The Epigenetic Silencing of ALKBH3 and the Epitranscriptomic Regulation of DNA Repair

Stefan Thor Hermanowicz

Thesis for the degree of Philosophiae Doctor

Supervisors:

Stefán Þ. Sigurðsson, Ph.D.
Þorkell Guðjónsson, Ph.D.

Doctoral committee:

Ólafur A. Stefánsson, Ph.D.
Claudia Lukas, Ph.D.
Jón G. Jónasson, MD

May 2019

UNIVERSITY OF ICELAND
SCHOOL OF HEALTH SCIENCES
FACULTY OF MEDICINE

Sviperfðabreytingar á ALKBH3 og áhrif sviperfða á RNA í stjórnun á DNA viðgerð

Stefán Þór Hermanowicz

Ritgerð til doktorsgráðu

Leiðbeinendur:

Stefán Þ. Sigurðsson Ph.D.,
Þorkell Guðjónsson Ph.D.

Doktorsnefnd:

Ólafur A. Stefánsson Ph.D.
Claudia Lukas Ph.D.
Jón G. Jónasson MD

Maí 2019

UNIVERSITY OF ICELAND
SCHOOL OF HEALTH SCIENCES
FACULTY OF MEDICINE

ISBN 978-9935-9455-9-4

Printing by Háskólaprent.

Reykjavik, Iceland 2019

Ágrip

Eitt af megin einkennum krabbameinsfruma er óstöðugt erfðamengi. Þessi óstöðugleiki er talinn stafa af uppsöfnun DNA skemmda sem fruman nær ekki að gera við. Til að koma í veg fyrir að þetta gerist býr fruman yfir mjög öflugu eftirlitskerfi sem skynjar og bregst við DNA skemmdum. Gallar í mikilvægum genum innan þessa kerfis eru algengir í krabbameinsfrumum. Í sumum tilfellum er hægt að nota slíka galla til að aðgreina krabbameinsfrumur frá heilbrigðum frumum, og nýta í lyfjameðferð. ALKBH3 er díoxigenasi sem tekur þátt í viðgerð á alkýlerandi DNA skemmdum. Rannsóknir okkar á gögnum frá ”The Cancer Genome Atlas” sýndu að stýrilsvæði ALKBH3 gensins er metýlerað í um 20 % brjóstakrabbameina. Þetta hefur í för með sér minnkaða ALKBH3 mRNA tjáningu og þar með minni framleiðslu á ALKBH3 próteini. Greining á sýnum úr íslenska þýðinu leiddu í ljós að tíðni ALKBH3 metýlunar er um 10 % í brjóstakrabbameini. Samanburður á sýnum úr heilbrigðum vef og krabbameinsvef úr sama einstaklingi sýndu að ALKBH3 metýlun greinist bara í krabbmeinsvef. Þar að auki er lifun verri hjá þeim brjóstakrabbameins-sjúklingum sem hafa mikið magn metýlunar á ALKBH3 stýrilsvæðinu. Frekari rannsóknir sýndu að tap á ALKBH3 tjáningu leiddi til minni tjáningar á RNF168, sem er mikilvægur stjórnþáttur í viðgerð á mjög alvarlegri gerð DNA skemmda, tvíþátta DNA rofi. Enn frekar gefa niðurstöður þessa verkefnis til kynna að ALKBH3 hafi áhrif á virkni RNF168 með því að að fjarlæga metýl hópa af mRNA sameind RNF168. Þetta er í fyrsta skipti sem sýnt er fram á að RNF168 er stjórnað með slíkum hætti.

Það að metýlun á mRNA sameindum geti haft áhrif á tjáningu próteina er tiltölulega ný uppgötvun sem hefur á skömmum tíma vakið mikla athygli. Sýnt hefur verið fram á að með því að bæta við eða fjarlægja metýl hópa á ákveðnum kirnum innan mRNA sameindarinnar er hægt að hafa áhrif á tjáningu próteina í gegnum nokkrar ólíkar leiðar. Þar má nefna RNA splæsingu, flutning mRNA sameinda úr kjarna og þýðingu mRNA í prótein. Greining á RNA metýlun gefa til kynna aukið magn N1-metýl-adenósín breytinga á RNF168 mRNA sameindinni eftir að tjáning ALKBH3 hefur verið bæld. Nánari greining á ALKBH3 bældum frumum leiddi í ljós aukið magn RNF168 mRNA sameinda í kjarna sem gefur til kynna vandamál við flutning út úr kjarna en þýðing mRNA sameinda í prótein á sér stað í umfrymi. Því má færa sterk rök fyrir því að ALKBH3 hafi áhrif á virkni RNF168 með því að stuðla að útflutningi RNF168

mRNA sameindarinnar úr kjarna í umfrymi, sem er nauðsynlegt fyrir RNF168 próteinmyndun. Þetta er í fyrsta skipti sem sýnt er fram á að DNA viðgerðarpróteini sé stjórnað með mRNA breytingum. Því er mikilvægt að skilgreina þessa áður óþekktu aðferð sem ALKBH3 notar til að stjórna genatjáningu til að auka skilning okkar á líffræðilegum orsökum krabbameina. Þar að auki má hugsanlega nýta þessa nýju þekking á hlutverki ALKBH3 í þróun nýrra meðferðarúrræða.

Lykilorð:

DNA viðgerð, sviperfðir, mRNA breytingar, brjóstakrabbamein, ALKBH3.

Abstract

The DNA damage response is crucial to maintaining the integrity of DNA and the health of a cell. Unrepaired lesions within the DNA can lead to genomic instability and potentially aid in the formation of diseases such as cancer. Some cancers possess dysfunctional DNA repair and chemotherapeutic treatments may aim to exploit this weakness that distinguishes cancer cells from normal healthy cells. ALKBH3 is a DNA repair protein involved in the repair of alkylation damage. Within The Cancer Genome Atlas, ALKBH3 displayed a hyper-methylated promoter in 20% of breast cancers. This hyper-methylation, a form of epigenetic regulation, lead to a dramatic reduction of ALKBH3 mRNA expression and therefore a decrease in total ALKBH3 protein levels. Within a sample of Icelandic breast tumors, the incidence of promoter methylation was 10%. Importantly, this methylation occurred only within the tumor tissue, but not the normal tissue of the same patients. Additionally, patients who contained high levels of promoter methylation had statistically significant decreased survival. When exploring the functional consequences of ALKBH3 silencing, the knockdown of ALKBH3 was found to cause a decrease in protein levels of RNF168, a protein crucial in ubiquitin signaling and effective DNA double-strand break repair. By eliminating changes in other forms of regulation, it was determined that RNF168 is being regulated by ALKBH3 through RNA methylation, a novel form of regulation of RNF168.

RNA methylation is an emerging field of control for protein expression. The addition and removal of methyl groups to nucleotides within mRNA has been implicated in having roles within alternative splicing, mRNA export and translational efficiency, thus causing differing expression of proteins and affecting cells accordingly. Using RNA immunoprecipitation, an increased level of RNF168 mRNA was pulled down following ALKBH3 knockdown, indicating increased levels of N1-methyladenosine on the mRNA are responsible for the change in protein levels. Furthermore, ALKBH3 is regulating cellular localization of RNF168 mRNA by impacting nuclear export. Following ALKBH3 knockdown, RNF168 mRNA is retained in the nucleus. The lack of available cytoplasmic mRNA to be translated is interpreted as the reason for the reduction of RNF168 in the absence of ALKBH3. This is a novel form of regulation of DNA repair. This mRNA regulatory mechanism by ALKBH3 could help elucidate a potential contributing factor to cancer development as well as

provide a potential target for chemotherapeutic treatment.

Keywords:

DNA Repair, Epigenetics, Epitranscriptomics, Breast Cancer, ALKBH3.

Acknowledgements

I first want to thank my instructor Stefán Sigurðsson for his guidance for the past 7 years. I couldn‘t have asked for a more patient, thoughtful, and better mentor. His door was always open to discuss our work, help guide my thoughts and encourage me to think critically. He would also continue to think about any issue I approached him with and come back to me with further insights, showing that he truly cared about my thoughts and ideas. Through him I’ve learned more in the past years then I could have possibly hoped.

I would like to thank Þorkell Guðjónsson. Without him the project simply would not have been possible. His ideas and work on this project are immeasurable and helped drive the project forward at all points. He too was always available to discuss the project and ready to provide helpful guidance and input. His expertise, thoroughness, and creativity helped form my scientific mind and transform me into the scientist I am today.

I would like to thank Kritika Kirty. She joined the project at a crucial juncture and helped develop some of the most exciting aspects of it. She always listened to my craziest scientific ideas and put an immense amount of thought and work into helping me investigate them. Without her work and mind this project would not have been possible to complete.

I would like to thank my PhD committee members. Ólafur Andri Stefánsson gave me guidance and assistance in the epigenetics aspect of the project. His contributions were essential to starting the project in the first place and responsible for our first publication of the project. Claudia Lukas provided invaluable feedback at key points in the project. Her expertise and professionalism helped improve the work in many ways. Jón Gunnlaugur Jónasson provided valuable expertise in the pathology sections of the project and paper.

I would like to thank my fellow lab members, current and former. Antón Ameneiro-Álvarez, Birna Þorvaldsdóttir, Drífa Hrund Guðmundsdóttir, Elisabet Alexandra Frick, Gunnlaug Ruth Guðmundsdóttir, Jasper van der Horst, Linda Hrönn Sighvatsdóttir, Mary Jean Lerry Figueroa Sicat, and Sigrún Guðjónsdóttir all contributed to this project in some manner and showed an incredible ability to put up with all my antics over the past years. An additional thank you is extended to all members of the BMC. I learned something from

everyone I interacted with during my time here and will take these lessons with me to whatever I do next.

I would like to thanks RANNIS and Eggertssjóður for their funding of this project.

I want to thank all my friends and family that have supported me throughout this process. A special thanks to my old friend Rob Lowell for his assistance with making the decoder tool.

A final special thank you to Mom, Dad and Ásdís, for supporting me for my whole life.

Contents

List of abbreviations

53BP1: Tumor protein p53 binding protein (TP53BP1)
ALKBHX: AlkB Homolog X, Alpha-Ketoglutarate dependent Dioxygenase
ASCC3: Activating Signal Cointegrator 1 Complex Subunit 3
ATM: Ataxia-telangiectasia mutated
ATR: Ataxia Telangiectasia and Rad3 related protein
BLM: Bloom syndrome protein
BRCA1: Breast cancer 1
BRCA2: Breast cancer 2
CHD4: Chromodomain helicase DNA binding protein 4
CLIP: Cross-linking immuno-precipitation
DDR: DNA damage response
DNA: Deoxyribonucleic acid
DNA-PK: DNA-dependent protein kinase
DR-GFP: two differentially mutated GFP genes oriented as direct repeats and separated by a drug selection marker
DSB: Double-strand break
EXO1: Exonuclease 1
FTO: Fat Mass And Obesity-Associated Protein
GC: Gene conversion
HR: Homologous recombination
ICL: Interstrand Cross-link
KU70: Ku autoantigen, 70kDa
KU80: Ku autoantigen, 80kDa
lncRNA: Long non-coding RNA
m1A: N1-methyladenosine
m1Am: N1,2'-O-dimethyladenosine
m3C: 3-methylcytosine
m5C: 5-methylcytosine
m6A: N6-methyladenosine
m6Am: N-6,2-dimethyladenosine
MDC1: Mediator of DNA-damage checkpoint 1
METTL: Methyltransferase Like
MMEJ: Microhomology-mediated NHEJ
MMS: Methyl methanosulfonate
MRE11: Meiotic recombination 11
MRN: MRE11-RAD50-NBS1
mRNA: Messenger RNA
NHEJ: Non-homologues end joining
NBS1: Nijmegen breakage syndrome
NXF1; Nuclear Export Factor 1
PALB2: Partner and localizer of BRCA2
PARP: Poly-(ADP-ribose) polymerase
PCR: Polymerase chain reaction

PTIP: PAX interacting (with transcription-activation domain) protein 1
qPCR: Quantitative polymerase chain reaction
RIF1: Rap1-Interacting Factor 1
RING: Really Interesting New Gene
RAD51: RAD51 recombinase
RIDDLE: Radiosensitvity, immunodeficiency, dysmorphic features and learning difficulties
RIP: RNA-immunoprecipitation
RNA: Ribonucleic acid
RNAi: RNA interference
RNA-IP: RNA-immunoprecipitation
RNF8: Ring finger protein 8
RNF168: Ring finger protein 168
RPA: Replication protein A
SAM: S-Adenosyl methionine
siRNA: Short interfering RNA
SSA: Single strand annealing
SSB: Single-strand break
ssDNA: Single stranded DNA
SUMO: Small ubiquitin-like modifier
TRMT: tRNA Methyltransferase
tRNA: Transfer RNA
UBE: Ubiquitin conjugating enzyme
USP: Ubiquitin-specific protease
UTR: Untranslated Region
VEGF: Vascular Endothelial Growth Factor
WTAP: Wilms Tumor 1-Associating Protein
γ-H2AX: H2AX serine 139 phosphorylation

List of figures

List of tables

List of original papers

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals

I. **CpG promoter methylation of the ALKBH3 alkylation repair gene in breast cancer** Stefansson O.A., Hermanowicz S., van der Horst J., Hilmarsdottir H., Staszczak Z., Jonasson J.G., Tryggvadottir L., Gudjonsson T., & Sigurdsson S. BMC Cancer. 2017; 17: 469.

