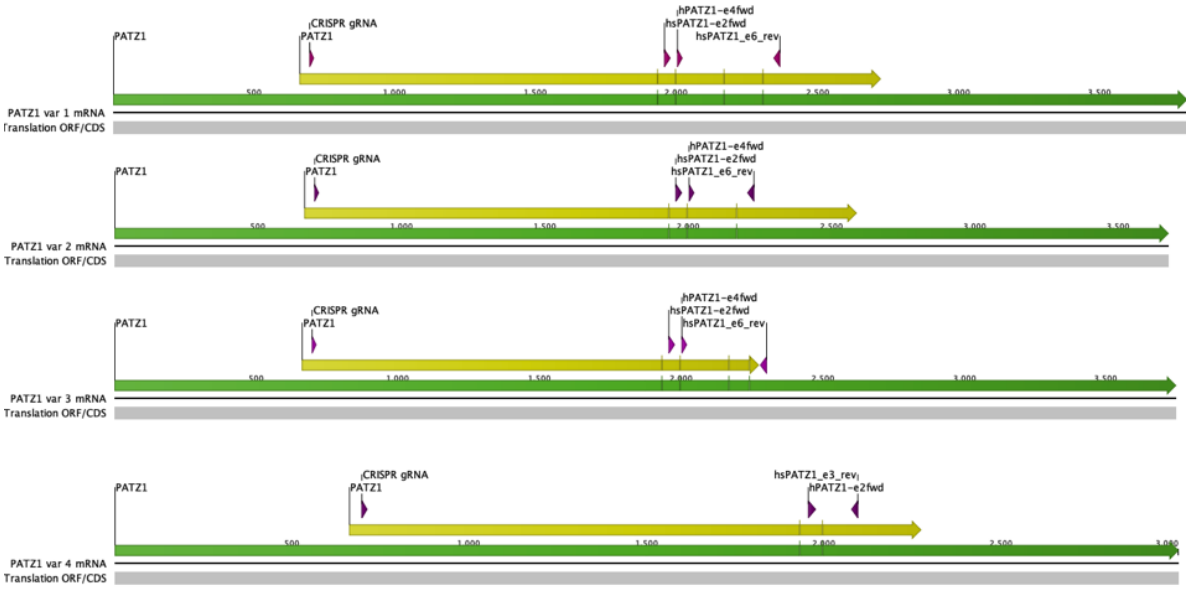
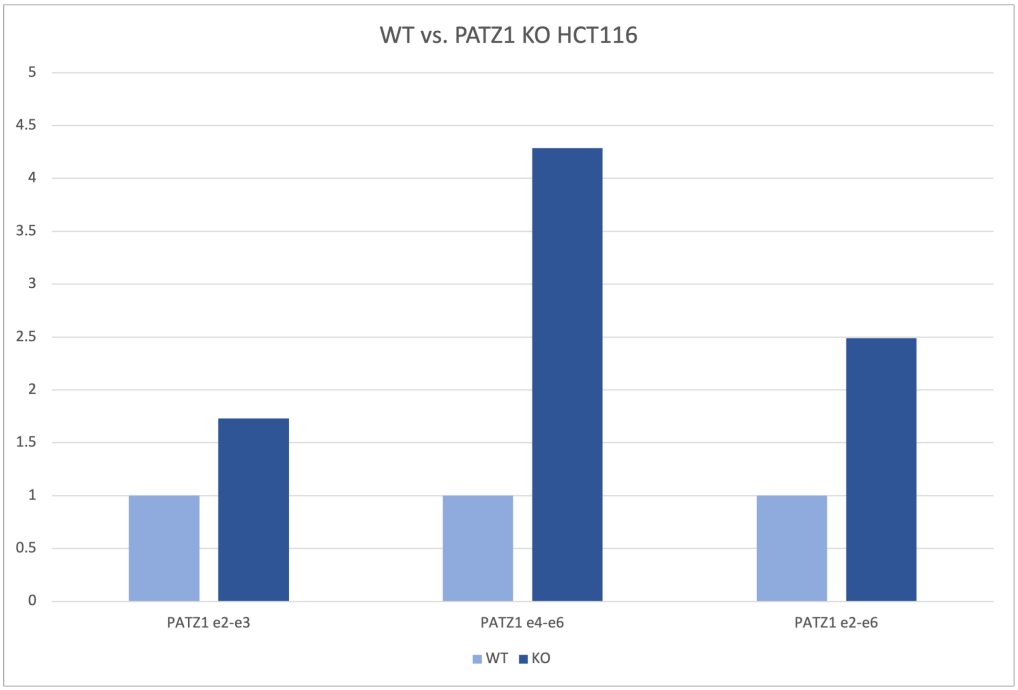


Figure 29. Specific primer pairs binding to different transcript variants of Patz1 gene designed to perform qPCR.

We performed qPCR on both WT and PATZ1 KO HCT116 and U87MG cell lines to see if there are any differences between the cells in the absence of the PATZ1 protein (Fig. 29).



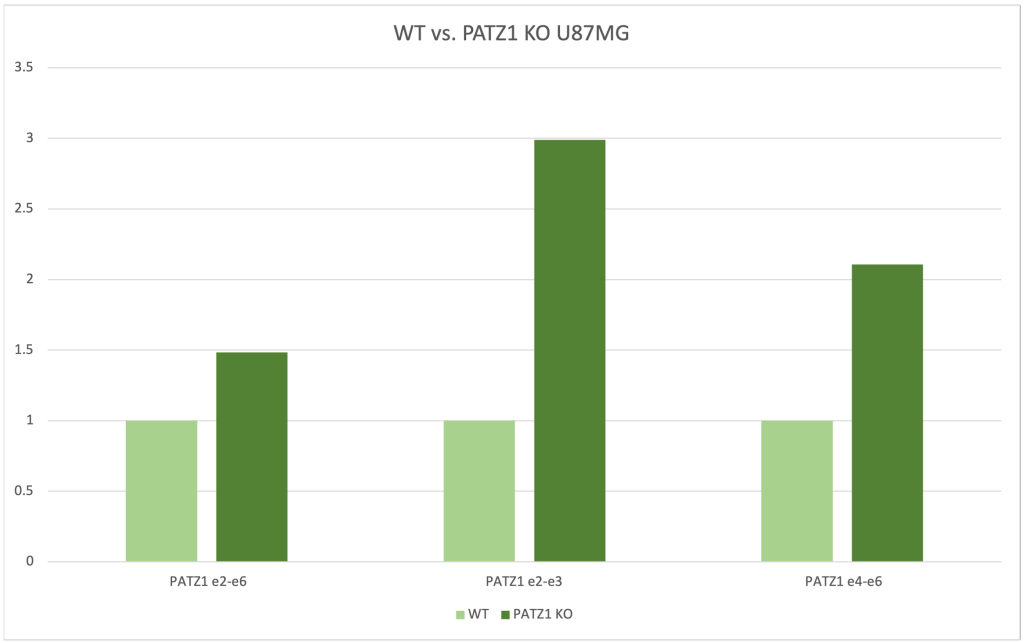


Figure 30. Quantification of Patz1 mRNA levels via qPCR analysis using specific primer pairs binding to different transcript variants.

In PATZ1 KO cell lines we would expect to see decreased Patz1 mRNA levels but in both PATZ1 KO HCT116 and PATZ1 KO U87MG cell lines patz1 mRNA levels were

increased in comparison to WT cell lines. When we checked CHIP-seq results we saw that the human PATZ1 protein binds to its own promoter/enhancer. You can see the possible binding sites and a PATZ1 homodimer binding model in the figure down below (Fig 30).

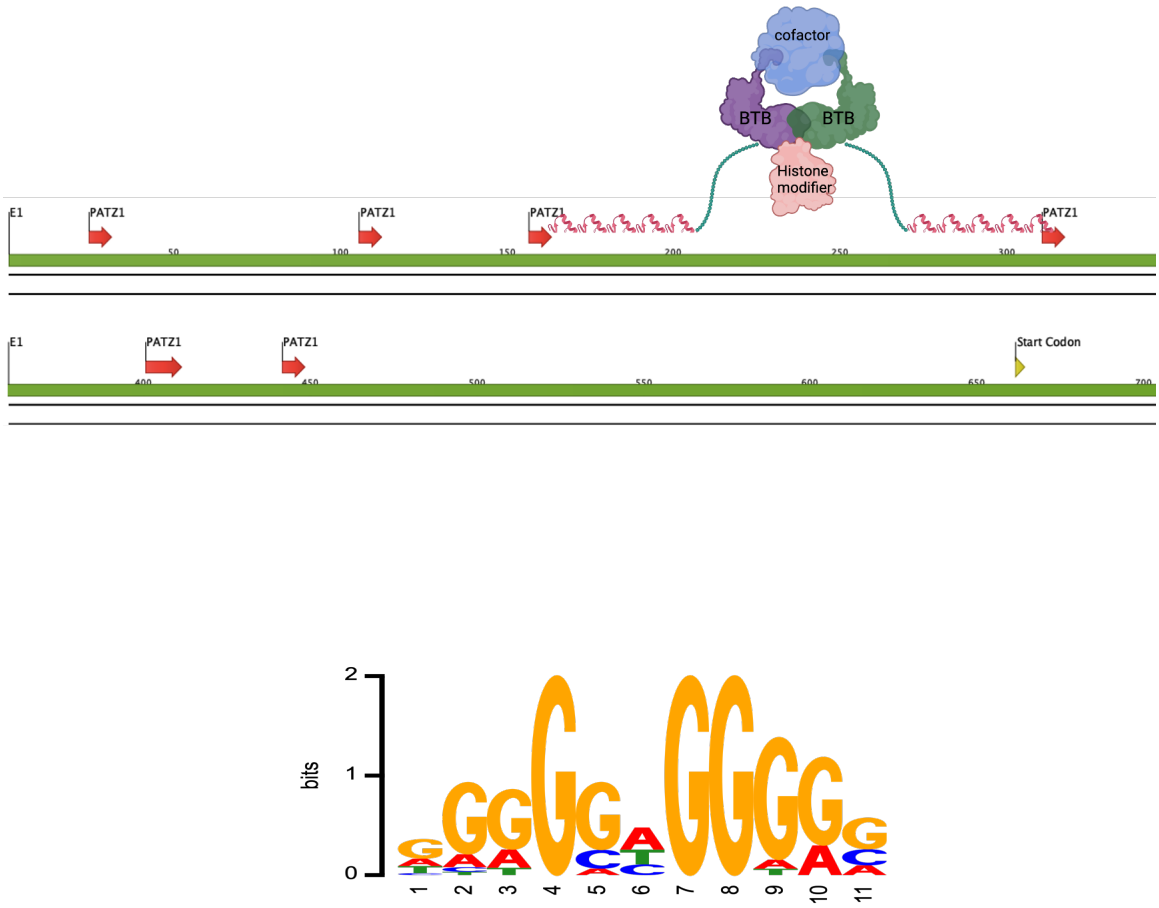


Figure 31. PATZ1 TF binding sites on the Patz1 promoter/enhancer region.

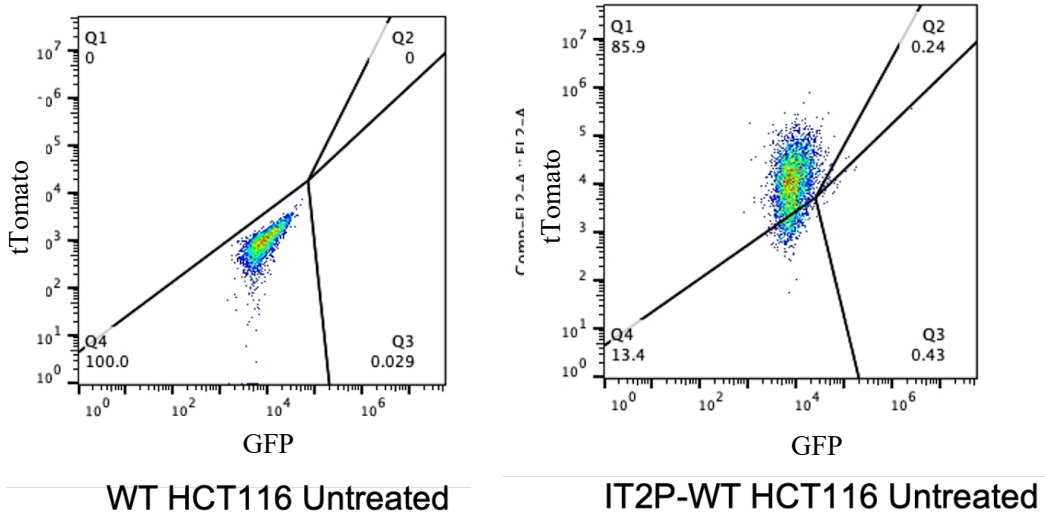
We thought that, due to the absence of PATZ1 protein in KO cells, promoter/enhancer of Patz1 gene cannot be bound by PATZ1 protein therefor the negative feedback mechanism is blocked. Since PATZ1 cannot suppress its own expression by binding to its own

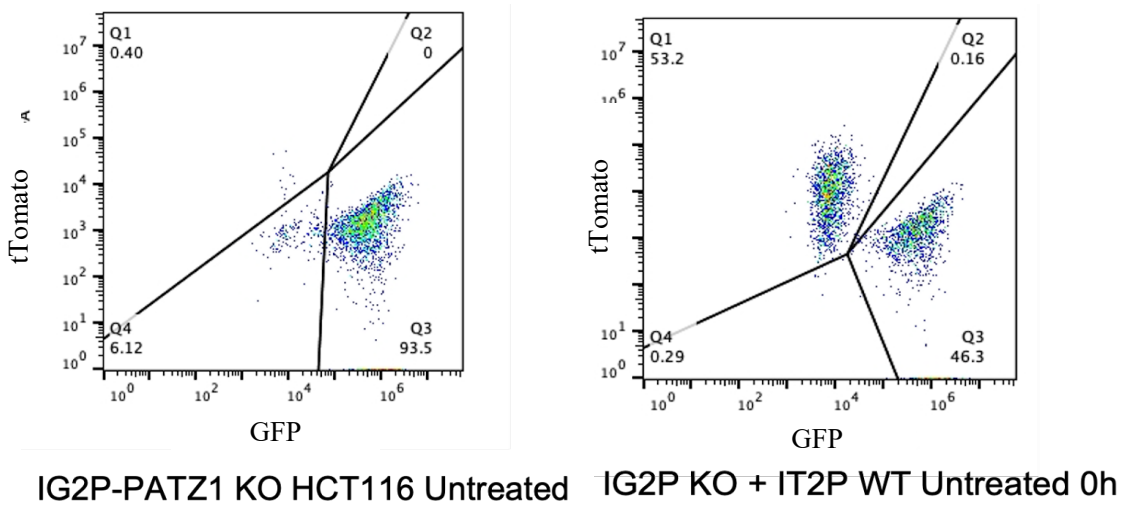
promoter/enhancer, we see the increase in the mRNA levels which are not coding PATZ1 protein since the gene was disrupted by CRISPR/Cas9 genome editing tool.

4.11. Effects of Hydrogen Sulfide on Doxorubicin Response

Patz1 was previously shown to be involved in the DNA damage response [59]. After we saw the difference in H_2S levels between WT and PATZ1 KO HCT116 cells, to see if this difference has any effect on the doxorubicin treatment which causes DNA damage, we decided to perform doxorubicin treatment on WT and PATZ1 KO HCT116 cells. Since hydrogen sulfide is also a gasotransmitter that can be secreted outside of the cell, we thought maybe increased levels of H2S in PATZ1 KO cells might have an effect on WT cells when they are growing together in the same environment. To see this effect, first we generated stably GFP expressing WT and tTomato expressing PATZ1 KO HCT116 cells. We seeded same number from each cell line into the same petri dish and treated them with and without different concentrations of doxorubicin. After 24 h of doxorubicin treatment we performed flow cytometry analysis to see if there is any change in GFP/tTomato ratio in mixed cell populations (Fig 31).

As you can see in the results, there is a slight change in the GFP/tTomato ratio in mixed cell populations treated with low concentrations of doxorubicin which is not significant. This change might be due to the H_2S difference between the WT and PATZ1 KO HCT116 cells since small increase in intracellular H_2S levels has an inducing effect on cell proliferation in colon cancer cells [60,61].