In addition, some unpublished data may be presented:

All papers are reprinted by kind permission of the publishers.

Declaration of contribution

Paper #1. CpG promoter methylation of the ALKBH3 alkylation repair gene in breast cancer. In this paper we describe the epigenetic silencing of ALKBH3 in breast cancer and the clinical consequences of such an event occurring. I performed cell line culture, qPCR of select samples, pyrosequencing, western blotting. Olafur Andri Stefansson was responsible for pyrosequencing, qPCR, stastistical, and informatics analyses. Thorkell Gudjonsson was responsible for developing and performing the m3C damage assay as well as western blots. Jasper van der Horst performed qPCR and pyrosequencing. Jon Gunnlaugur Jonasson and Laufey Tryggvadottir contributed information on clinical parameters. Jon Gunnlaugur Jonasson and Olafur Andri Stefansson were responsible for the TMA expression analyses. Olafur Andri Stefansson and Stefan Sigurdsson conceived of the study and were in charge of its design, coordination and writing of the manuscript. I am the second author and worked on the final preparation of the manuscript.

Unpublished Data. The functional consequences of ALKBH3 depletion. The majority of the work in this thesis. Experimental design, coordination, and data interpretation was performed by myself, Thorkell Gudjonsson and Stefan Sigurdsson. I was responsible for cell culture, western blots, micro-irradiation experiments, RNA-IP design and execution, decoder tool design, FLAG and siRNA resistant ALKBH3 plasmids, RNA Scope, confocal microscopy and RNA-fractionation. Thorkell Gudjonsson was responsible for cell culture, western blots, single-strand annealing assay and confocal microscopy. Kritika Kirty was responsible for performing and analyzing qPCR, RNA-IP design and execution. Drifa Hrund Gudmundsdottier was responsible for the FACS of the DR-GFP for GC.

1 Introduction

1.1 Cancer

Cancer is a major human disease condition and is the second leading cause of death worldwide, accounting for 8.8 million deaths in 2015, or 1 in 6 of all global deaths (World Health Organization, https://www.who.int/). With aging populations the prevalence of cancer is expected to increase and therefore the need for effective and better treatment will also increase. A complication with treating cancer is that it is not simply one disease, but many with an enormous variety of factors responsible for the development of the disease. Due to this variety in origin and disease development, treatments need to become more specific to the individual patients.

1.1.1 Breast Cancer

Breast cancer is the most common form of cancer among women. In 2015, 570,000 women died globally from breast cancer, accounting for approximately 15% of all cancer deaths among women. In Iceland alone, more than 200 women are diagnosed with breast cancer each year. This accounts for ~30% of all cancer diagnoses in Icelandic women (Krabbameinskra, http://www.krabbameinsskra.is/). In a subset of cases, the development of breast cancer is linked with defective DNA damage repair (Lord et al., 2016; Saal et al, 2008). Within the Icelandic population there is a founder mutation in a key DNA damage repair protein BRCA2. This mutation is found in 7-8% of all breast cancer patients and 40% of males with breast cancer in Iceland (Thorlacius et al., 1997; Tryggvadottir et al., 2006). Individuals harboring this mutation have increased risk of developing the disease as well as worse prognosis for survival (Tulinius 2002; Jonasson et al., 2016). Germline mutations in other DNA damage response genes such as ATM, BRCA1 and XRCC1 have been implicated in increased breast cancer risk (Choi et al., 2016; Couch et al., 2014; Patrono et al., 2014; Lord et al., 2012). Some chemotherapeutic drugs, such as PARP inhibitors or platinum agents, aim to exploit weaknesses in the DNA damage response in order to treat breast cancer (Bryant et al., 2005; Farmer et al., 2005; Ashworth et al., 2008; Tutt et al., 2018). Research from many laboratories has tremendously advanced our current understanding how breast cancer and DNA repair are linked and how

this knowledge can be therapeutically exploited. A wider understanding of the molecular mechanisms of DNA repair and their defects in cancer patients hold great promises for better tools in future cancer treatment.

1.2 DNA Damage

DNA is under constant assault from various exogenous and endogenous forces. Exogenous source include exposure to sunlight or cigarette smoke, while endogenously caused DNA lesions can be attributed to stochastic errors of the cell´s own machinery. In proliferating cells, DNA replication has been identified as the major cause of stochastic DNA damage (Tomasetti & Vogelstein, 2015). But also errors in mitosis can cause DNA damage (Pedersen et al., 2016) and have furthermore been strongly associated with cancer development through whole-genome doubling events (Zack et al., 2013). A common theme among all these sources is that they may result in DNA lesions. If these lesions persist within the DNA, they may have mutagenic effects that lead to even more DNA damage. Such genomic instability, which if located in certain parts of the genome may aid in the formation of a disease such as cancer (Vijg et al., 2013; Janssens et al., 2006). The damage to the DNA can exist in various forms such as single-strand breaks (SSBs), double-strand breaks (DSBs), interstrand-crosslinks (ICLs), insertions, deletions, and mismatched base pairs. Each of these types of damage have specific proteins and pathways through which the cell can repair the damage and avoid the adverse effects that result from the damage. There are however many instances in which these pathways do not function correctly and the damage is able to persist and potentially disrupt the proper function of a cell. This thesis will focus primarily on DNA alkylation damage and DNA double strand breaks.

1.2.1 Alkylation Damage

Alkylation damage is when a alkyl group is placed on a nucleotide and the location of the alkylation is considered a lesion. As with many DNA damage causing agents, alkylating agents are present in the environment. Exposure to alkylating agents can come from cigarette smoke, chemotherapeutic treatments, or exposure to industrial chemicals. Examples of such agents are methyl halides, such as methyl chloride and methyl bromide, or drugs like Methyl-methanesulfonate (MMS) (Sedgwick et al., 2004). Alkylation damage can also result from endogenous sources such as S-Adenosyl methionine

(SAM). A major form of this damage is the addition of a methyl group to the DNA or RNA. There are however many instances of methylation on DNA and RNA that are functional marks rather than damaging, such as 5-methylcytosine, a key regulatory modification in epigenetics.

Methylation inducing agents can react with DNA and RNA at 12 different sites, including all the exocyclic oxygens and most ring nitrogens. They can also methylate oxygen atoms in phosphates of the sugar–phosphate backbone, thereby generating methylphosphotriesters (Sedgwick et al., 2004). The type of methylation that occurs is more specifically dependent on the type of reaction in which the methylation is added. There are two mechanisms, SN1 and SN2, through which this can happen. Methylation from SN1 agents have preference to add 7-methylguanine and 3-methyladenine (Bodell et al., 1979). SN2 agents, such as the previously mentioned methyl halides, SAM and MMS, are more likely to target the 1' nitrogen in adenine and 3' nitrogen in cytosine, producing N1-methyladenine (m1A) and 3-methylcytosine (m3C) (Bodell et al., 1979). m1A and m3C tend to occur only in ssDNA and RNA due to the protective effect of a paired base when in dsDNA.

The persistence of m3C and m1A within DNA can prove to be problematic. If left unrepaired, m3C can lead to inflammation, nucleotide misincorporation and most notably replicative stress (Dango et al., 2011; Soll et al., 2012). Less accurate polymerases, such as translesion polymerase POLQ, may read through the alkylation damage and allow for continued growth despite the presence of alkylation damage (Ukai et al., 2006). Additionally the polymerases used can be at times error prone and allow for misincorporporations. For example, Pol β is a crucial polymerase for translesion synthesis is found to be commonly mutated in cancers and allows for misincorporations (Starcevic et al., 2004; Krokan et al., 2013). Proficiency for alkylation repair is therefore critical to protect the cell against accumulation of genetic mutations. Recent cancer genome sequencing studies have revealed a profound impact from treatment with alkylating agent temozolomide leading to a specific mutational pattern (Alexandrov et al., 2013). This observation provides a solid link between alkylation damage and the formation of mutations and potentially an eventual cancer.

In recent years, several researchers have reported loss of MGMT in breast cancer thereby implicating defective alkylation repair in breast cancer development (Spitzwieser et al., 2015; Drablos et al., 2004; Fumagalli et al., 2012). The MGMT gene has an important role in removing cytotoxic adducts from O(6)-guanine in DNA (Esteller et al., 2000).

1.2.2 Double Strand Breaks

Double strand breaks are one of the most severe forms of DNA damage. A double strand break is defined, as implied by the name, by the breakage of both strands of the DNA. The persistence of 1-2 of these breaks in the genome can cause genomic translocations, instability, mutations and eventually lead to cellular senescence or death (d'Adda di Fagagna et al., 2008). It is therefore of paramount importance that these breaks be recognized and repaired quickly in order to prevent any of these forms of genomic stress from occurring. Mutations within genes involved in DNA DSB repair, such as BRCA1, BRCA2, PALB2 or ATM have been implicated in several types of cancer (Prakash et al., 2015). However prior to any repair, the break must first be recognized.

1.3 DNA Damage Recognition and DNA Repair

There are multiple pathways used for the detection and the transduction of the signal to activate DNA repair proteins in response to DNA damage. The detection, signal transduction and repair of DNA damage is collectively known as DNA damage response (DDR) (Harper & Elledge, 2007). This thesis will focus primarily on the DDR to DSBs. The primary method of detecting DSBs is through the MRN complex, which is composed of MRE11, RAD50, and, NBS1 (MRN). The MRN complex is responsible for initially recognizing the presence of a DSB within the DNA and then recruiting the Ataxia Telangiectasia Mutated (ATM) kinase that phosphorylates several key proteins triggering DNA damage checkpoints in the cell cycle (Paull & Lee, 2005). A key target of the ATM kinase is histone H2AX, that is phosphorylated on the serine 139 residue, changing it into the γ-H2AX. This phosphorylation can spread along large distances on the chromatin that surrounds the DSB (Rogakou et al., 1999). This amplifies the response to the DSB. Functionally H2AX has been shown to orchestrate the recruitment of various repair factors to the sites of DNA damage (Celeste et al., 2002, Polo & Jackson, 2011). The cell will then choose how it will repair the DSB, of which there are several options.

1.3.1 DNA DSB Repair Pathway Choice

Following the previously described pathway of initial recognition, one of the proteins recruited by γ-H2AX is MDC1, an additional target of ATM (Stucki et al., 2005). MDC1 will recruit RNF8, an E3 ubiquitin ligase (Mailand et al., 2007). E3 ubiquitin ligases will recruit E2 ubiquitin-conjugating enzymes, which generally fall under the UBE2 naming convention, for example UBE2A, UBE2H and UBE2V1. The E2 ubiquitin-conjugating enzymes will interact with E1 ubiquitin activating enzymes which initially recognize the target and through the chain of E1, E2 and E3 will deposit a ubiquitin on the target protein.

At this point there is a choice of which repair pathway will be used in order to repair the DSB (Chapman, Taylor, & Boulton, 2012). There are two major repair pathways available to be used for repair. One option is Homologous Recombination (HR), a pathway that utilizes available sister chromatids as a template for repair, therefore limiting it to S and G2 phase of the cell cycle. The other option is Non-Homologous End Joining (NHEJ), which involves the direct ligation of the broken DNA ends and is available in all phases of the cell cycle. The decision of which pathway to choose is clearly dependent on the cell cycle, but is largely decided by DNA end resection. Two proteins, BRCA1 and 53BP1, are key effectors of the DDR and determine the levels of resection that occurs which in turn determines the choice of repair pathway (Panier & Boulton, 2013; Lukas et al., 2011). It has been suggested there is constant competition between 53BP1 and BRCA1 at the site of DSBs as both proteins co-exist in nuclear foci at sites of DNA DSB (Bunting et al., 2010).

In S and G2 phases, the HR pathway is preferentially deployed for accurate repair (Kass et al., 2010). This depends on BRCA1 binding the break, thereby displacing and preventing binding by 53BP1. 53BP1 however has since been shown to still be essential by promoting the fidelity of HR (Ochs et al., 2016). BRCA1 binding allows for the recruitment of the MRN/CTIP complex which promotes initial end-processing. This initial end-processing results in the formation of partially resected intermediates, which is followed by a more extensive processing step by the nucleases DNA replication helicase/nuclease 2 (DNA2), Exonuclease 1 (EXO1) and the Bloom syndrome helicase (BLM) (Gravel et al., 2008; Zhu et al., 2008). The fine-tuning of end resection is necessary for the next proteins in the HR pathway to function.

For NHEJ, 53BP1 is the preferred protein to bind the DNA ends over BRCA1. Following the recruitment of RNF8 and subsequent deposition of a ubiquitin on the γ-H2AX, RNF168 is subsequently recruited to add further ubiquitin allowing for the recruitment of 53BP1 (Mailand et al., 2007; Doil et al.,

2009). 53BP1 itself is an ATM target, and when phosphorylated, 53BP1 will recognize the ubiquitin signal and together with RIF1 prevent extensive end-resection at the site of the DSB (Feng et al., 2013). The regulation of end resection is crucial for protecting the DNA ends to allow for eventual ligation. This process channels the cell towards NHEJ (Hustedt et al., 2017). RIF1 has furthermore been shown to be a key effector protein of 53BP1, as without RIF1, 53BP1 can not suppress extensive DNA resection (Chapman et al., 2011). RIF1 accumulation at DSBs in S-phase is antagonized by BRCA1, suggesting that BRCA1 inhibits the pro-NHEJ activity of 53BP1-RIF1 in a cell cycle specific fashion (Escribano-Diaz et al., 2013). Another reported effector of 53BP1 is PTIP. PTIP is recruited to phosphorylated 53BP1 where it counteracts resection, thereby inhibiting HR. However, PTIP does not seem to promote the use of NHEJ (Callen et al., 2013).

Further proteins are dependent on the RNF8-RNF168-53BP1 pathway in order to carry out successful NHEJ. REV7 (also known as MAD2L2) is a protein that prevents resection and promotes the use of NHEJ (Xu et al., 2015; Boersma et al., 2015). The shieldin complex, comprised of SHLD1, SHLD2, SHLD3, and REV7, is recruited to DSBs in a 53BP1 and RIF1 dependent manner. SHLD2 binds the ssDNA and allows the complex to prevent end-resection from occurring. The loss of shieldin complex members or inefficient recruitment will impair NHEJ and cause hyper-resection to occur (Greenberg, 2018).

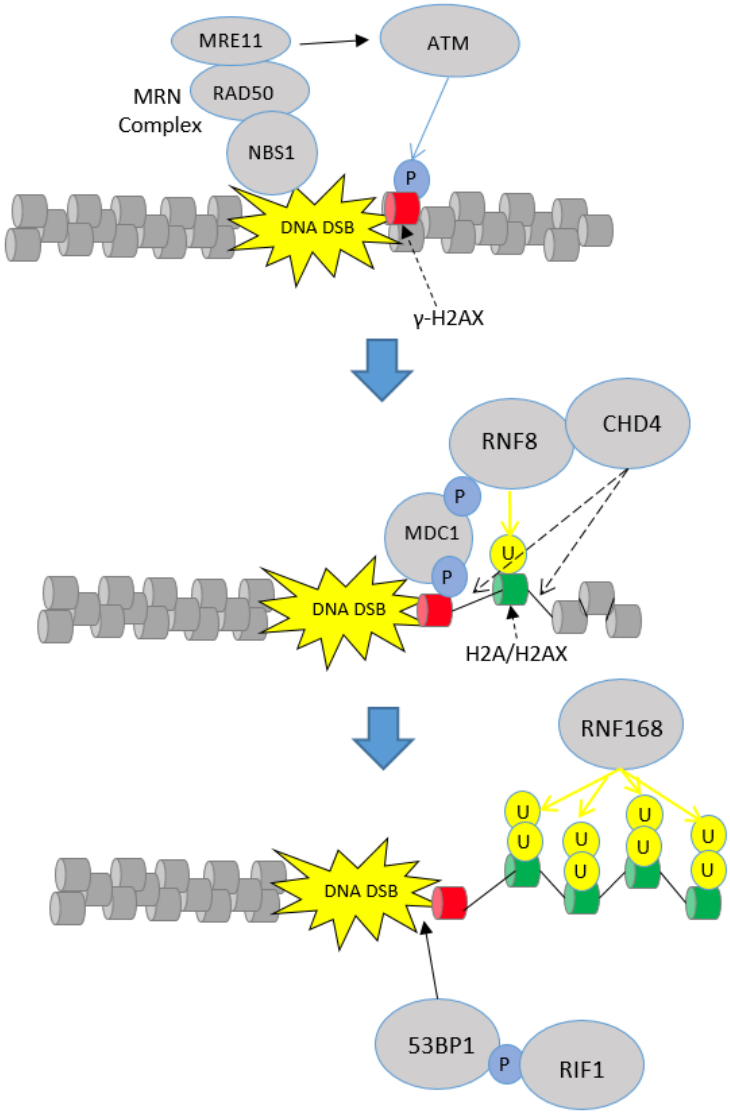


Figure 1. A snapshot of the ubiquitin dependent DSB signalling pathway following a DSB. A DNA DSB occurs and is recognized by MRN Complex, which activates ATM, phosphorylating H2AX and making γ-H2AX. MDC1 then recognizes phosphorylated γ-H2AX, bringing RNF8 and additional NBS1. RNF8 ubiquitylates H2A/H2AX. RNF168 recognizes the ubiquitylation put in place by RNF8, meanwhile CHD4 opens up chromatin to allow for further ubiquitylation chain formation by RNF168. RNF168 adds additional ubiquitylation, further opening the DNA to access by repair machinery. 53BP1 and RIF1 will bind at the site of the damage, preventing extensive resection and channeling the cell towards NHEJ.

1.3.2 Homologous Recombination

Homologous Recombination (HR) is dependent on the availability of sister chromatids to be used as a template for repair and is therefore restricted to S and G2 phases of the cell cycle. Following a DSB, the single-stranded 3' overhangs generated by resection are coated by RPA, forming a nucleoprotein filament. RPA is responsible for preventing the DNA both from forming potentially harmful secondary structures and from degradation by single stranded nucleases (Chen et al., 2013). BRCA2, together with PALB2, recognizes RPA and binds to the end of the single strand via protein protein interactions bringing along the recombinase RAD51, which is then loaded onto the DNA, displacing RPA (Wilson & Elledge 2002; Buisson et al, 2010). The RAD51 nucleoprotein filament, along with other proteins, then facilitates the pairing of the broken chromosomal end with undamaged sister chromatid and stimulates the invasion into the sister chromatid followed by DNA repair synthesis, using the intact sister chromatid as a template (Wilson & Elledge, 2002). BRCA2 additionally promotes the recruitment of ATR Interaction Protein (ATRIP) that recruits Ataxia Telangiectasia and Rad3 related protein (ATR). Activated ATR phosphorylates Chk1 that causes a signal cascade leading to arrest of the cell cycle (Zou & Elledge, 2003). It is of extreme importance to stop the cell cycle when DNA damage is present. This allows the cell to repair the DNA damage and if the damage is extensive undergo programmed cell death or apoptosis. When DDR or cell cycle checkpoints malfunction the damage is allowed to persist within the DNA, which could be detrimental to the cell and the host (Polo & Jackson, 2011).

1.3.3 Non-Homologous End Joining

The predominantly used form of repair in human cells is Non-Homologous End Joining (NHEJ) (Beucher et al., 2009). Unlike homologous recombination it does not require sister chromatids as a template, meaning NHEJ can be active in all phases of the cell cycle. There is classical NHEJ as well as several alternative forms of NHEJ. Classical NHEJ is initiated by the binding of the heterodimer of Ku70/80 to the ends of the DNA break. Ku70/80 binding recruits DNA-PK to the site of the DNA DSB. Ku70/80 interacts with the catalytic domain of DNA-PK, thereby activating its catalytic activity. This is followed by the synapsis of the two ends of the broken DNA by DNA ligase IV and its cofactor XRRC4 (Smogorzewska et al., 2004). XLF (XRCC4-like factor), also

known as Cernunnos or NHEJ1, has also been shown be part of the effective assembly of the NHEJ ligase complex at DNA ends (Chasseval et al., 2006). The absence of XLF leads to decreased end-joining efficiency (Tsai and Chu, 2013).

As mentioned previously, 53BP1 is a key regulator in this pathway. While not playing an enzymatic role in the repair process, it acts to limit DNA end-resection, thereby fostering accurate NHEJ to occur. In the absence of 53BP1, the DNA undergoes increased resection and therefore disrupts the classical NHEJ repair mechanism (Bothmer et al., 2010). Loss of 53BP1 will also cause the cell to turn to alternative forms of repair that are less accurate and often deleterious (Ochs et al., 2016).

1.3.4 Alternative forms of DNA DSB repair

One alternative form of DSB repair is single-strand annealing (SSA). Following a DSB, the 5' ends of the strands strands on each side will undergo end resection by EXO1. The process follows the initial steps of HR by having ssDNA overhangs being bound by RPA, but instead of RAD51 recognition, the RPA bound strands are recognized by RAD52. This process continues until matching repeated sequences are available on both sides of the DNA. These sequences are then annealed together by RAD52 resulting in a deletion of the DNA located between the matching sequences. This is clearly a very mutagenic process due to deletions always resulting from the mechanism and is therefore not a preferable form of repair. (Chang et al., 2017; Lee et al. 2014). Interestingly, cells using this type of repair can increase in occurrence following the depletion of 53BP1. In S and G2 phases, the absence of 53BP1 will allow the broken DNA ends to undergo excessive end-resection. This thereby prevents the cell from being able to effectively load RAD51 and undergo the error-free form of HDR. The cell will instead need to rely on RAD52, allowing the use of SSA to repair the break and causing increased likelihood of mutations due to the use of this error-prone form of repair (Ochs 2016).

Another form of DSB repair, called alternative NHEJ (alt-NHEJ), is available should canonical NHEJ be dysfunctional. This pathway is also referred to as Microhomology-Mediated End Joining (MMEJ). MMEJ occurs during S-phase and utilizes short homologous regions as templates. In MMEJ, the ssDNA overhangs resulting from a DSB are utilized and undergo limited processing to produce 3´ overhangs. The overhangs are annealed at very short 5-25 nucleotide homologous sequences. This method of repair is

considered to be very error prone due to the amount of annealing, resecting and insertion of bases during the process (Mcvey & Lee, 2008; Bennardo et al., 2008).

In the following sections, the proteins which are most relevant for the work in this thesis, will be described in detail.

1.4 RNF168

RNF168, as mentioned previously, is an E3 ubiquitin ligase and a key protein in the DNA DSB repair pathway. Knockout mice for RNF168 show defective recruitment of key DNA DSB repair proteins such as 53BP1 to sites of DNA damage. These RNF168 deficient mice display increased radiosensitivity, immuno-defects and decreased spermatogenesis. Additionally, the dual inactivation of RNF168 and the tumor suppressor p53, promotes survival of such cells and propagation of chromosomal translocations. In combination with the other symptoms listed above, this can explain the increased cancer risk within these RNF168 deficient mice (Bohgaki et al., 2011). Within humans, mutations in RNF168 underlie a genetic disorder known as RIDDLE (radiosensitivity, immunodeficiency, dysmorphic features, and learning difficulties) syndrome (Stewart et al., 2009; Devgan et al., 2011). RIDDLE syndrome is a very rare inherited autosomal recessive disease with only 4 cases identified and detailed (Stewart et al., 2009; Devgan et al., 2011; Pietrucha et al., 2017). Patients suffering from the disease have varied clinical presentations of the disease such as gait ataxia, microcephaly, learning defects and ataxia telangiectasia. While there was some variance in the symptoms of the disease, a commonality among all the patients was the inability to recruit 53BP1 to the sites of DSBs.

RNF8 is the primary E3 ubiquitin ligase in the ubiquitin-dependent DNA repair pathway (Mailand et al., 2007). However, RNF8 alone is not sufficient for K63 ubiquitylation of H2A and H2AX. RNF168 is required to amplify the signal by depositing additional ubiquitin onto the initial ubiquitin chains from RNF8 (Doil et al., 2009). RNF168 is however completely dependent on the initial lysine 63 ubiquitylation by RNF8 for proper RNF168 recruitment. In the absence of RNF168, ubiquitin chains can not be formed and therefore 53BP1 will not be recruited (Doil et al., 2009). This would disrupt the recruitment of many of the previously mentioned proteins that depend on 53BP1. Ubiquitin chains have also been suggested to be required for the recruitment of BRCA1, a protein often found mutated and dysfunctional in

cancers (Doil et al., 2009). The inactavation of the RNF168-53BP1 disrupts accurate repair and promotes the use of SSA (Ochs et al., 2016). Furthermore, the combined absence of BRCA1 and RNF168 may force the cell to use SSA form of repair for survival (Ochs et al., 2016). In BRCA1 heterozygous cells, the absence of RNF168 has been shown to disrupt the ability of the cell to form RAD51 foci, indicated RNF168 is necessary for HR (Zong et al, 2019).

There is a multitude of ways in which RNF168 is regulated. There is the aforementioned dependence of RNF8 for proper function. Another example shows USP7 (Ubiquitin Specific Peptidase 7) binding RNF168 and regulating the stability of the protein (Zhu et al., 2015). RNF168 has also been shown to be regulated by TRIP12 and UBR5, which target RNF168 for degradation. (Gudjonsson et al., 2012). Four other deubiquitinases; USP3, USP16, BRCC36 and OTUB1 have been identified as counteractors to the function of RNF8 and RNF168 (Bartocci et al., 2013). Many other examples of regulation have been identified but the key takeaway is that RNF168 is a highly regulated protein.

The deposition of ubiquitin by RNF168 clearly plays a key role in DNA repair as it is necessary for the recruitment of proteins that help determine which repair pathway the cell will take. There are however further process RNF168 has been implicated in. RNF168-mediated recruitment of 53BP1 has been shown to be crucial in the formation of 53BP1 nuclear bodies in G1 that help protect the DNA during times of replicative stress (Lukas et al., 2011; Harrigan et al., 2011). RNF168 has been shown to be important in inhibiting transcription around DSBs (Shanbhag et al., 2010). In the absence of the telomere regulating protein TRF2, RNF168 has been shown to promote the fusion of dysfunctional telomeres which may have deleterious consequences (Arnoult & Karlseder, 2015; Okamoto et al., 2013; Peuscher & Jacobs, 2011).