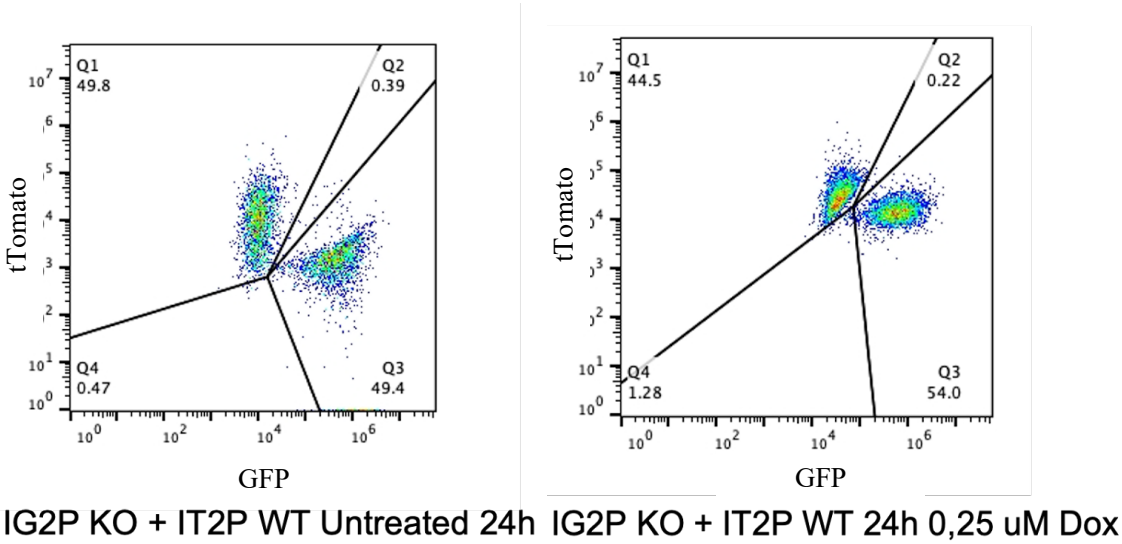


Figure 32. Proliferation analysis of stable GFP expressing WT and tTomato expressing PATZ1 KO HCT116 cells following Doxorubicin treatments using flow cytometry.

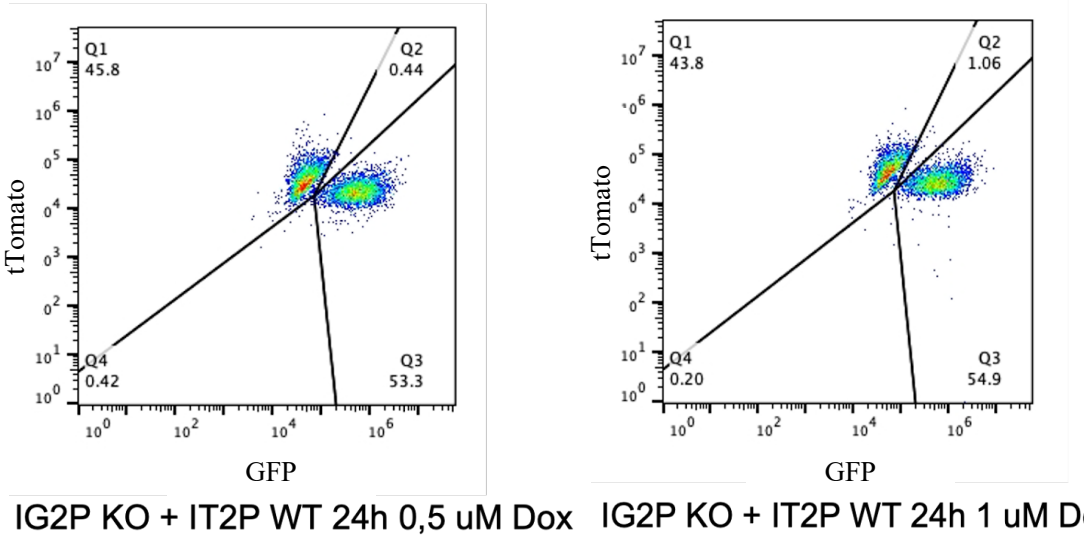


Figure 33. (Continued) Proliferation analysis of stable GFP expressing WT and tTomato expressing PATZ1 KO HCT116 cells following Doxorubicin treatments using flow cytometry.

5. DISCUSSION AND CONCLUSION

Transcription factors are the main regulators of gene expression. Clarifying the mechanisms through which TFs function and how they affect the activity of other proteins or regulatory molecules could provide information on many intracellular signaling pathways and on the development of the phenotypic outcomes. To our knowledge there are almost 1639 TFs encoded by the human genome. There are many different studies published and continuing to understand the functions of human TFs. Most transcription factors bind to DNA by recognizing specific DNA sequences called transcription factor binding sites (TFBSs), and thus, they regulate the transcription of adjacent genes through their promoters or of genes that are distant through their enhancers. In all cell types of a particular species, some of cis-regulatory regions such as promoters and enhancers are same at the DNA sequence level for a specific TF. As a result, the transcriptional setup of each cell type is the consequence of the activity of specific TFs expressed in that cell type, and how genes are selected for transcriptional activation or repression. The same transcription factors (TFs) can be expressed at the same rate in various cell types, but they could have distinct binding sites because of variant chromatin structures and epigenetic alterations that modify TF activity and regulatory pathways. TFs are essential regulators of many cellular processes, including the processes of tissue growth and embryonic development which can result in cancer and other diseases when aberrantly controlled. Therefore, gaining regulatory knowledge of the TF network would lay a crucial foundation for both future research and treatment methods.

Zinc finger domains are the most common DNA-binding domain family which are generally found in combination with brick-a-brac, trim-track, broad complex (BTB) or kruppel-associated box domains. It is well known that kruppel-associated box domain containing TFs recruit transcriptional repressors whereas the transcriptional role of BTB

domain containing TFs is yet to be understood. ZNF proteins participate in several physiological processes, such as cell proliferation, differentiation, and death, by controlling gene expression [11]. For example, KLF4 plays a crucial role in the intestines in addition to its well-known function in skin homeostasis. The terminally differentiated epithelial cells and goblet cells in this tissue express KLF4, which promotes differentiation and suppresses proliferation. KLF4 specifically inhibits the expression of genes regulated by β-catenin by interacting with it and repressing intestinal epithelial growth[62]. Furthermore, in KLF4 mutant animals absence of the goblet cells provides that KLF4 is crucial for the differentiation of goblet cells[63]. ZF containing transcription factor SLUG (C2H2-type, transcription factor) is involved in *in vitro* and *in vivo* adipocyte differentiation. SLUG -/- mice shows decreased levels of white adipose tissue mass in comparison to wild-type mice, whereas the white adipose tissue mass level is higher in SLUG overexpressing mice. SLUG potentially regulates the development of white adipose tissue via controlling Histone deacetylase 1 (HDAC1) recruitment to the PPARG promoter, creating a chromatin state that is more accessible to PPARG transcriptional activators[64].

What makes BTB-ZF transcription factors different than other TFs is the effect of dimerization on transcriptional activity by increasing the affinity and specificity to specific DNA sequences or by keeping the dimers inactive below a certain binding threshold [65]. Therefore, hetero- or/and homo-dimerization mediated by BTB-domains is a crucial mechanism in gene regulation. A dimer's ability to inhibit or activate transcription can be changed by switching out one of its components. It has been demonstrated that BTB-ZF TFs have structural and organizational roles, particularly through promoting the formation of DNA loops to mediate enhancer-promoter interactions[20]. Additionally, BTB-ZF TFs has a new mechanism of chromatin compartmentalization. For example, Zbtb7b, also known as ThPox, interacts with HDAC3 and Lap2b to bind specific DNA sequences to the inner nuclear membrane, which results in the transcriptional silencing of particular genomic regions [66].

BTB domain containing proteins can bind to transcriptional coregulators and chromatin remodeling proteins such as HATs and HDACs. BTB domain containing proteins form homodimers or/and heterodimers with other BTB domain containing

proteins to bind to the coregulators. While BTB domain is necessary for protein-protein interactions, ZF domains are necessary for protein-DNA interactions. We can divide BTB domain containing transcription factors into three groups according to which DNA-binding domain they carry. These are Zbtb TFs with zinc finger domains, Bach TFs with bZIP domains and Nacc TFs carrying both ZNF and bZIP domains [18]. In human we have two members from each Bach and Nacc TF families and 49 members of the Zbtb TF family. Functional studies of Zbtb TFs are mainly focused on Bcl6 (Zbtb27), Lrf (Zbtb7a) which are protooncogenes and on the tumor suppressor Plzf (Zbtb16). Studies showed that these transcription factors act as repressors via recruiting silencing complexes such as NCOR/SMRT, BCOR, SIN3A/B and NuRD [67]. Transcriptional and structural roles of other members of the BTB-ZF family remains to be investigated.