1.5 ALKBH3

ALKBH3 (alkB homolog 3, alpha-ketoglutarate dependent dioxygenase) is one of the 9 mammalian homologs of the bacterial AlkB gene. This includes ALKBH1-8 and FTO (Fat mass and obesity associated protein), which is informally known as ALKBH9. In *E. Coli*, AlkB is responsible for the direct reversal of alkylation damage on ssDNA by SN2 methylating agents such as methyl methanesulfonate (MMS) (Dinglay et al., 2000; Wyatt et al., 2008). This is done by an iron and 2-oxoglutarate dependent oxidative demethylation reaction that reverts the base to its unmodified state (Duncan et al., 2002). This

reaction produces succinate and carbon dioxide, with the methyl group being released as formaldehyde (Trewick et al., 2002). While there are 9 homologs, only 2 of them, ALKBH2 and ALKBH3, are considered to be true functional homologs of AlkB. This is due to the type of methylation mark targeted for repair by the proteins. AlkB most efficiently removes methyl groups from N1-methyladenosine (m1A) and 3-methylcytosine (m3C) from ssDNA (Trewick et al., 2002). When these methylation marks exist on DNA they are considered to be improperly located and therefore classified as a form of DNA damage, specifically alkylation damage. ALKBH2 and ALKBH3 were shown to have the same methylation targets as AlkB, however ALKBH2 had a preference for removing the substrates from dsDNA, where as ALKBH3, along with the helicase ASCC3, were responsible for removing the methylation from ssDNA (Dango et al., 2011). For ALKBH3, the removal of these methylation marks from ssDNA is dependent on the presence of ASCC3, and in its absence cells suffer from increased levels of m3C damage, which is thought to reduce cell proliferation and lead to spontaneous DNA DSB damage in a cell-type specific manner (Dango et al., 2011).

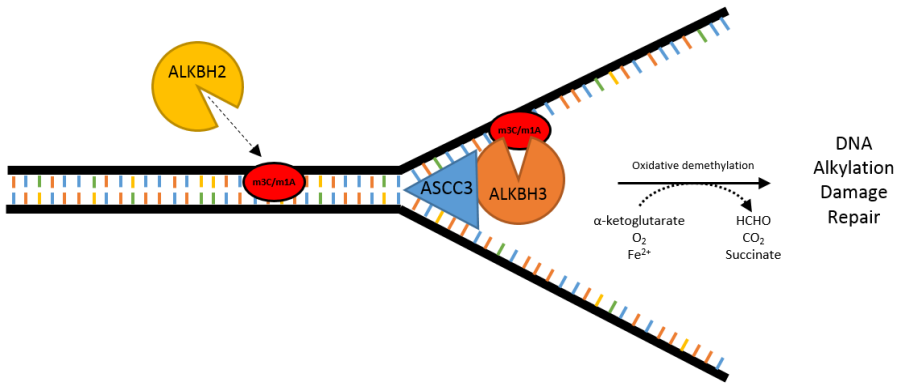


Figure 2. DNA alkylation repair through oxidative demethylation by ALKBH2 and ALKBH3. ALKBH2 shows preference for m1A/ m3C in dsDNA where as ALKBH3 shows preference for m3C/m1A in ssDNA. ALKBH3 is dependent on helicase ASCC3 for unwinding DNA and allowing access to alkylation damage (Sedgwick et al., 2004; Dango et al., 2011).

ALKBH3 has been shown to be overexpressed in various cancers, including lung, pancreatic and prostate cancer (Tasaki et al., 2011; Yamato et al., 2012). The Yamato (2012) study looked into overexpression of ALKBH3 in

pancreatic cancer. They found that the knockdown of ALKBH3 caused increased apoptosis and a 50-70% reduction in tumor growth *in vivo*. Additionally the paper showed the knockdown of ALKBH3 decreasing the expression of VEGF (Vascular Endothelial Growth Factor), a protein involved in angiogenesis, therefore implicating ALKBH3 has some role in the regulation of angiogenesis. Angiogenesis has long been linked to cancer formation and growth due to the need for cancer cells to have increased blood supply and anti-angiogenic drugs, such as bevacizumab are commonly used in anti-cancer chemotherapies (Willett et al., 2004).

ALKBH3 was first shown to act as a RNA demethylase in 2003 (Aas et al., 2003). Aas et al. described ALKBH3 acting in its known repair capacity for alkylation damage, but now on RNA rather than it's known function in DNA repair. Importantly Aas distinguishes a difference between ALKBH3 and ALKBH2, showing that only ALKBH3 has a capability of targeting the single stranded RNA. This is in line with the finding that while ALKBH2 and ALKBH3 showed the same repair targets, ALKBH2 acts on dsDNA where as ALKBH3 acts on ssDNA (Dango et al., 2011). *AlkBh2* knockout mice were shown to have increased levels of m1A in their genomic DNA as well as showing increased sensitivity to MMS, where as these phenotypes were not observed in *AlkBh3* knockout mice (Ringvoll et al., 2006).

In recent years, ALKBH3 has become a gene of interest due to the increasing research into epitranscriptomics. ALKBH3 has had its role in mRNA demethylation expanded upon in several papers discussed later in the thesis.

1.5.1 ALKBH3 in DSB repair

The previously referenced 2011 paper from Dango contained a LC-MS/MS for a FLAG-tagged constuct of ALKBH3 in order to identify interactors. Among the list of top 30 protein interactors were the proteins RIF1 and CHD4 (Dango et al., 2011). CHD4 is involved in the remodeling of chromatin in order to allow for increased access to the DNA by RNF168, which further ubiquitylates histone H2A/H2AX (Larsen et al., 2010; Luijsterburg et al., 2012). As previously described, RIF1 is a binding partner of 53BP1 and PTIP, proteins involved in DNA repair pathway choice.

There is another indication ALKBH3 may play some role in DSB repair. In *E.Coli*, AlkB was shown to interact with RecA to stimulate alkylation repair (Shivange et al., 2016). RecA is the *E. coli* homolog of human RAD51, a protein that is crucial for effective DNA DSB repair through the homologous

recombination pathway (Wilson & Elledge, 2002).

1.6 FTO

FTO is a Alpha-Ketoglutarate Dependent Dioxygenase, commonly known as Fat Mass and Obesity-Associated Protein. It belongs to the AlkB family proteins and is also known as ALKBH9. Mutations in the FTO have been associated with increased obesity within humans (Chu et al., 2008, Thorleifsson et al., 2008). Following these studies, FTO became a notable target of research due to scientific interest in causes for obesity and the link between obesity and cancer. FTO had been shown to oxidatively demethylate 3-methyluridine and 3-methylthymine in ssRNA and ssDNA in vitro, but at a considerably lower level than that of the other AlkB-family proteins (Lee et al., 2005).

FTO has an established, but complex role as a RNA demethylase. A 2011 study from Jia et al. identified N6-methyladenosine in nuclear RNA as a major target of FTO (Jia et al., 2011). For years it was assumed that both FTO and ALKBH5 demethylated N6-methyladenosine exclusively. It was however recently shown that FTO actually has a preference for N6,2′-O-dimethyladenosine, henceforth referred to a m6Am (Mauer et al., 2016). The Mauer paper showed using Thin-Layer Chromatography that FTO has an approximately 100-fold higher catalytic efficiency towards m6Am than m6A. The significance of this is separating the targets of FTO from the targets of ALKBH5, and therefore possibly separating FTO from the implications of m6A in differentiation and splicing.

The RNA demethylation targets were expanded on in 2018 by He and colleagues. He showed that knockdown of FTO increased global m6A levels as well as m6Am (Wei et al., 2018). This may overcome the preference of m6A for m6Am due to the sheer amount of m6A present in mRNA, as m6A has been detailed as the most abundant RNA modification (Meyer et al., 2012; Dominissini et al., 2012). Additionally the He group identified N1-methyladenosine as a target of FTO, but only within tRNAs. The group suggested that FTO was acting both in the nucleus and cytoplasm, but the m6A and m1A demethylation activity was largely nuclear and needed to be further investigated (Wei et al., 2018).

1.7 Epigenetics

Epigenetics is the study on mechanisms through which changes in phenotype

are established and stably maintained following cellular divisions without involving any changes in genotype. The fundamental unit of chromatin is the nucleosome representing a short stretch of DNA wrapped around a protein complex consisting of histone proteins arranged as octamers (Cutter et al., 2015). Nucleosome occupancy over regulatory regions in DNA associates with transcriptional activity as densely occupied regions are poorly accessible to transcription factors (Hesson et al., 2014). The degree of nucleosome occupancy, or chromatin packaging, is regulated by the use of so-called epigenetic marks of which the best studied is undoubtedly DNA methylation involving the addition of methyl groups to the 5-position on cytosine bases forming 5-methylcytosine (m5C). Where m5C is followed by guanine is referred to as a CpG site (Jones et al., 2012). Methyl-Binding Domain-containing proteins recognize and bind to methylated CpGs and recruit histone modifiers to mediate or maintain repressed chromatin structure (Jones et al., 2012). The collective term for all these regulatory modifications on the DNA is called the epigenome.

1.7.1 Epigenetics in Cancer and DNA Repair

In cancer cells, the epigenome is frequently disrupted. This is characterized by a global loss in repressive marks and localized modifications over regulatory elements (Baylin et al., 2014). It is now known that genes functionally involved in shaping the epigenome of human cells are often mutated in breast cancer and various other cancers, for example MLL3, MLL2, ARID1A and SETD2 (Simó-Riudalbas et al., 2014). The discoveries of recurrent mutations in epigenetic genes provided a important link between disruptions in the epigenome and the development of cancer. In addition to this, earlier observations had already established repressive epigenetic marks over regulatory regions of known tumor suppressor genes in cancer cells. In breast cancer, this catalogue includes CpG promoter methylation of BRCA1, RAD51C, FOXC1, RUNX3 and L3MBTL4 (Stefansson et al., 2011; Cunningham et al., 2014; Muggerud et al., 2010). Of these, the BRCA1 gene is well established as a cancer predisposition gene where germline mutations are found in association with greatly increased risk for breast and ovarian cancer (Lalloo et al., 2012). Other high-risk breast cancer susceptibility genes such as BRCA2, PALB2, BARD1, FANCM, ATM, CHEK2 and TP53, however, are not found epigenetically silenced (Collins et al., 1997, Petrovics et al., 2012).

1.8 Epitranscriptomics

Epitranscriptomics is the study of biochemical marks on RNA (Zhao, Roudtree and He, 2017). Similar to epigenetics, epitranscriptomics involves modifications to nucleotides that may alter their function, but has no impact on the sequence of nucleotides. While epitranscriptomics covers modifications on all forms of RNA, this thesis will focus specifically on modifications found on mRNA. According to an online database of modifications (http://modomics.genesilico.pl/, Boccaletto et al., 2018), there are currently 170 known modifications on RNA, and 168 methyltransferases responsible for depositing methylation onto RNA.

1,2′-O-dimethyladenosine	m1Am
1,2′-O-dimethylguanosine	m1Gm
1,2′-O-dimethylinosine	m1Im
1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine	m1acp3Y
1-methyladenosine	m1A
1-methylguanosine	m1G
1-methylinosine	m1I
1-methylpseudouridine	m1Y
2,8-dimethyladenosine	m2,8A
2-geranylthiouridine	ges2U
2-lysidine	k2C
2-methyladenosine	m2A
2-methylthio cyclic N6-threonylcarbamoyladenosine	ms2ct6A
2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine	ms2io6A
2-methylthio-N6-hydroxynorvalylcarbamoyladenosine	ms2hn6A
2-methylthio-N6-isopentenyladenosine	ms2i6A
2-methylthio-N6-methyladenosine	ms2m6A
2-methylthio-N6-threonylcarbamoyladenosine	ms2t6A
2-selenouridine	se2U
2-thio-2′-O-methyluridine	s2Um
2-thiocytidine	s2C
2-thiouridine	s2U
2′-O-methyladenosine	Am
2′-O-methylcytidine	Cm
2′-O-methylguanosine	Gm
2′-O-methylinosine	Im
2′-O-methylpseudouridine	Ym
2′-O-methyluridine	Um
2′-O-methyluridine 5-oxyacetic acid methyl ester	mcmo5Um
2′-O-ribosyladenosine (phosphate)	Ar(p)
2′-O-ribosylguanosine (phosphate)	Gr(p)
2′3′-cyclic phosphate end	(pN)2′3′>p
3,2′-O-dimethyluridine	m3Um
3-(3-amino-3-carboxypropyl)-5,6-dihydrouridine	acp3D
3-(3-amino-3-carboxypropyl)pseudouridine	acp3Y
3-(3-amino-3-carboxypropyl)uridine	acp3U
3-methylcytidine	m3C
3-methylpseudouridine	m3Y
3-methyluridine	m3U
4-demethylwyosine	imG-14
4-thiouridine	s4U
5,2′-O-dimethylcytidine	m5Cm
5,2′-O-dimethyluridine	m5Um
5-(carboxyhydroxymethyl)-2′-O-methyluridine methyl ester	mchm5Um
5-(carboxyhydroxymethyl)uridine methyl ester	mchm5U
5-(isopentenylaminomethyl)-2-thiouridine	inm5s2U
5-(isopentenylaminomethyl)-2′-O-methyluridine	inm5Um
5-(isopentenylaminomethyl)uridine	inm5U
5-aminomethyl-2-geranylthiouridine	nm5ges2U
5-aminomethyl-2-selenouridine	nm5se2U
5-aminomethyl-2-thiouridine	nm5s2U
5-aminomethyluridine	nm5U
5-carbamoylhydroxymethyluridine	nchm5U
5-carbamoylmethyl-2-thiouridine	ncm5s2U
5-carbamoylmethyl-2′-O-methyluridine	ncm5Um
5-carbamoylmethyluridine	ncm5U
5-carboxyhydroxymethyluridine	chm5U
5-carboxymethyl-2-thiouridine	cm5s2U
5-carboxymethylaminomethyl-2-geranylthiouridine	cmnm5ges2U
5-carboxymethylaminomethyl-2-selenouridine	cmnm5se2U
5-carboxymethylaminomethyl-2-thiouridine	cmnm5s2U
5-carboxymethylaminomethyl-2′-O-methyluridine	cmnm5Um
5-carboxymethylaminomethyluridine	cmnm5U
5-carboxymethyluridine	cm5U
5-cyanomethyluridine	cnm5U
5-formyl-2′-O-methylcytidine	f5Cm
5-formylcytidine	f5C
5-hydroxycytidine	ho5C
5-hydroxymethylcytidine	hm5C
5-hydroxyuridine	ho5U
5-methoxycarbonylmethyl-2-thiouridine	mcm5s2U
5-methoxycarbonylmethyl-2′-O-methyluridine	mcm5Um
5-methoxycarbonylmethyluridine	mcm5U
5-methoxyuridine	mo5U
5-methyl-2-thiouridine	m5s2U
5-methylaminomethyl-2-geranylthiouridine	mnm5ges2U
5-methylaminomethyl-2-selenouridine	mnm5se2U
5-methylaminomethyl-2-thiouridine	mnm5s2U
5-methylaminomethyluridine	mnm5U
5-methylcytidine	m5C
5-methyldihydrouridine	m5D
5-methyluridine	m5U
5-taurinomethyl-2-thiouridine	tm5s2U
5-taurinomethyluridine	tm5U
5′ (3′ -dephospho-CoA)	CoA(pN)
5′ (3′ -dephosphoacetyl-CoA)	acCoA(pN)
5′ (3′ -dephosphomalonyl-CoA)	malonyl-CoA(pN)
5′ (3′ -dephosphosuccinyl-CoA)	succinyl-CoA(pN)

5′ diphosphate end	p(pN)
5′ hydroxyl end	5′-OH-N
5′ monophosphate end	(pN)
5′ nicotinamide adenine dinucleotide	NAD(pN)
5′ triphosphate end	pp(pN)
7-aminocarboxypropyl-demethylwyosine	yW-86
7-aminocarboxypropylwyosine	yW-72
7-aminocarboxypropylwyosine methyl ester	yW-58
7-aminomethyl-7-deazaguanosine	preQ1tRNA
7-cyano-7-deazaguanosine	preQ0tRNA
7-methylguanosine	m7G
7-methylguanosine cap (cap 0)	m7Gpp(pN)
8-methyladenosine	m8A
N2,2′-O-dimethylguanosine	m2Gm
N2,7,2′-O-trimethylguanosine	m2,7Gm
N2,7-dimethylguanosine	m2,7G
N2,7-dimethylguanosine cap (cap DMG)	m2,7Gpp(pN)
N2,N2,2′-O-trimethylguanosine	m2,2Gm
N2,N2,7-trimethylguanosine	m2,2,7G
N2,N2,7-trimethylguanosine cap (cap TMG)	m2,2,7Gpp(pN)
N2,N2-dimethylguanosine	m2,2G
N2-methylguanosine	m2G
N4,2′-O-dimethylcytidine	m4Cm
N4,N4,2′-O-trimethylcytidine	m4,4Cm
N4,N4-dimethylcytidine	m4,4C
N4-acetyl-2′-O-methylcytidine	ac4Cm
N4-acetylcytidine	ac4C
N4-methylcytidine	m4C
N6,2′-O-dimethyladenosine	m6Am
N6,N6,2′-O-trimethyladenosine	m6,6Am
N6,N6-dimethyladenosine	m6,6A

N6-(cis-hydroxyisopentenyl)adenosine	io6A
N6-acetyladenosine	ac6A
N6-formyladenosine	f6A
N6-glycinylcarbamoyladenosine	g6A
N6-hydroxymethyladenosine	hm6A
N6-hydroxynorvalylcarbamoyladenosine	hn6A
N6-isopentenyladenosine	i6A
N6-methyl-N6-threonylcarbamoyladenosine	m6t6A
N6-methyladenosine	m6A
N6-threonylcarbamoyladenosine	t6A
Qbase	Qbase
adenosine	A
agmatidine	C+
alpha-dimethylmonophosphate cap	mm(pN)
alpha-methylmonophosphate cap	m(pN)
archaeosine	G+
cyclic N6-threonylcarbamoyladenosine	ct6A
cytidine	C
dihydrouridine	D
epoxyqueuosine	oQtRNA
galactosyl-queuosine	galQtRNA
gamma-methyltriphosphate cap	mpp(pN)
glutamyl-queuosine	gluQtRNA
guanosine	G
guanosine added to any nucleotide	pG(pN)
guanylylated 5′ end (cap G)	Gpp(pN)
hydroxy-N6-threonylcarbamoyladenosine	ht6A
hydroxywybutosine	OHyW
inosine	I
isowyosine	imG2
mannosyl-queuosine	manQtRNA

Figure 3. List of currently known RNA modifications according to Modomics database from The Laboratory of Bioinformatics and Protein Engineering at the International Institute of Molecular and Cell Biology in Warsaw (Boccaletto et al., 2018)

The existence of these types of modifications have been known since the 1970s. N6-methyladenosine was shown to be present in mRNA and 2'-O-methylnucleosides were shown to be present in rRNA and tRNA (Desrosiers 1974). However for many years following their identification in RNAs, these modification remained largely unstudied due to lack of appropriate methods. In 2012, papers from the Rechavi group and Jaffrey group detailed a methylated RNA immunoprecipitation technique that allowed for mapping the modifications and subsequently revived interest in the field (Dominissini et al., 2012; Meyer et al., 2012).

The addition and removal of these methyl groups from nucleotides within mRNA has been implicated in having roles within alternative splicing (Zhao et al., 2014), mRNA export (Wickramasinghe et al., 2015), stability (Wang et al., 2014; Mauer et al., 2016) and translational efficiency (Lin et al., 2016; Slobodin et al., 2017), thus causing differing expression of proteins and affecting cells accordingly. This thesis will primarily focus on a few types of RNA methylation and their role in epitranscriptomics. The focus will be on

known demethylation targets of ALKBH3 (Aas et al., 2003; Duncan et al., 2002), 3-methylcytosine (m3C) and N1-methyladenosine (m1A), as well the modifications N6-methyladenosine (m6A) and N6,2'-O-dimethyladenosine (m6Am).

1.8.1 Implications of mRNA Methylation

The scope of the ramifications of mRNA methylation is still somewhat nebulous. Examples of their impact have been seen in pathological behavior (Engel et al., 2017), stem cell development (Liu et al., 2018), neural development (M. Li et al., 2018) and sex determination (Haussman et al., 2016). There are likely many yet undiscovered roles of RNA methylation, yet there is a general consensus that these modifications impact expression of proteins. The methods in which they are controlling expression are numerous and not perfectly understood, yet research is being actively performed in the field and our understanding is growing quickly. Here the main ways methylation is influencing expression will be introduced.

1.8.1.1 Splicing

Control of mRNA splicing was one of the earliest forms of regulation found to be a result of mRNA methylation. As early as the 1980s, methylation was being shown to have an impact on the splicing and function of mRNAs by the inhibition of S-adenosylmethionine (SAM), a compound involved in methyl group transfer (Stoltzfus et al., 1982). Since then, the role of m6A and other methylation marks has been greatly expanded upon. Drosophila gene *Sxl* was shown to depend on m6A when determining gender. The presence of m6A was required for alternative splicing that allowed for female sex determination. The removal of a methyltransferase complex showed sex bias towards males (Haussman et al., 2016). The methyltransferase complex described in that paper has a human homolog in the METTL3, METTL14 and WTAP complex, to be discussed later in the thesis.

1.8.1.2 mRNA export

The export of proteins may be affected by the presence of a methylation mark on mRNA. An example of this is seen with mRNA marked with 5-methylcytosine (m5C). The m5C marked mRNA is recognized by the protein ALYREF, an mRNA transport adaptor, which then shuttles the mRNA out of

the nucleus. mRNA transcripts lacking the m5C modification are not recognized and therefore not shuttled out (Yang et al., 2017). This has been implicated with the modifications to be discussed in this thesis. An example is seen in the absence of ALKBH5, mRNA transcripts had increased levels of m6A and displayed accelerated nuclear export, seen as increased levels of mRNA in the cytoplasm (Zheng et al., 2013). Nuclear export of m6A marked mRNAs was also shown to be decreased following the knockdown of YTHDC1, a protein that contains a YTH domain that recognizes m6A (Roundtree et al., 2017). YTHDC1 was also recently shown to target m1A, indicating there may be a role for m1A in the nuclear export process (Dai et al., 2018).

1.8.1.3 Translation promotion and inhibition

The impact methylation has on translation has been trickier to decipher. There is currently much debate as to whether methylation promotes or inhibits translation efficiency, but the answer is probably both due to varying mechanisms of processing methylated mRNA as well as the location of methylation potentially having an impact. In 2017, Slobodin et al. proposed that m6A was being added to mRNA cotranscriptionally, and that slower transcription allowed for increased presence of m6A on the mRNA. These increased levels would then induce decreased translation efficiency (Slobodin et al., 2017). In the same year, Coots et al. published a paper showing that m6A presence in the 5' UTR of mRNA facilitated and promoted translation (Coots et al., 2017). In the same year, Shi et al. showed the importance of m6A presence for recognition by YTH Domain Family proteins, specifically YTHDF1 and YTHDF3. These proteins recognized m6A bound transcripts and would promote translation. The knockdown of these proteins resulted in the accumulation of m6A bound transcripts in the cytoplasm (Shi et al., 2017). Clearly the exact mechanisms of regulating translation are still being defined, but the consensus is that these modifications may be important for the translation of certain mRNA targets.

1.8.1.4 Stability

Similar to translation, mRNA methylation can have effects on both ends of the spectrum. Transcripts marked with m6A have been shown to be selectively targeted by YTHDF2, which promotes degradation of mRNA through delivery to P-bodies (Wang et al., 2014). Conversely, m6Am (N6,2'-O-dimethyladenosine) has been shown to enhance the stability of transcripts

marked with m6Am by repressing the activity of the decapping enzyme DCP2 (Mauer et al., 2016). Again, the extent to which these modifications regulate stability appear to be varied and dependent on the type and location of the modification.

1.8.2 N1-methyladenosine

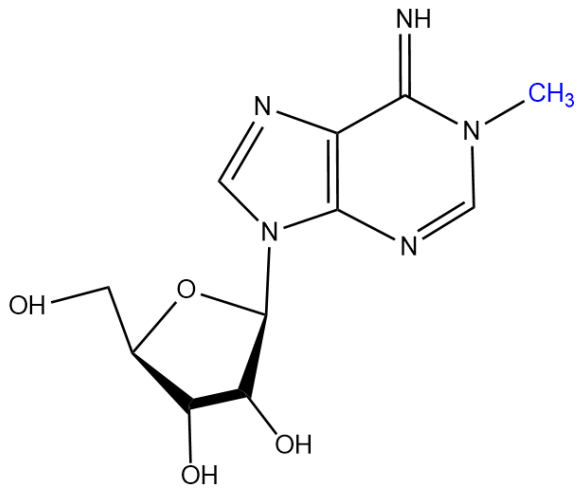


Figure 4. Structure of N1-methyladenosine. The methyl group is bound at the N1 position of the ribose component.

N1-methyladenosine (m1A) is a modified adenosine nucleotide. The modification is achieved through the addition of a methyl group onto the 1' nitrogen of the ribose component of an adenosine nucleotide. It is also worth noting that this modification of the nucleotide results in a positively charged adenosine. The modification itself was found in RNA several decades ago (Desrosiers 1974), and has since been shown to exist more specifically in mRNA, tRNA and lncRNA. Among RNA modifications, m1A has currently a rather unclear impact on cellular function when it is present. When m1A is located near the 5' cap of mRNA, it is thought to promote translation of mRNA (Dominissini et al., 2017). While it is also found across all other segments of mRNA, the impact of the modification in these positions is unknown.

There is also currently debate as to the prevalence of the modification in mRNA. In 2016, a paper detailing the presence and dynamic nature of m1A

in mRNA was published. This paper concluded that ~20% of human mRNA transcripts contained a m1A modification, but the impact was unknown (Li et al., 2016). This was later refined down to a much lower number by two different groups (Safra et al., 2017; Li et al., 2017). The same group who published the initial paper used a method involving the presence and absence of a isolated AlkB enzyme followed by the use of a reverse transcriptase, Superscript III, that would produce a truncated cDNA product due to its inability to read through the m1A modification. This was then compared against another reverse transcriptase, TGIRT, that was able to read through the modification and produce non-truncated cDNA products. By comparing the truncated and non-truncated products, and then applying their own calculations, they were able to identify positions which were likely to be modified by m1A. This reduced the number of transcripts modified from ~20% of all mRNA to a mere 17, largely mitochondrial, transcripts (Li et al., 2017). Another group performed a similar experiment, except rather than use isolated AlkB they relied on Dimroth rearrangement, a phenomenon that causes N1-methyladenosine to transition into N6-methyladenosine (m6A) in alkaline conditions (described in further detail later in the thesis). Whereas m1A causes truncation in cDNA synthesis, m6A does not. This method described only 15 mRNAs as being modified by m1A, 10 of which were cytosolic and only 5 of which were in the mitochondria (Safra et al., 2017). The debate today largely revolves around this abundance and whether the methods used for detection are reliable. Some groups have said m1A exists quite abundantly within mRNA as it is detected by LC/MS techniques, whereas others dismiss those signals as noise resulting from tRNA contamination. While the debate continues as to what the best method for m1A site identification is, there clearly is no consensus. Pitfalls within the analyses of each method likely lead to the exclusion of lower expressing RNAs. The methods used by Safra and Li, while potentially flawed or lacking sensitivity, provide several important ideas. One of them being that m1A, but not m6A, has an impact on the production of full-length cDNA when using certain reverse-transcriptases. Another important takeaway is that RNA modification tends to be a very specific and targeted event, rather than just a random result of some chemical interaction.