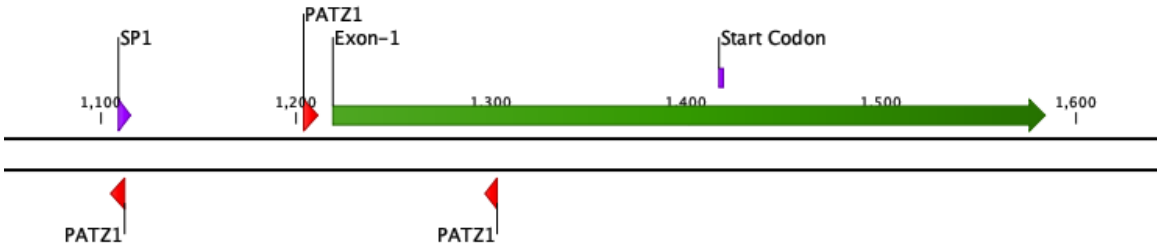
In this study, we focused on the human transcription factor PATZ1 which is a member of BTB-ZNF TF family. We identified its transcriptional targets in human colon cancer cell line HCT116 for the first time in literature.

PATZ1 protein is a context dependent transcription factor that belongs to the family of BTB-ZNF TFs. It has been shown that the PATZ1 protein participates in a wide range of physiological and pathological processes by acting as either an activator or a repressor in the control of gene expressions, depending on the cellular context [36]. Human Colon Cancer cell line HCT116 is a widely used, highly aggressive cell line with little or no capacity to differentiate. In many different research areas from anti-cancer drug screenings to transcriptome analysis, HCT116 cell line is used as an experimental model with being quite easy to grow and proliferate in vitro. In literature there are no studies exploring the transcriptional role of PAZT1 in HCT116 cell line, thus we chose this cell line as our main model system to identify transcriptional targets of the human TF PATZ1.

To identify differentially expressed genes between the absence and the presence of the PATZ1 transcription factor in our model system, after generating PATZ1 knockout HCT116 cell line by using CRISPR/Cas9 genome engineering tool RNA-Seq analysis was performed. There were 156 DEG of which 89 these DEGs were downregulated and 67 were upregulated in the absence of PATZ1 protein (APPENDIX D). GO term analysis showed us that the PATZ1 protein is involved in biological processes like neurogenesis,

cell adhesion, tissue development, neuron differentiation, regulation of muscle cell differentiation were affected (Fig.11b). Our data demonstrates that PATZ1 transcription factor may act both as an activator and a repressor in this cell line. Even though in many of the earlier studies PATZ1 was defined as a transcriptional repressor, recent studies show us that PATZ1 protein is a context dependent, tissue specific transcription factor. For example, in one of the studies FADS1 (fatty acid desaturase 1) expression was shown to be downregulated by the PATZ1 TF in HepG2 cell line [1] whereas in a different study it was demonstrated that the PATZ1 binds to the DNA sequence flanking the TSS of CDKN1B (p27) and positively regulates its transcription in liver cancer cells [37]. RNA-Seq data analysis provides us the list of total DEG between WT and PATZ1 KO HCT116 cells. However, changes in the expression levels of some genes might be the result of indirect effects of PATZ1 or a PATZ1-independent mechanism. To determine which DEGs were directly bound by the PATZ1 TF we decided to compare our RNA-seq data to CHIP-seq data publicly available on NCBI-GEO (Gene Expression Omnibus) database [57] (GEO accession GSE92195). CHIP-seq data provides us the information of which genomic regions are bound by the human PATZ1 transcription factor directly in human embryonic kidney cell line HEK293. As a result of the comparison of these two data sets, we found 15 genes (3 upregulated and 12 downregulated genes in the absence of the PATZ1 protein) out of 156 DEGs were bound by the PATZ1 TF. To further analyze the difference in the expression levels of these 15 genes we performed qPCR and decided to focus on Cth gene which had the most significant change in its expression level. Considering the fact that CHIP-seq was performed on a different cell line (HEK293), to confirm the direct binding of the PATZ1 protein to the promoter/enhancer region of Cth gene we used the Dual Luciferase Assay System. We demonstrated that the PATZ1 transcription factor binds to and negatively regulates the expression of the Cth gene for the first time in the literature. When we analyze the sequence of the Cth promoter region we saw several possible PATZ1 binding sites. Interestingly one of this PATZ1 binding sites overlaps with the SP1 binding site (Fig. 32a). The transcription factor specificity protein 1 (SP1) increases Cth gene expression in cells as response to biological processes such as inflammatory stress[50] (Fig.32b).

a.



b.

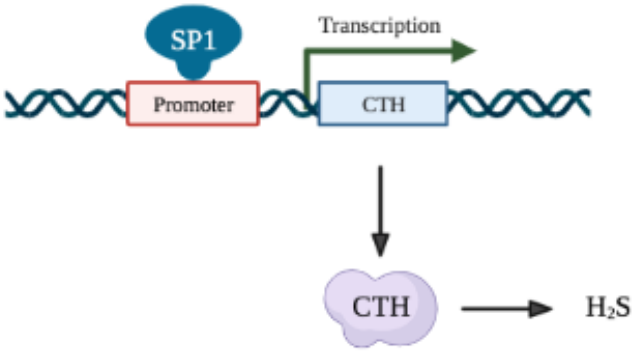


Figure 34. Representative figures of the regulation of the Cth gene promoter a. Promoter regions of the Cth gene (blue) annotated with the possible PATZ1 binding sites (red), SP1 binding site (purple), CHIP-Seq binding peak start and end sites (yellow) and exon 1 of the Cth gene (green). SP1 binding site overlaps with the PATZ1 binding site. b. SP1 TF binds to the promoter of Cth gene and activates the expression.

In a study published in 2017, Pan et al. demonstrated that the PATZ1 binding site at the enhancer region of FADS1 gene overlaps with the SP1 binding site and in the presence of rs174557 genetic variation binding of the PATZ1 protein is diminished which allows activating SP1 complex to determine the activity of this region. This study shows the competitive binding between the suppressive PATZ1 TF and the activating SP1 complex [1]. By using this approach, as a future work we can focus on the binding

competition between PATZ1 TF and SP1 TF at the Cth gene promoter region to understand the transcriptional regulation of the gene by this specific sequence.

Generating a knockout/knockdown cell line is one of the main approaches to study regulatory effects of a transcription factor. In this study, we successfully generated stable Patz1 knockout cell lines via CRISPR/Cas9 genome engineering tool. Following the analysis of DEGs, rescue experiments needed to be performed to see if re-expression of the transcription factor would recover the expression phenotype. In our study we performed transient and stable re-expression of the Patz1 gene, yet we could not observe the expected recovery of the expression profile. One of the reasons might be due to the unintended insertion of the puromycin resistant plasmid coding Cas9/gRNA complex into the genome with or without including the selection marker. Under the selective pressure of long term puromycin selection, insertions might occur during transient transfection. The risk of stable integration of the plasmid containing the CRISPR-Cas9 components into the host cell's genome exists regardless of whether DNA is transfected through lentivirus or plasmid. This has been found to promote chromosomal instability and heterogeneity, impart in vivo metastatic potential, and affect cell behavior and function. To overcome this unwanted effect, we are planning to perform rescue experiments via using expression plasmids carrying PATZ1 cDNA with introduced point mutations to the CRISPR targeted PAM sequence. With this approach we can prevent Cas9/gRNA complex to recognize and cut the re-expression plasmids. Another reason could be the above-mentioned competitive binding of SP1 and PATZ1 TFs to the promoter region of the Cth gene (Fig.33). Co-occupancy of different transcription factors and differences in their DNA binding affinities determine the outcome of the regulatory element activity. The affinity of a transcription factor for its binding sequence does not explain its genomic occupancy, instead, combinatorial action of multiple TFs and complexes determine the precise control of binding at regulatory elements. Transcription factor complexes can create reversible and irreversible changes on gene regulatory elements via histone modifications especially if they are interacting with histone modifiers. Thus, studying the competitive binding between transcription factors or affinity differences between binding partners in complexes would broaden our knowledge about how PATZ1 TF functions to control its targets.