1.8.3 ALKBH3 in RNA demethylation

As mentioned previously, N1-methyladenosine on RNA is a target of ALKBH3. Interest in this role, which has been known about for years, increased along with the increased focus on the impact of RNA methylation in recent years. The previously mentioned papers from Li (2016) and Dominissini (2016) that

focused on the presence of m1A within mRNA additionally identified ALKBH3 as the demethylase responsible for the dynamic regulation of this modification. Dominissini showed the overexpression of wild-type ALKBH3 to cause a decrease global levels of m1A in mRNA (Dominissini et al., 2016). Li (2016) showed a similar result, but additionally showed ALKBH3 had a slight preference for m1A on RNA to ssDNA, allowing for a divergence from the established idea of the main role of ALKBH3 being a DNA alkylation damage repair protein. Li also suggests that some m1A sites are induced by stress, in their case caused by UV damage (Li et al., 2016). Still the impact of this interaction between protein and modification remain nebulous and there are yet to be specific examples of proteins regulated by this pathway.

1.8.4 N6-methyladenosine

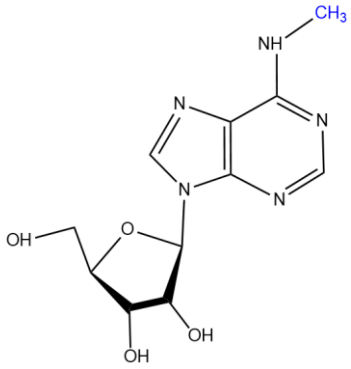


Figure 5. Structure of N6-methyladenosine without phosphate component.

Of the RNA modifications discussed in this thesis, N6-methyladenosine (m6A) is by far the most studied. It was initially identified in 1974, but was more or less forgotten about until the invention of a methylated RNA immunoprecipitation sequencing technique was established in 2012 (Desrosiers, 1974; Dominnissini et al., 2012; Meyer et al., 2012). Since then interest in the modification has greatly expanded and it has become a key part of understanding epitranscriptomic regulation. The modification is highly pervasive and has been shown to exist in the transcripts of more than 7000 human genes (Dominnissini et al., 2012). It has been implicated in a vast variety of roles, such as cellular differentiation, mRNA stability, translation control and alternative splicing (Batista et al., 2014; Zhou et al., 2018; Alarcon

et al., 2015; Xiao et al., 2016; Tang et al., 2017). Similar to m1A, m6A has been shown to be dynamically regulated within mRNA, primarily by ALKBH5 (Zheng et al., 2013). FTO was originally thought to demethylate m6A but has since been shown to preferentially target a separate modification, N6,2'-O-dimethyladenosine (Mauer et al., 2016). Despite the implications of m6A in various processes, ALKBH5 knockout animals appear normal aside from defects in spermatogenesis (Zheng et al., 2013; Tang et al., 2017). This possibly suggests that m6A is not as crucial as many believed, but also certainly implies that the process in which RNA methylation is regulated is a far more complex process than currently understood.

The research done on m6A provides an important framework for what methods are available to study RNA modifications as well as what the potential regulatory implications are of mRNA methylation.

1.8.5 Dimroth Rearrangement

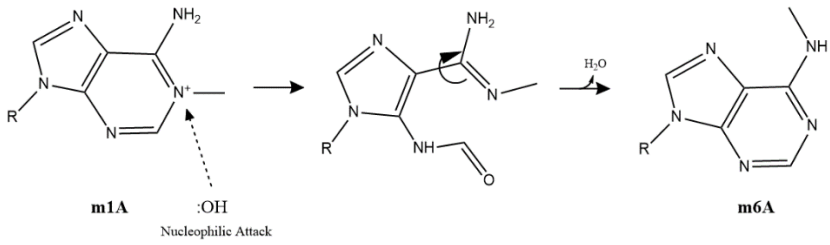


Figure 6. Dimroth rearrangement. Nucleophilic attack on the N1 position of the ring causes the ring to open and spin around, moving the methyl group to the N6 position (Macon 1968).

Dimroth rearrangement is the event in which m1A is able to transition into m6A when in the presence of a alkaline environment. This occurs when there is a nucleophilic attack that opens up the carbon ring in the ribose and reforms into an aromatic ring after spinning around. Importantly, this reaction is known to only proceed from the m1A to m6A direction and not the reverse. This is possibly due to the formation of an aromatic ring following the switch which is much more chemically stable and thermodynamically difficult to open. While this event has been proven to occur *in vitro*, it is difficult to prove *in vivo* but it

is reasonable to assume it occurs *in vivo* as well.

1.8.6 N6,2′-O-dimethyladenosine

N6,2'-O-dimethyladenosine is another form of modified adenosine. It is structurally identical to N6-methyladenosine, but has an added methyl group at the 2' oxygen of the adenosine ribose component.

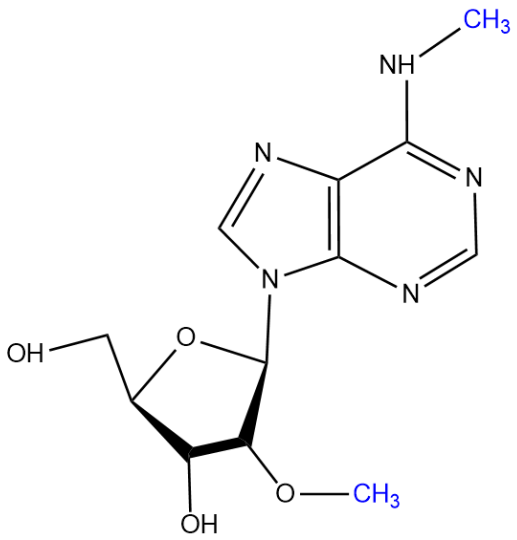


Figure 7. Structure of N6-2'-O-dimethyladenosine (m6Am). It is structurally identical to N6-methyladenosine but has a additional methyl group at the 2'-Oxygen. Shown to be preferentially demethylated by FTO (Mauer et al., 2016).

Functionally, m6Am has been identified as having a couple roles. When located in the 5' cap of the mRNA, m6Am prevents DCP2 mediated decapping of mRNAs, therefore increasing the mRNA stability (Mauer et al., 2016). The same paper showed m6Am also affects miRNA binding and regulation by providing resistance miR-155- mediated RNA degradation (Mauer et al., 2016).

As mentioned previously, FTO has been shown to preferentially demethylate m6Am over m6A. To reiterate, this is important in that it separates FTO demethylation targets from that of ALKBH5, currently solely known as a m6A RNA demethylase.

1.9 Methylation Writers

The methylation of RNA does not seem to be a random event. While some misplaced alkylation damage may be incurred from exogenous sources, most RNA methylation appears to be an extremely targeted and regulated event. Methylation writers are complexes that contain a methyltransferase. Methyltransferases can be split into several classes. Class I methyltransferases function by having a Rossman domain that allows for the binding of S-adenosyl methionine, a natural donor of methyl groups (Kozbial 2005). There are multiple complexes involved in methyltransferase activity with mRNA, each with a specific area of methyl addition.

1.9.1 METTL3, METTL14, RBM15 and WTAP

METTL3 (methyltransferase-like 3) and METTL14 (methyltransferase-like 14) are two proteins involved in the transfer of m6A onto RNA. They, together with RBM15 (RNA Binding Motif 15) and WTAP (Wilms Tumor 1-associated protein), form a complex that allows for the depositing of m6A onto specific mRNAs. METTL14 was originally thought to be a methyltransferase, but has since been shown to be enzymatically inactive, leaving METTL3 as the sole methyltransferase component of the complex (Sledz et al., 2016; Wang et al., 2016; Jaffrey et al., 2017). METTL14 has instead a adaptor role, binding the RNA substrate and promoting the activity of METTL3 (Jaffrey et al., 2017). WTAP has a role as an adaptor, coupling METTL3/METTL14 with RBM15 (Patil 2016). RBM15 has a role in the recognition of RNA motifs, allowing for the targeting of specific sequences within RNA transcripts by the m6A writer complex. The knockdown of METTL3, METTL14 and WTAP all show decreased (30-50%) levels of global m6A, suggesting each play a key role in the proper function of the complex (Liu et al., 2017). Additionally, due to there not being complete reduction, this data may suggest there are yet to be researched proteins involved with m6A deposition, indicating these pathways are more complex than currently understood.

1.9.2 TRMT6, TRMT61A and TRMT10C

The methyltransferases responsible for depositing methylation onto the N1 position of adenine are TRMT6, TRMT61A and TRMT10C. TRMT6 and TRMT61A are thought to complex together where as TRMT10C likely has a redundant methyltransferase function but is localized differently in the cell (Safra et al., 2017; Li et al., 2017). TRMT stands for tRNA methyltranferase,

as these proteins were initially thought of as targeting exclusively tRNAs. In Saccharomyces cerevisiae, GCD10 (TRMT6 homolog) and GCD14 (TRMT61A homolog) are responsible for transferring methyl groups onto tRNA. These methyl groups are believed to be important for tRNA stability (Ozanick et al., 2005). Since then, TRMT6 and TRMT61A have been shown to target mRNA as well as tRNA (Safra et al., 2017). TRMT10C seems to preferentially target mRNAs located in the mitochondria. Intriguingly, the mRNAs targeted have been shown to have a tRNA T-loop motif structure within the mRNA itself (Safra et al., 2017).

1.9.3 Consensus motifs

What is common between both of these methyltransferase complexes is their recognition of specific nucleotide sequences referred to as a consensus motif. These motifs, possibly in combination with mRNA folding patterns allow for the specific targeting of sequences on certain mRNAs to have a regulatory effect. The agreed upon consensus motif for the METTL3-METTL14-WTAP complex is the DRACH motif (D= A, G, or U; R = A or G; H = A, C, or U) (Linder 2016). The motif identified for the TRMT complex is GUUCRA (Li et al., 2017; Safra et al., 2017). Consensus motifs for specific recognition extend beyond methyltransferase complexes as well. For example the Serine-Arginine Rich Splicing Factor (SRSF) protein family consists of 12 different proteins, each having a unique consensus motif (Long et al., 2009). These motifs are still being defined for a great number of proteins, but likely play a role of growing importance for the specificity in which RNA interacting proteins operate.

1.9.4 T-loop motif in tRNA and mRNA

TRMT6 and TRMT61A in particular are interesting because of GUUCRA motif in T-loop structures. While the proteins were initially thought to just target tRNAs (hence the name tRNA Methyltransferase 6), they have since been shown to target secondary structures within mRNA

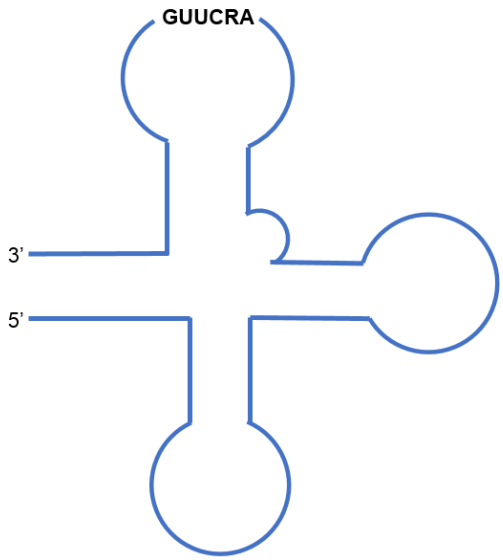


Figure 8. GUUCRA motif in T-loop of tRNA targeted by TRMT6/TRMT61A/TRMT10C methyltransferase complex to deposit N1-methyladenosine.

Safra et al. identified GUUCRANNC motif, which was refined to GUUCRA by Li et al. (Safra et al., 2017; Li et al., 2017). Safra importantly identified several mRNAs that were being recognized from this motif and modified with m1A. Interestingly, it is suggested that these mRNAs contained secondary "T-loop-structures" that were targeted by the TRMT6/TRMT61A/TRMT10C complex. This opened up the idea that the m1A methyltransferases were targeting more than tRNAs, but also mRNAs as well (Safra et al., 2017).

Figure 9. T-loop structures containing the GUUCRA motif exist in mRNA. This allows mRNA to be targeted by the m1A methyltransferase complex

1.10 Methylation Readers

1.10.1 YTH Domain proteins

Multiple proteins that belong to the YTH Domain Family of proteins have been established as readers of m1A and m6A (Dai et al., 2018). The YTH (YT521-B homology) domain is present within 174 different proteins and is evolutionarily conserved among many eukaryotic species (Stoilov et al., 2002). The domain itself is typically located in the middle of proteins and has been found to be involved in the recognition of RNA molecules (Stoilov et al., 2002; Liao et al., 2018). The 5 YTH domain proteins relevant to this thesis can be split into 3 categories. These are the YTH DF Family Proteins 1-3 (YTHDF1, YTHDF2, YTHDF3), YTH Domain Containing Proteins YTH DC1 family (YTHDC1) and YTH DC2 family (YTHDC2) (Liao et al., 2018). These proteins were initially identified to interact with m6A containing mRNAs but have been, with the exception of YTHDC2, shown to interact with m1A as well (Dai et al., 2018). The proteins themselves serve different functional roles.

1.10.1.1 YTHDF1

YTHDF1 is primarily implicated in translation efficiency for m6A bound mRNA transcripts. It has been shown to actively promote translation by recognizing m6A and then recruiting eIF3 (eukaryotic initiation factor 3), which in turn recruits eIF4G (eukaryotic initiation factor 4G) and ribosomal subunit 40S to begin translation (Wang et al., 2015).

1.10.1.2 YTHDF2

YTHDF2 is involved in the stability and degradation of m6A tagged mRNAs. YTHDF2 has been shown to recognized m6A on mRNAs in the cytoplasm and then proceed to shuttle those mRNAs to processing bodies (P-bodies) (Sheth et al., 2003; Wang et al., 2014). P-bodies are cytoplasmic and considered mRNPs (messenger ribonucleotide proteins) because they are always found to be bound to mRNAs (Martinez et al., 2013). The importance of the interaction between YTHDF2 and P-bodies may be that it promotes the recruitment of the CCR4-NOT deadenylase complex to promote the degradation of mRNA (Du et al., 2016). The interaction of YTHDF2 with m6A

has also been implicated as a key part of neural development in mice, with conditional depltion of *Ythdf2* in mice causing lethality in late developmental stages (M. Li et al., 2018).

1.10.1.3 YTHDF3

This is perhaps the least studied of the YTH Domain Family proteins. YTHDF3 has been shown to interact with both YTHDF1 and YTHDF2, therefore involving it in both mRNA translation as well as mRNA stability. YTHDF3 is proposed to interact with YTHDF1 in the recognition of m6A modified mRNAs, and then recruit the 40S and 60S ribosomal subunits to promote translation (Li et al., 2017). It has been proposed that it may interact with additional unknown proteins to inhibit translation as well (Shi et al., 2017). Shi (2017) proposes that YTHDF3 may be present to recognize m6A modified mRNAs and help accessibility to those mRNAs by YTHDF1 and YTHDF2.

1.10.1.4 YTHDC1

YTHDC1 is a nuclear reader of m6A (Xiao et al., 2016) and m1A (Dai et al., 2018). It binds methylation marked mRNAs and promotes the recruitment of exon including pre-mRNA splicing factor SRSF3 (Serine-Argenine Rich Splicing Factor 3) and blocks the recruitment of exon excluding SRSF10 (Serine-Argenine Rich Splicing Factor 10) (Xiao et al., 2016). SRSF3 recruits and complexes with NXF1 (Nuclear Export Factor 1), involving an aspect of mRNA export into the recognition of methylation marks (Roundtree et al., 2016). NXF1 and YTHDC1 showed no interaction. Knockdown of YTHDC1 resulted in extended nuclear residence of m6A containing mRNAs, implicating that methylation marked mRNAs undergo a more specific and specialized nuclear export process (Roundtree et al., 2016).

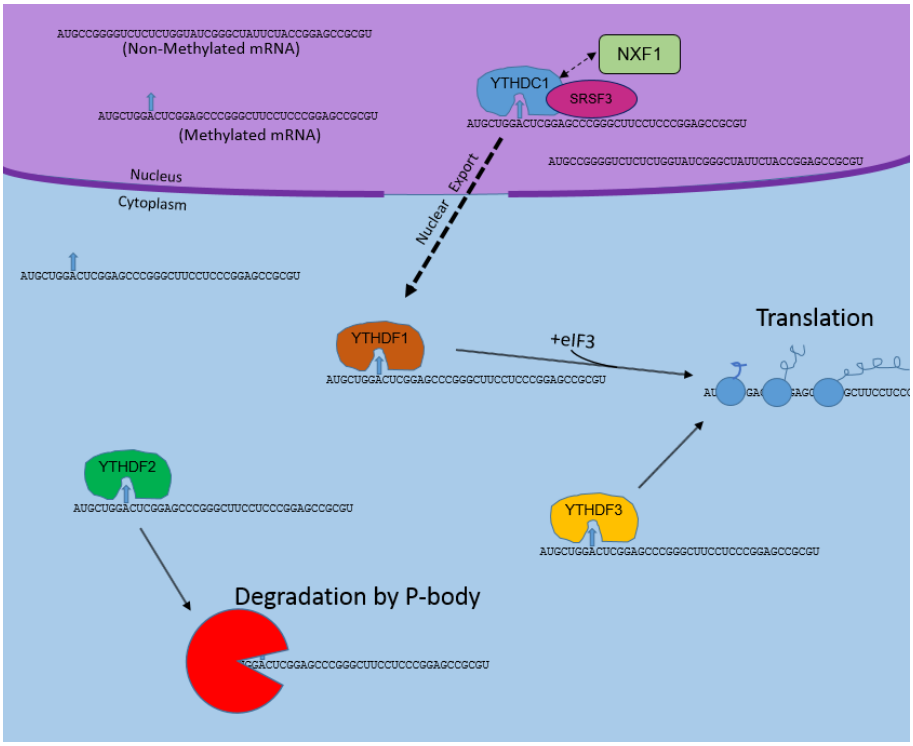


Figure 10. The differing roles of YTH proteins. YTHDC1 involved in nuclear export. YTHDF1 and YTHDF3 involved in translation. YTHDF2 involved in mRNA degradation.

2 Aims

The initial aim of the project was to examine the epigenetic silencing of ALKBH3 in tumor tissue, as discovered through The Cancer Genome Atlas. Instances of DNA repair genes being epigenetically silenced by promoter methylation have been linked to cancer (Stefansson et al., 2011; Cunningham et al., 2014; Lalloo et al., 2012; Muggerud et al., 2010), yet no data had been published on the epigenetic silencing of ALKBH3. The first part of the project aimed to characterize to what extent promoter hypermethylation affects the expression of ALKBH3. Furthermore we aimed to characterize this specifically within a subset of Icelandic breast cancer tumors.

The second aim of the project was then to determine what role, if any, ALKBH3 is playing in the development and progression of these tumors. This involved elucidating what the functional role of ALKBH3 is and what the consequences of its absence, through events like promoter methylation, may have on cells. ALKBH3 had a previously established functional role in alkylation repair (Aas et al., 2003; Dango et al., 2013) and we aimed to examine what the consequences of decreased alkylation repair may be. We additionally discovered that ALKBH3 may have a role in DSB repair. This novel connection shifted the focus of this aim to finding out how ALKBH3 was involved in this pathway and furthermore examine once again what the functional consequence of ALKBH3 depletion in the context of promoter hypermethylation may have on a cell's ability to perform effective DSB repair.

We identified ALKBH3 as a regulator of RNF168, a key protein in DSB signalling and repair pathway choice. The focus was then to determine the mechanism behind this regulation and if it shared similar phenotypic consequences to those that had been previously detailed in literature. By eliminating changes in other forms of regulation, we were able to show that RNF168 is being regulated by ALKBH3 through RNA methylation. This led the project into a field recently dubbed epitranscriptomics. We then further explored how this phenomenon actually regulates RNF168 and elucidate this new role of ALKBH3 in the removal of methylation marks from mRNA as a regulatory mechanism.

2.1 Specific Aims

1 Confirm ALKBH3 promoter hyper-methylation within Icelandic breast cancer samples and measure impact on protein expression in cell lines.

2 Determine potential functional impact of ALKBH3 silencing on tumors.

3 Elucidate mechanism of regulation of RNF168 by ALKBH3.

3 Materials and methods

3.1 Cell cultures

The cell lines used were obtained from the American Type Culture Collection. The cells were cultured at 37°C in DMEM or RPMI with added 10% fetal bovine serum (+penicillin/strepto-mycin (1.5 µg/mL)), or DMEM with 2mM Glutamine, 250 ng/mL insulin, 10 µgmL transferrin, 10 E-8M Sodium Selenite, 10 E-10M 17 beta-estradiol, 0.5 µg/mL hydrocortisone, 5 µg/mL ovine prolactin and 10 ng/mL EGF (H14) as detailed in Table 1.

Table 1. Cell lines used for study.

Cell Line	Tissue Type	Media
BT-20	Breast	DMEM
BT-474	Breast	RPMI
CAMA-1	Breast	DMEM
EFO-27	Ovarian	RPMI
HCC-1428	Breast	RPMI
HCC-1500	Breast	RPMI
HCC-38	Breast	RPMI
HMT-3522 S1	Breast	H14
MCF-10A	Breast	DMEM
MCF-7	Breast	DMEM
MDAMB-134VI	Breast	DMEM
MDAMB-231	Breast	DMEM
MDAMB-435	Breast	DMEM
MDAMB-436	Breast	DMEM
MDAMB-468	Breast	DMEM

OVCAR-3	Ovarian	RPMI
OVCAR-8	Ovarian	RPMI
PA-1	Ovarian	DMEM
SkBr3	Breast	DMEM
T47D	Breast	DMEM
U2OS	Bone	DMEM
UACC-3199	Breast	DMEM

3.2 Study cohort

DNA samples were derived from primary breast tumors (n = 265) and adjacent normal breast tissue (n =30). The normal breast tissue was obtained from non-tumorous regions of the breast. The DNA was previously isolated from freshly frozen tissue following a standard protocol based on phenol-chloroform (+proteinase K) extraction. RNA samples were available for a subset of the tumor (n = 36) and normal breast (n = 10) tissue samples isolated using Tri-Reagent (Thermo Fisher Scientific). Clinical parameters, including tumor size, nodal status, histological grade along with disease-specific follow-up times were obtained from the nationwide Icelandic Cancer Registry. This work was carried out according to permits from the Icelandic Data Protection Commission (2006050307) and Bioethics Committee (VSNb2006050001/03–16). Informed consent (written) was obtained from all patients.

3.3 Informatics

Information on CpG methylation over the promoter region of ALKBH2 and ALKBH3 was obtained from preexisting methylome analyses published by Stefansson et al.. (GSE52865), Dedeurwaerder et al.. (GSE20713) and Teschendorff et al.. (GSE69914) available through the Omnibus repository at NCBI's website (www.ncbi.nlm.-nih.gov/gds/) (Stefansson et al., 2015; Dedeurwaerder et al., 2011; Teschendorff et al., 2016). The normalized data were extracted from the SOFT formatted files using the GEOquery package in R and analysed by comparing normal breast tissue samples and breast cancers. This was carried out using the Student's t-test on M–values computed using Mi = log2(Bi / (1-Bi)) where B represents the β-value coupled with the Benjamini-Hochberg adjustment procedure to account for multiple hypothesis

testing making use of the p.adjust function in R. The multiple hypothesis adjustment accounted for the total number of CpGs represented on the array platform, i.e. adjusting for the entire >450 thousand CpGs in GSE52865 and GSE69914 and >27 thousand CpGs in GSE20713.

DNA methylation (450 K Infinium) and RNAseq (V2) level 3 data were downloaded from the Cancer Genome Atlas data repository (http://cancergenome.nih.gov/) (Koboldt et al., 2012). Firstly, the analysis of differential ALKBH3 mRNA expression levels between normal breast tissue and breast cancers was carried out using the Wilcoxon's rank sum hypothesis test taking into account adjustment for multiple hypothesis testing including the entire set of >20 thousand genes included in the RNAseqV2 dataset. This was carried out using the Benjamini-Hochberg (BH) procedure through the p.adjust function in R. Secondly, the relation between CpG methylation for each site represented over either ALKBH2 and ALKBH3 were studied with respect to ALKBH3 mRNA expression using Spearman's correlation analysis and, as before, with genome-wide adjustment of the P-values using the BH procedure to account for multiple hypothesis testing.

Information on epigenetic marks for the ALKBH3 promoter region in variant human mammary epithelial cells (vHMEC) was obtained from the Roadmap Epigenomics browser (egg2.wustl.edu/roadmap/web_portal/) (Kundaje et al., 2015). This includes information on ALKBH3 expression based on RNA sequencing and chromatin marks based on ChIPseq along with data on chromatin accessibility based on DNA sequencing. Data on CpG methylation for the vHMEC cells was derived from methylCRF computational analysis using MeDIP-seq and MRE-seq data to infer whole-genome 5-methylcytosine states as carried out and provided by the Roadmap Epigenomics project (Kundaje et al., 2015). Information on nucleotide positions for ALKBH3 gene structure (introns and exons) was downloaded from Ensembl (GRCh37 browser; HG19). Data on transcriptional start site (TSS) and CpG islands were obtained from the FANTOM5 promoterome (Lizio et al., 2015). Using the UCSC genome browser, the chromStart/chromEnd fields in the hg19.cpgIslandExt table provided the CpG island positional information. The R statistical software (R 3.1.0) was then used to graphically represent the ALKBH3 promoter with respect to the TSS, 1st Exon and CpG island.

3.4 siRNA knockdown

All siRNAs were designed by Thermo Fisher Scientific. The siRNAs belonged to the Silencer Select product line. Below are the siRNAs listed with their Assay

ID number (Table 2).

Table 2. siRNAs and their ID#.