In our study we focused on the Cth gene as our candidate PATZ1 target. Cystathionine gamma-Lyase (CTH) enzyme which is encoded by the Cth gene is involved in the trans-sulfuration pathway from methionine to cysteine as an enzyme that produces H_2S. We decided to check if increased levels of CTH enzyme in Patz1 knockout HCT116 cells created any change in the intracellular H_2S levels. For this purpose, we used a chemical fluorogenic probe 7Az4Mc and both in spectrophotometer and fluorescence microscope we detected an increase of intracellular H_2S levels in the absence of PATZ1 protein. This result supports the previous findings and proves that the PATZ1 TF regulates the expression of Cth gene in HCT116 human colon cancer cell line. Considering that the CBS enzyme is the main H_2S producing enzyme in HCT116 colon cancer cell line, these results can direct researchers towards an alternative approach to study the regulation of intracellular H_2S levels in HCT116 cells.

Figure 35. Model for PATZ1 TF binding to the promoter region of the Cth gene to suppress the expression of the gene and possible mechanism of competition between suppressor PATZ1 and activator SP1 transcription factors.

In conclusion our current study identifies transcriptional target genes of the human PATZ1 transcription factor in the human colon cancer cell line HCT116. For the first time in the literature, we demonstrated that the human PATZ1 TF binds to and negatively regulates the promoter of Cth gene which encodes the H_2S producing enzyme CTH. Our study also supports the hypothesis that PATZ1 is a context dependent transcription factor which acts both as an activator and a repressor on its target genes.

6. REFERENCES

1. Pan G, Ameur A, Enroth S, et al. PATZ1 down-regulates FADS1 by binding to rs174557 and is opposed by SP1/SREBP1c. *Nucleic Acids Res*. 2017;45(5):2408-2422. doi:10.1093/nar/gkw1186

2. Cheatle Jarvela AM, Hinman VF. Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. *Evodevo*. 2015;6(1). doi:10.1186/2041-9139-6-3

3. Luscombe NM, Austin SE, Berman HM, Thornton JM. An overview of the structures of protein-DNA complexes. *Genome Biol*. 2000;1(1):1-37. doi:10.1186/gb-2000-1-1-reviews001

4. Gibson TJ, Postma JPM, Brown RS, Argos P. A model for the tertiary structure of the 28 residue dna-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins. *Protein Eng Des Sel*. 1988;2(3):209-218. doi:10.1093/protein/2.3.209

5. Razin S V., Borunova V V., Maksimenko OG, Kantidze OL. Cys2his2 zinc finger protein family: Classification, functions, and major members. *Biochem*. 2012;77(3):217-226. doi:10.1134/S0006297912030017

6. Peng Y, Clark KJ, Campbell JM, Panetta MR, Guo Y, Ekker SC. Making designer mutants in model organisms. *Dev*. 2014;141(21):4042-4054. doi:10.1242/dev.102186

7. Urnov FD, Miller JC, Lee YL, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 2005;435(7042):646-651. doi:10.1038/nature03556

8. Dovat S, Ronni T, Russell D, Ferrini R, Cobb BS, Smale ST. A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains.

Genes Dev. 2002;16(23):2985-2990. doi:10.1101/gad.1040502

9. Lu D, Searles MA, Klug A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature*. 2003;426(6962):96-100. doi:10.1038/nature02088

10. Mackay JP, Crossley M. Zinc fingers are sticking together. *Trends Biochem Sci*. 1998;23(1):1-4. doi:10.1016/S0968-0004(97)01168-7

11. Cassandri M, Smirnov A, Novelli F, et al. Zinc-finger proteins in health and disease. *Cell Death Discov*. 2017;3(1). doi:10.1038/cddiscovery.2017.71

12. Kim SH, Kim EJ, Hitomi M, et al. The LIM-only transcription factor LMO2 determines tumorigenic and angiogenic traits in glioma stem cells. *Cell Death Differ*. 2015;22(9):1517-1525. doi:10.1038/cdd.2015.7

13. Harrison SD, Travers AA. The tramtrack gene encodes a Drosophila finger protein that interacts with the ftz transcriptional regulatory region and shows a novel embryonic expression pattern. *EMBO J*. 1990;9(1):207-216.

14. DiBello PR, Withers DA, Bayer CA, Fristrom JW, Guild GM. The Drosophila Broad-Complex encodes a family of related proteins containing zinc fingers. *Genetics*. 1991;129(2):385-397.

15. Godt D, Couderc JL, Cramton SE, Laski FA. Pattern formation in the limbs of Drosophila: bric à brac is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus. *Development*. 1993;119(3):799-812.

16. Numoto M, Niwa O, Kaplan J, et al. Transcriptional repressor ZF5 identifies a new conserved domain in zina finger proteins. *Nucleic Acids Res*. 1993;21(16):3767-3775. doi:10.1093/nar/21.16.3767

17. Bardwell VJ, Treisman R. The POZ domain: a conserved protein-protein interaction motif. *Genes Dev*. 1994;8(14):1664-1677.

18. Stogios PJ, Downs GS, Jauhal JJS, Nandra SK, Privé GG. Sequence and structural analysis of BTB domain proteins. *Genome Biol*. 2005;6(10):1-18. doi:10.1186/gb-2005-6-10-r82

19. Perez-Torrado R, Yamada D, Defossez PA. Born to bind: The BTB protein-protein interaction domain. *BioEssays*. 2006;28(12):1194-1202. doi:10.1002/bies.20500

20. Katsani KR, Hajibagheri MAN, Verrijzer CP. Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology. *EMBO J*. 1999;18(3):698-708. doi:10.1093/emboj/18.3.698

21. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev*. 2012;247(1):172-183. doi:10.1111/j.1600-065X.2012.01112.x

22. Fedele M, Crescenzi E, Cerchia L. The POZ/BTB and AT-hook containing zinc finger 1 (PATZ1) transcription regulator: Physiological functions and disease involvement. *Int J Mol Sci*. 2017;18(12). doi:10.3390/ijms18122524

23. Cho JH, Kim MJ, Kim KJ, Kim JR. POZ/BTB and AT-hook-containing zinc finger protein 1 (PATZ1) inhibits endothelial cell senescence through a p53 dependent pathway. *Cell Death Differ*. 2012;19(4):703-712. doi:10.1038/cdd.2011.142

24. Sakaguchi S, Hombauer M, Bilic I, et al. The zinc-finger protein MAZR is part of the transcription factor network that controls the CD4 versus CD8 lineage fate of double-positive thymocytes. *Nat Immunol*. 2010;11(5):442-448. doi:10.1038/ni.1860

25. Kobayashi a, Yamagiwa H, Hoshino H, et al. A combinatorial code for gene expression generated by transcription factor Bach2 and MAZR (MAZ-related factor) through the BTB/POZ domain. *Mol Cell Biol*. 2000;20(5):1733-1746.

26. Fedele M, Benvenuto G, Pero R, et al. A novel member of the BTB/POZ family, PATZ, associates with the RNF4 RING finger protein and acts as a transcriptional repressor. *J Biol Chem*. 2000;275(11):7894-7901.

27. Ow JR, Ma H, Jean A, et al. Patz1 regulates embryonic stem cell identity. *Stem Cells Dev*. 2014;23(10):1062-1073. doi:10.1089/scd.2013.0430

28. Sakaguchi S, Hombauer M, Bilic I, Naoe Y. The zinc finger protein MAZR is part of the transcription factor network controlling CD4 / CD8 cell fate decision of DP thymocytes. *Nat Inmunol*. 2010;11(5):442-448. doi:10.1038/ni.1860.

29. Bilic I, Koesters C, Unger B, et al. Negative regulation of CD8 expression via Cd8 enhancer-mediated recruitment of the zinc finger protein MAZR. *Nat Immunol*. 2006;7(4):392-400. doi:10.1038/ni1311

30. Valentino T, Palmieri D, Vitiello M, et al. Embryonic defects and growth alteration in mice with homozygous disruption of the Patz1 gene. *J Cell Physiol*. 2013;228(3):646-653. doi:10.1002/jcp.24174

31. Mancinelli S, Vitiello M, Donnini M, et al. The Transcription Regulator Patz1 Is Essential for Neural Stem Cell Maintenance and Proliferation. *Front Cell Dev Biol*. 2021;9(April):1-15. doi:10.3389/fcell.2021.657149

32. Schilder BM, Raj T. Fine-mapping of Parkinson's disease susceptibility loci identifies putative causal variants. *Hum Mol Genet*. 2021:0-30. doi:10.1093/hmg/ddab294

33. Guadagno E, Vitiello M, Francesca P, et al. PATZ1 is a new prognostic marker of glioblastoma associated with the stem-like phenotype and enriched in the proneural subtype. 2017;8(35):59282-59300.

34. Cui H, Bansal V, Grunert M, et al. Muscle-relevant genes marked by stable H3K4me2/3 profiles and enriched MyoD binding during myogenic differentiation. *PLoS One*. 2017;12(6):1-19. doi:10.1371/journal.pone.0179464

35. Mastrangelo T, Modena P, Tornielli S, et al. A novel zinc finger gene is fused to EWS in small round cell tumor. *Oncogene*. 2000;19(33):3799-3804. doi:10.1038/sj.onc.1203762

36. Zhao C, Yan M, Li C, Feng Z. POZ/BTB and AT-hook-containing zinc finger protein 1 (PATZ1) suppresses progression of ovarian cancer and serves as an independent prognosis factor. *Med Sci Monit*. 2018;24:4262-4270. doi:10.12659/MSM.908766

37. Ng ZL, Siew J, Li J, et al. PATZ1 (MAZR) Co-occupies Genomic Sites With p53 and Inhibits Liver Cancer Cell Proliferation via Regulating p27. *Front Cell Dev Biol*. 2021;9(February):1-16. doi:10.3389/fcell.2021.586150

38. Tian X, Sun D, Zhang Y, Zhao S, Xiong H, Fang J. Zinc finger protein 278, a potential oncogene in human colorectal cancer. *Acta Biochim Biophys Sin (Shanghai)*. 2008;40(4):289-296. doi:10.1111/j.1745-7270.2008.00405.x

39. Wallace JL, Wang R. Hydrogen sulfide-based therapeutics: Exploiting a unique but ubiquitous gasotransmitter. *Nat Rev Drug Discov*. 2015;14(5):329-345. doi:10.1038/nrd4433

40. Coletta C, Papapetropoulos A, Erdelyi K, et al. Hydrogen sulfide and nitric oxide are mutually dependent in the regulation of angiogenesis and endothelium-dependent vasorelaxation. *Proc Natl Acad Sci U S A*. 2012;109(23):9161-9166. doi:10.1073/pnas.1202916109

41. Bhatia M. Hydrogen sulfide and substance P in inflammation. *Antioxidants Redox Signal*. 2010;12(10):1191-1202. doi:10.1089/ars.2009.2927

42. Ma DK, Vozdek R, Bhatla N, Horvitz HR. CYSL-1 Interacts with the O 2-Sensing Hydroxylase EGL-9 to Promote H 2S-Modulated Hypoxia-Induced Behavioral Plasticity in C. elegans. *Neuron*. 2012;73(5):925-940. doi:10.1016/j.neuron.2011.12.037

43. Polhemus DJ, Lefer DJ. Emergence of Hydrogen Sulfide as an Endogenous Gaseous Signaling Molecule in Cardiovascular Disease David. *Changes*. 2012;29(6):997-1003. doi:10.1161/CIRCRESAHA.114.300505.Emergence

44. Mani S, Li H, Untereiner A, et al. Decreased endogenous production of hydrogen sulfide accelerates atherosclerosis. *Circulation*. 2013;127(25):2523-2534. doi:10.1161/CIRCULATIONAHA.113.002208

45. Robert Cronin Yung Peng, Rose Khavari ND. Cystathionine γ-lyase deficiency mediates neurodegeneration in Huntington's disease. *Physiol Behav*. 2017;176(3):139-148. doi:10.1159/000444169.Carotid

46. Szabõ C. Hydrogen sulphide and its therapeutic potential. *Nat Rev Drug Discov*. 2007;6(11):917-935. doi:10.1038/nrd2425

47. Sbodio JI, Snyder SH, Paul BD. Regulators of the transsulfuration pathway. *Br J Pharmacol*. 2019;176(4):583-593. doi:10.1111/bph.14446

48. Doeller JE, Isbell TS, Benavides G, et al. Polarographic measurement of hydrogen sulfide production and consumption by mammalian tissues. *Anal Biochem*. 2005;341(1):40-51. doi:10.1016/j.ab.2005.03.024

49. Cao X, Ding L, Xie ZZ, et al. A Review of Hydrogen Sulfide Synthesis, Metabolism, and Measurement: Is Modulation of Hydrogen Sulfide a Novel Therapeutic for Cancer? *Antioxidants Redox Signal*. 2019;31(1). doi:10.1089/ars.2017.7058

50. Paul BD, Snyder SH. H2S: A Novel Gasotransmitter that Signals by Sulfhydration. *Trends Biochem Sci*. 2015;40(11):687-700. doi:10.1016/j.tibs.2015.08.007

51. Li S, Yang G. Hydrogen Sulfide Maintains Mitochondrial DNA Replication via Demethylation of TFAM. *Antioxidants Redox Signal*. 2015;23(7):630-642. doi:10.1089/ars.2014.6186

52. Sen N, Paul BD, Gadalla MM, et al. Hydrogen sulfide-linked sulfhydration of NF-κB mediates its antiapoptotic actions. *Mol Cell*. 2012;45(1):13-24. doi:10.1016/j.molcel.2011.10.021

53. Mosharov E, Cranford MR, Banerjee R. The Quantitatively Important Relationship between Homocysteine Metabolism and Glutathione Synthesis by the Transsulfuration Pathway and Its Regulation by Redox. 2000;39:13005-13011.

54. Ishii I, Akahoshi N, Yamada H, Nakano S, Izumi T, Suematsu M. Cystathionine γ-lyase-deficient mice require dietary cysteine to protect against acute lethal myopathy and oxidative injury. *J Biol Chem*. 2010;285(34):26358-26368. doi:10.1074/jbc.M110.147439

55. Frimpter WG, Haymovitz A, Horwith M. Cystathioninuria. *N Engl J Med*. 1963;268(7):268-339.

56. Yin P, Zhao C, Li Z, et al. Sp1 is involved in regulation of cystathionine γ -lyase gene expression and biological function by PI3K / Akt pathway in human hepatocellular carcinoma cell lines. 2012;24:1229-1240. doi:10.1016/j.cellsig.2012.02.003

57. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57-74. doi:10.1038/nature11247

58. Saiz-Baggetto S, Méndez E, Quilis I, Igual JC, Bañó MC. Chimeric proteins tagged with specific 3xHA cassettes may present instability and functional problems. *PLoS One*. 2017;12(8):1-12. doi:10.1371/journal.pone.0183067

59. Keskin N, Deniz E, Eryilmaz J, et al. PATZ1 Is a DNA Damage-Responsive Transcription Factor That Inhibits p53 Function. *Mol Cell Biol*. 2015;35(10):1741-1753. doi:10.1128/MCB.01475-14

60. Cai W, Wang M, Ju L, Wang C, Zhu Y. Hydrogen sulfide induces human colon cancer cell proliferation: role of Akt, ERK and p21. *Cell Biol Int*. 2010;34(6):565-572. doi:10.1042/cbi20090368

61. Wu D, Li M, Tian W, et al. Hydrogen sulfide acts as a double-edged sword in human hepatocellular carcinoma cells through EGFR/ERK/MMP-2 and PTEN/AKT signaling pathways. *Sci Rep*. 2017;7(1):1-14. doi:10.1038/s41598-017-05457-z

62. Shields JM, Christy RJ, Yang VW. Identification and Characterization of a Gene Encoding a Gut- enriched Krüppel-like Factor Expressed during Growth Arrest. 1996;271(33):20009-20017.

63. Katz JP, Perreault N, Goldstein BG, et al. The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. 2002;129(11):2619-2628.

64. Bermejo-rodrı C, Pe PA, Sa I. Adipose tissue mass is modulated by SLUG (SNAI2). 2007;16(23):2972-2986. doi:10.1093/hmg/ddm278

65. Marianayagam NJ, Sunde M, Matthews JM. The power of two : protein dimerization in biology. 2004;29(11). doi:10.1016/j.tibs.2004.09.006

66. Zullo JM, Gaffney DJ, Epstein CB, et al. DNA Sequence-Dependent Compartmentalization and Silencing of Chromatin at the Nuclear Lamina. 2012;149:1474-1487. doi:10.1016/j.cell.2012.04.035

67. Maeda T. Regulation of hematopoietic development by ZBTB transcription factors Linker region. *Int J Hematol*. 2016;104(3):310-323. doi:10.1007/s12185-016-2035-x

7. APPENDIX

7.1. APPENDIX A: Chemicals Used in the Study

Chemicals and Media Components	Supplier Company
2-Mercaptoethanol	Sigma, Germany
Acetic Acid	Merck, Germany
Acid Washed Glass Beads	Sigma, Germany
Acrylamide/Bis-acrylamide	Sigma, Germany
Agarose	peQLab, Germany
Ammonium Persulfate	Sigma, Germany
Ammonium Sulfate	Sigma, Germany
Ampicillin Sodium Salt	CellGro, USA
Bacto Agar	BD, USA
Bacto Tryptone	BD, USA
Boric Acid	Molekula, UK
Bradford Reagent	Sigma, Germany
Bromophenol Blue	Sigma, Germany
Chloramphenicol	Gibco, USA
D-Glucose	Sigma, Germany
Distilled water	Milipore, France
DMEM	PAN, Germany
DMSO	Sigma, Germany
DNA Gel Loading Solution, 5X	Quality Biological, Inc, USA
DPBS	CellGro, USA
EDTA	Applichem, Germany
Ethanol	Riedel-de Haen, Germany
Ethidium Bromide	Sigma, Germany
Fetal Bovine Serum (FBS)	Biological Industries, Israel
Glycerol Anhydrous	Applichem, Germany
Glycine	Applichem, Germany
HBSS	CellGro, USA

HEPES	Applichem, Germany
Hydrochloric Acid	Merck, Germany
Isopropanol	Riedel-de Haén, Germany
Kanamycin Sulfate	Gibco, USA
LB Agar	BD, USA
LB Broth	BD, USA
L-Glutamine	Hyclone, USA
Liquid nitrogen	Karbogaz, Turkey
Magnesium Chloride	Promega, USA
Methanol	Riedel-de Haen, Germany
Monoclonal Anti-HA Antibody	Sigma, Germany
p53 Mouse Antibody	Cell Signalling Technology, USA
p53 Polyclonal Antibody	Cell Signalling Technology, USA
PATZ1 (H-300) Antibody	Santa Cruz, USA
Penicillin-Streptomycin	Sigma, Germany
Phenol-Chloroform-Isoamylalcohol	Amersco, USA
PIPES	Sigma, Germany
Potassium Acetate	Merck, Germany
Potassium Chloride	Fluka, Germany
Potassium Hydroxide	Merck, Germany
Protease Tablets (EDTA-free)	Roche, Germany
ProtG Sepharose	Amersco, USA
RNase A	Roche, Germany
RPMI 1640	PAN, Germany
SDS Protein Gel Loading Pack	Fermentas, Germany
SDS Pure	Applichem, Germany
Skim Milk Powder	Fluka, Germany
Sodium Azide	Amresco, USA
Sodium Chloride	Applichem, Germany
SuperSignal Chemiluminescent Substrate	Thermo Scientific, USA
TEMED	Applichem, Germany
Tris Buffer Grade	Amresco, USA
Tris Hydrochloride	Amresco, USA

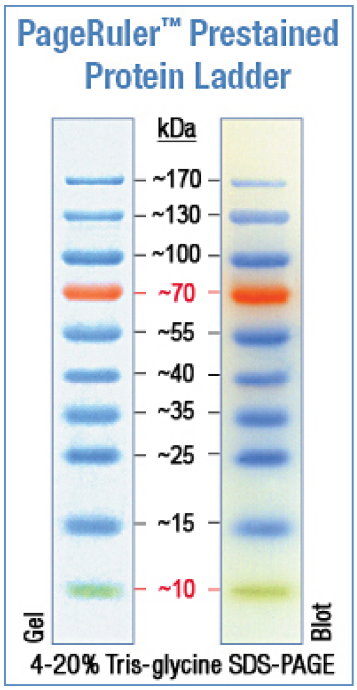
Triton X100	Promega, USA
Tween20	Sigma, Germany

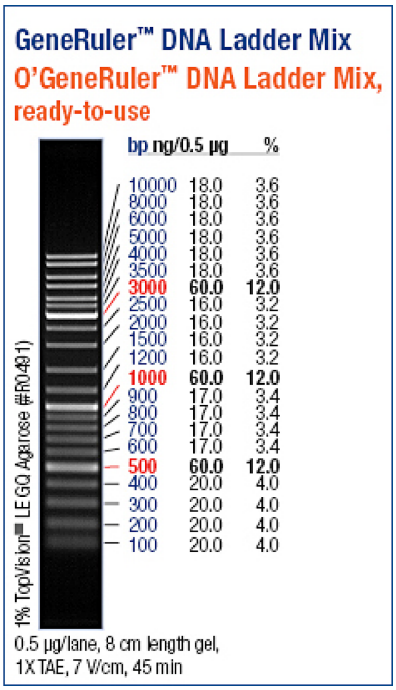
7.2. APPENDIX B: Equipment Used in the Study

Equipment	Company
Autoclave	Hirayama,Hiclave HV-110,Japan
Balance	Sartorius, BP221S, Germany
	Schimadzu, Libror EB-3200 HU, Japan
Biomolecular Imager	ImageQuant LAS 4000 mini - GE Healthcare, USA
Cell Counter	Cole Parmer, USA
Centrifuge	Eppendorf, 5415D, Germany
	Hitachi, Sorvall RC5C Plus, USA
CO2 Incubator	Binder,Germany
Deepfreeze	-80oC, Forma,Thermo ElectronCorp.,USA
	-20oC,Bosch,Turkey
Distilled Water	Millipore, Elix-S, France
Electrophoresis Apparatus	Biogen Inc., USA
Electroporator	Neon Transfection System - Life Technologies, USA
Filter Membranes	Millipore,USA
Flow Cytometer	BDFACSCanto,USA
Gel Documentation	Biorad, UV-Transilluminator 2000, USA
Heater	ThermomixerComfort,Eppendorf,Germany
Hematocytometer	Hausser Scientific,Blue Bell Pa.,USA
Ice Machine	Scotsman Inc., AF20, USA
Incubator	Memmert, Modell 300, Germany
Laminar Flow	Kendro Lab. Prod., Heraeus, HeraSafe HS12, Germany
Liquid Nitrogen Tank	Taylor-Wharton,3000RS,USA
Magnetic Stirrer	VELP Scientifica, ARE Heating Magnetic Stirrer, Italy
Microliter Pipettes	Gilson, Pipetman, France
	Eppendorf, Germany
Microscope	Olympus CK40,Japan
	Olympus CH20,Japan
	Olympus IX70,Japan
	Zeiss Confocal LSM710, German

Microwave Oven	Bosch,Turkey
pH meter	WTW, pH540 GLP MultiCal, Germany
Power Supply	Biorad, PowerPac 300, USA
Refrigerator	Bosch,Turkey
Shaker Incubator	New Brunswick Sci., Innova 4330, USA
Spectrophotometer	Schimadzu, UV-1208, Japan
	Schimadzu, UV-3150, Japan
Thermocycler	Eppendorf, Mastercycler Gradient, Germany
Vortex	Velp Scientifica,Italy

7.3. APPENDIX C: DNA and Protein Molecular Weight Markers





7.4. APPENDIX D: List of the DEGs from RNA-Seq analysis in HCT116 WT vs. PATZ1 KO cells.

Gene	Chromosome	Region	Fold change
SMIM22	16	4788397..4796491	12.0142678
RIMKLB	12	8681600..8783095	8.280685953
CALB1	8	complement(90058608..90095475)	7.112263257
COLEC11	2	3594832..3644644	5.858774921
ANGPT2	8	complement(6499651..6563409)	5.243966717
LINC01405	12	110936585..110958208	5.002712327
TACSTD2	1	complement(58575423..58577773)	4.844054341
RND3	2	complement(150468195..150539011)	4.815577387
CSTA	3	122325244..122341972	4.261305574
H19	11	complement(1995163..2001470)	4.247415436
KRT7	12	52232520..52252186	4.230342128
STXBP6	14	complement(24809656..25050297)	4.1503115
ALPPL2	2	232406843..232410714	3.968850981
KLK11	19	complement(51022216..51028039)	3.959207479
EXO5	1	40508741..40516556	3.814230558
NCAM1	11	112961247..113278436	3.665520018
AC011747.4	2	complement(8559833..8583792)	3.636764794
ZBED2	3	complement(111592900..111595443)	3.555221974
AL162759.1	14	complement(53768942..53850882)	3.507143287
TMEM139	7	143279957..143288048	3.461628922
KLRC4-KLRK1	12	complement(10372353..10410146)	3.203144934
LY6D	8	complement(142784880..142786592)	3.133169493
RP11-114H23.2	12	complement(75694010..75698816)	2.981628789
ANKFN1	17	56110958..56511659	2.937238852
URGCP-MRPS24	7	complement(43866558..43906589)	2.926082317
BMP4	14	complement(53949736..53958761)	2.789624625
AL136419.6	14	complement(24209646..24215987)	2.777587429
KRT80	12	complement(52168996..52192000)	2.768743917
EDAR	2	complement(108894471..108989372)	2.660330682
UBE2Q2P6_1	15	complement(82372196..82372912)	2.644733383

KCNK9	8	complement(139600838..139703056)	2.640910767
PVRL4	1	complement(161070995..161089599)	2.567650471
ZFP42	4	187995771..188005050	2.564670215
LTN1	21	complement(28928144..28992956)	2.556273494
TNNC1	3	complement(52451102..52454070)	2.387252887
PRKCQ	10	complement(6427143..6580301)	2.385208456
BDNF	11	complement(27654893..27722058)	2.362263604
GPR87	3	complement(151294086..151316952)	2.345448242
TMEM44-AS1	3	194584011..194590260	2.320022273
CASC10	10	complement(21492658..21497262)	2.245193925
AGAP11	10	86970741..87010126	2.21359022
LINC01468	10	complement(52450874..52470533)	2.209192073
C1orf116	1	complement(207018521..207032756)	2.20904747
HLA-DQB1	6	complement(32659467..32668383)	2.206857765
RPL36A-HNRNPH2	X	101391011..101412297	2.202280821
COL4A6	X	complement(108155607..108439497)	2.195218041
RBAK	7	5045821..5069488	2.189487184
LOXL4	10	complement(98247690..98268250)	2.185340442
SKIDA1	10	complement(21513478..21526368)	2.182028854
MEX3B	15	complement(82041778..82046141)	2.181873275
TRIM29	11	complement(120111275..120185529)	2.137716312
RP11-114H23.1	12	complement(75563202..75984015)	2.124607646
WDR66	12	121917862..122003927	2.111656146
MROH7-TTC4	1	54641786..54742308	2.102738642
NOSTRIN	2	168786539..168865514	2.092711655
RP5-940J5.9	12	complement(6537794..6538370)	2.091423484
WNT10A	2	218880363..218899581	2.088784391
ADGRF1	6	complement(46997703..47042363)	2.077689339
SPATA18	4	52051331..52097292	2.077570896
SOX4	6	21592769..21598619	2.067785427
HIST1H2BD	6	26158146..26171349	2.049532808
ALOX5	10	45374176..45446119	2.047442444
CTH	1	70411218..70439851	2.045779233

ARPC4-TTLL3	3	9793082..9835401	2.040861324
TBC1D3D	17	37924458..38139051	2.022674552
TGM1	14	complement(24249114..24264432)	2.020674629
CTHRC1	8	103371515..103382997	2.015629425
ISPD	7	complement(16087527..16421322)	2.007317632
EFNA2	19	1286154..1300237	-2.024073442
POPDC3	6	complement(105158280..105179995)	-2.054104911
UHRF2	9	6413151..6507054	-2.059562087
KCNS3	2	17877847..18361616	-2.065337408
CHRM4	11	complement(46385098..46386608)	-2.073706373
TNIK	3	complement(171061339..171460408)	-2.082724832
CALB2	16	71358713..71390438	-2.11410147
MBNL3	X	complement(132369317..132489968)	-2.119604334
NR3C2	4	complement(148078762..148444698)	-2.13596851
INSIG1	7	155297776..155310235	-2.137731321
ARHGAP29	1	complement(94148988..94275068)	-2.141575625
PDLIM7	5	complement(177483394..177497606)	-2.159971502
HCN2	19	589893..617159	-2.161172997
NR2F1	5	93583337..93594615	-2.174870965
RAVER1	19	complement(10316212..10333640)	-2.176434159
GLIS3	9	complement(3824127..4348392)	-2.17843338
ACOX2	3	complement(58505136..58537319)	-2.202277265
PDGFB	22	complement(39223359..39244751)	-2.223246649
ZDHHC2	8	17156029..17224799	-2.235534595
AMIGO2	12	complement(47075707..47079951)	-2.261216762
SAMD5	6	147508927..147737547	-2.287585825
VSNL1	2	17539126..17657018	-2.297860373
CORO2B	15	68578969..68727806	-2.300494879
RAB27B	18	54717860..54895516	-2.328976377
RUNX2	6	45328157..45664349	-2.349865521
CEMIP	15	80779343..80951776	-2.365795672
SPP1	4	87975650..87983426	-2.371371957
RP1-37E16.12_2	22	37734401..37755656	-2.378049164
ZEB1	10	31318495..31529814	-2.383348431
LINC01270	20	50292720..50314922	-2.408668112
HOXA1	7	complement(27092993..27095996)	-2.424666758
KLF2	19	16324817..16327874	-2.443526951

VCAN	5	83471465..83582303	-2.444780333
MMP14	14	22836557..22849027	-2.451601154
ABCB1	7	complement(87503633..87713323)	-2.475131092
SNTB1	8	complement(120535745..120813273)	-2.485953182
ZNF525	19	53365693..53392217	-2.508765695
TRHDE	12	72087266..72670757	-2.589212133
HSP90AA1	14	complement(102080738..102139699)	-2.591680972
COX6B2	19	complement(55349306..55354814)	-2.713929301
SAMD9	7	complement(93099513..93118023)	-2.722362371
ETV1	7	complement(13891228..13991425)	-2.75141531
WNT16	7	121325367..121341104	-2.857929909
NAV3	12	77830905..78213008	-2.911371841
NPIPA8	16	complement(18317942..18336736)	-2.940006261
METTL7A	12	50923472..50932517	-2.958304226
FGF2	4	122826708..122898236	-3.056652648
BHLHE41	12	complement(26120026..26125127)	-3.057909025
DCLK1	13	complement(35768652..36131306)	-3.275659724
FAM198B	4	complement(158124474..158173318)	-3.284789408
TGFB2	1	218346235..218444619	-3.569604006
DSC3	18	complement(30990008..31042815)	-3.63756802
HOXA3	7	complement(27106184..27152581)	-3.647267733
TTN	2	complement(178525989..178830802)	-3.70802307
VIM	10	17228259..17237593	-3.762379539
RP11-977G19.10	12	complement(56285916..56316059)	-3.768844758
NRP1	10	complement(33177492..33336262)	-3.84073177
CD33	19	51225064..51243860	-3.879491789
RTEL1-TNFRSF6B	20	63659300..63698698	-3.905799282
SENP3-EIF4A1	17	7563287..7578715	-3.937997532
ZNF391	6	27374615..27403904	-3.939222846
ADAM19	5	complement(157395534..157575775)	-3.983654625
CPED1	7	120988677..121297444	-4.253970549
NT5E	6	85449584..85495791	-4.543502033
EPS8L3	1	complement(109750080..109764027)	-4.763357537
PPP1R1B	17	39626740..39636626	-4.798510726
CD68	17	7579467..7582113	-4.827439956
CAPN6	X	complement(111245103..111270523)	-4.827847508

SEMA3A	7	complement(83955777..84492724)	-5.172671403
MILR1	17	64449037..64468643	-5.28592432
FAM131B	7	complement(143353400..143362770)	-5.384622841
EFNA5	5	complement(107376889..107670895)	-5.783385652
C10orf10	10	complement(44970981..44978810)	-5.909767127
ZNF426	19	complement(9523224..9538627)	-6.379156936
ZNF804A	2	184598366..184939492	-6.786936576
SPARC	5	complement(151661096..151687165)	-6.871027033
PLA2R1	2	complement(159932006..160062610)	-7.085747004
COL6A3	2	complement(237324003..237414375)	-7.458933563
TMEM200A	6	130365734..130443063	-7.479167322
KIRREL2	19	35855861..35867109	-7.756616656
UTS2B	3	complement(191267168..191330536)	-7.973780553
SAMD3	6	complement(130144315..130365425)	-8.524621092
RGS4	1	163068775..163076802	-8.920229221
ABCA1	9	complement(104781002..104928237)	-8.921188527
MNS1	15	complement(56421544..56465137)	-13.47273161
CACNB4	2	complement(151832776..152099079)	-15.35550793
LARGE	22	complement(33162226..33922841)	-19.62419399
TTC22	1	complement(54779712..54801267)	-28.53254707