Gene	Assay ID#
ALKBH2	s42494
ALKBH3 #67	s47967
ALKBH3 #68	s47968
ALKBH3 #69	s47969
ALKBH5	s29688
ASCC3	s21605
FTO	s35511
hnRNP K	s6737
METTL14	s33680
METTL3	s32141
Negative Control #2	4390846
RNF168	s126171
SRSF3	s12733
THRAP3	s19361
TRMT10C	s29784
TRMT6	s28398
TRMT61A	s41857
WTAP	s18433
YTHDC1	s40757
YTHDF1	s29743
YTHDF2 #1	s29147

YTHDF2 #2	s28148
YTHDF3	s48464

All siRNAs were diluted to a concentration of 10 µM. Transfections were optimized for each size of culture dish for optimal cell growth and always treated with equal volumes of RNAiMax (Thermo Scientific, Cat# 13778075). Cells were seeded 24 hours prior to transfection in order to be at 50% confluency at the time of transfection. Transfection was carried about by diluting the siRNA in Opti-MEM (Thermo Scientific, Cat# 31985062), , then diluting RNAiMAX with the same volume of Opti-MEM and then combining the two volumes and incubating for 15 minutes. The cells were incubated in the presence of siRNA for 48 or 72 hours before performing fixation or extraction dependent on the experiment. Refer to Table 3 and indicate final concentration of siRNA in the medium.

Table 3. Transfection conditions

Culture Dish	Vol of 10µM siRNA/RNAiMAX	Vol Opti-MEM (each)	Vol Media
60mm	5 µL	250 µL	4.5 mL
100mm	10 µL	500 µL	9 mL
8 well glass chamber	0.25 µL	25 µL	150 µL
24 well plate	1.5 µL	50 µL	400 µL
6 well plate	2.5 µL	125 µL	1.75 mL

3.5 Western Blot

Cells for western blot experiments were grown in 60mm culture dishes. Samples were extracted directly into loading buffer. The loading buffer contained NuPAGE™ LDS Sample Buffer (4X) (Thermo Scientific, Cat# NP0007), diluted to 1X in ddH2O along with 10X dithiothreitol (DTT) reducing agent, also diluted to 1X. Samples were washed once with PBS and put on ice, any excess liquid was removed. The loading buffer was heated to 95°C

and then 200µL was placed onto samples followed by cell scraping. 0.2µL of Benzonase Nuclease (Santa Cruz, sc-202391) was added to each sample and incubated at room temperature for 1 minute. Samples were then placed in a heat block at 95°C for 10 minutes. Polyacrylamide gels were prepared prior to each experiment, acrylamide concentration varying from 6% to 12% based on target protein size, with a stacking gel of 15%, and gel thickness of 1.5mm. 10x Running Buffer (composed of 30.0g Tris base, 144.0 g of glycine and 10.0 g of SDS in 1000 mL of H2O) was diluted to 1X in a total volume of 1L. 45µL of samples were loaded alongside 7.5µL of ladder (Color Prestained Protein Standard, NEB, Cat# P7712S). The gel was run at 90V until ladder was sufficiently resolved and target proteins were in the middle of the gel (typically 90-120 minutes). For transfer, 10X transfer buffer(25mM Tris, 192 mM Glycine and 20% methanol in H2O) was diluted to 1X, combined with 20% methanol and H20. In the case of larger protein western blots 0.05% SDS was added. Whatmann filter paper and sponges were wet in the buffer prior to the transfer and chilled at -20°C. The transfer was onto a Nitrocellulose Pure Transfer membrane (Santa Cruz, sc-3724). Transfers were done for 75 minutes at 150mA (per membrane). Following transfer the membrane was rinsed in PBS, then blocked in 5% milk in PBS for 1 hour. Antibodies were then diluted to their respective dilution (Table 4) in 5% milk and left on the membrane overnight. The membrane was then washed 4 times, for 5 minutes each time, in PBS with 0.1% Tween. The HRP secondary antibodies were diluted in PBST 1:10000 and placed on the membrane for 1 hour. This was followed by an additional 4, 5 minute washes in PBST. Western Blotting Luminol Reagent (Santa Cruz, sc-2048) was added for 1 minute. The membrane was imaged in a Biorad Universal Hood II Gel Doc System using the software Image Lab taking an image every 30 seconds for 15 minutes.

3.5.1 Western Antibodies

Table 4. Antibodies used for detection in western blots.

Gene	Company	ID #	Dilution
PRIMARY			
Actin	AMD Millipore	MAB1501R	1:10000
ALKBH3	AMD Millipore	09-882	1:500

ASCC3	Bethyl	A300-569A	1:500
FLAG M2	Sigma Aldrich	088K6018	1:1000
FTO	Abcam	ab126605	1:1000
RIF1	Bethyl	A300-569A	1:1000
RNF168	AMD Millipore	ABE367	1:500
RNF8	Santa Cruz Biotechnology	sc-133971	1:1000
SMC1	Abcam	ab32219	1:10000
53BP1	Santa Cruz Biotechnology	sc-22760	1:2000
Vinculin	Santa Cruz	sc-5573	1:250
SECONDARY			
HRP Secondary Anti-mouse	Santa Cruz Biotechnology	sc-2096	1:10000
HRP Secondary Anti-rabbit	Santa Cruz Biotechnology	sc-2313	1:10000

3.6 Bisulfite conversion and Pyrosequencing

Bisulfite conversion was carried out using the EZ-96 DNA Methylation-Gold kit(Zymo Research, D5008). The bisulfite conversion was done in a PCR for 16 cycles of (95 °C for 30s and 50 °C for 1 h) to then hold at 4 °C until samples were added to the DNA columns for completing the conversion following the manufacturer's guidelines (Zymo Research).

The PyroMark Assay Design 2.0 software was used to design primers for the analysis of ALKBH3 promoter methylation. The following primer sequences were used: 5′-(Btn)-GTGGGATTATTAGGATTGAG GATT-3′ (5-biotin labelled) and 5′-CTCCAACAACTCC CAATCAC-3′. The pre-amplification PCR reaction was carried out using a hot-start PCR polymerase (Immolase DNA polymerase from Bioline; Bio-21,047). The PCR conditions were as follows: 96 °C for 10 min, 45 cycles of (96 °C for 30s, 60 °C for 30s and 72 °C for 30s) followed by 15 min hold at 72 °C and then 4 °C. The PCR products were then captured using streptavidin coated agarose beads (GE Healthcare,

Streptavidin Sepharose High Performance 34 µm beads) under denaturing conditions to obtain single-stranded DNA. The pyrosequencing reaction was then carried out using the PyroMark Q24 machine (Qiagen) and PyroMark Gold-Q24 Reagents kit (Qiagen) using the following sequencing-primer: 5′-ACATCAAA CACTTCCT-3′.

CpG methylation for three CpG's were assessed (−58, −53 and −50 bp upstream of the TSS (p1) given the FANTOM5 promoterome database) (Lizio 2015). The output data (obtained from PyroMark Q24 sequencing reactions), representing percent methylated cytosines over each of these three CpGs, was averaged for each sample analysed. This yielded a single measure representing a proxy for CpG methylation levels over the ALKBH3 promoter region. The statistical analysis of paired tumor and normal breast tissue samples made use of a paired Wilcoxon's test using the wilcox.test function in R.

3.7 Immunofluorescence and confocal microscopy

Cells were grown on coverslips (Thermo CB00100RA020MNT0). The cells were fixed for 10 minutes in 4% paraformaldehyde and washed in PBS (phosphate buffered saline) twice. Cells were then permeabilized in 0.2% Triton x100 in PBS for 5 minutes and rinsed 3 times with PBS. Cells were then blocked in DMEM media +10% FBS for 60 minutes. The cover slides were then incubated with the primary antibodies for 60 minutes. The slides were washed 3 times in PBS and incubated with the secondary antibodies and DAPI for 60 minutes. The slides were then washed 3 more times in PBS followed by a wash in ddH2O. The samples were allowed to dry for 15 minutes before being mounted to slides using mounting medium (Santa Cruz sc-516212)

3.7.1 Antibodies for immuno-fluorescence Staining

Table 5. Antibodies used in immufluorescent staining.

Antibody	Company & Catalog #	Dilution
PRIMARY		
3-methylcytosine	Active Motif 61111	1:250

5-methylcytosine	Abcam ab10805	1:250
53BP1	Santa Cruz sc-22760	1:1000
γH2AX	Abcam ab22551	1:1000
ALKBH3	Millipore 09-882	1:250
CyclinA	Santa Cruz sc-271682	1:200
DAPI	Thermo Fisher D1306	1:10000
MDC1	Abcam ab11171	1:2000
RNF168	Millipore ABE367	1:250
SECONDARY		
Alexafluor 488 anti mouse	Life Technologies A21121	1:1000
Alexafluor 555 anti rabbit	Life Technologies A21434	1:1000

The primary antibodies used were γH2AX Mouse IgG1 (ab22551 from Abcam; UK) and RAD51 Rabbit IgG (ab63801 from Abcam; UK). Secondary antibodies were Alexaflur 488 anti-mouse IgG1 (A21121; Life technologies) and Alexafluor 546 Anti-rabbit IgG (A21434; Life technologies).

All confocal microscopy was performed in a Olympus FLV1200 confocal microscope. Images were analyzed using ImageJ and Cell Profiler 3.0. ImageJ was used to quantify signal intensity while Cell Profiler was used to quantify foci.

3.8 DNA alkylation damage detection

CAMA1 and MDA-MB-468 were grown on coverslips and fixed with freshly prepared 4% para-formaldehyde solution for 15 min. After fixation, cells were treated with 1.5 M HCL for 20 min in order to gain access to single stranded DNA, followed by a 2 minute treatment with Sodium Borate (pH 8.5) to neutralize the acid. After permeabilization (5 min, 0.2% TritonX) and 1 h of blocking (DMEM (Thermo Scientific 12491-015) with 10% FBS (Thermo Scientific A38401)), cells were stained with antibodies against 3-methylcytosine (m3C) (rabbit, Active Motif, 61111) and 5-methylcytosine (m5C) (mouse, abcam, ab10805) for 1 h at room temperature. Both antibodies

were diluted 1:250 in blocking buffer. Next, the samples were incubated with secondary antibodies, Alexa-Fluor 488 goat anti rabbit (A11008, Thermo Fisher Scientific) and Alexa-Fluor 555 goat anti-mouse (A21422, Thermo Fisher Scientific), diluted in blocking buffer (1:1000) for 1 h. Nuclear DNA was stained by DAPI (SIGMA, D9542). The DAPI stain was added directly to the secondary antibody solution (diluted 1:5000). After drying, the coverslips were mounted on glass slides using Fluoroshield (SIGMA, F6182) mounting medium.

Images were acquired using the FV1200 Olympus inverted confocal microscope. Dual color confocal images were acquired with standard settings using laser lines 488 nm and 543 nm for excitation of Alexa Fluor 488 and Alexa Fluor 568 dyes, respectively. Nuclear DAPI staining was imaged using excitation by the 405 nm laser. For each condition 10 images were randomly acquired with the 20X/ 0.75 objective and imported into CellProfiler for downstream image analysis. For each data point, 400–600 nuclei (identified by DAPI staining) were analysed for 3meC and 5meC nuclear intensity (mean integrated intensity).

The m3C and m5C values presented are based on four independent staining experiments. The Wilcoxon's rank sum hypothesis test was used to assess differences in m3C and m5C values (in R 3.1.0).

3.9 DNA Damage Induction

3.9.1 Neocarzinostatin.

Neocarzinostatin (NCS) is a radiomimetic drug that rapidly induces DNA DSBs. Neocarzinostatin itself is a chromoprotein that consists of two noncovalently bound components. These are the chromophore component (NCS-chrom), which has biological activities such as DNA cleavage, and a protein component that stabilizes NCS-chrom. After activation by thiols, the NCS-thiol adduct collapses to form a putative biradical that can remove a hydrogen from the deoxyribose of DNA, causing the DNA to be cleaved (Edo et al., 1997).

3.9.2 405nm laser

Laser micro-irradiation was performed with the 405nm laser of an Olympus FLV1200 confocal microscope. Targeted areas were pulsed at 100 nJ to induce DSBs and DSB response. Cells were fixed in 4% PFA after 10, 30 and 60 minutes to check for recruitment of proteins to the site of irradiation.

3.10 Proteasome Inhibition

Proteasome inhibition was performed through treatment with MG132 (Sigma 1211877-36-9). MG132 is a peptide aldehyde, which effectively blocks the proteolytic activity of the 26S proteasome complex, being a potent inhibitor of the chymotryptic-like activity (Lee and Goldberg, 1998).

3.11 RNA Immunoprecipitation

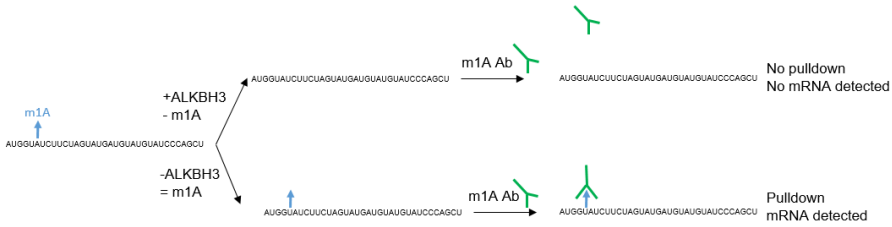


Figure 11. RNA immunoprecipitation for methylation marks. The removal of a modification from a target mRNA will cause for decreased pulldown

Cells were initially seeded to 50% confluency in a 100mm dish and allowed to attach overnight. 10uL siRNA was combined with 10µL of transfection reagent RNAiMAX, and allowed to incubate for 15 minutes. The siRNA was then placed onto the cells and left for 48 hours. Cells were lysed with 1mL RIPA lysis buffer (50 mM Tris-Cl, 1% NP-40, 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NacL) with added 100 units of RNAsin to prevent RNA degradation and centrifuged at 12000 xg to remove cellular debris. 100µL of the supernatant was removed to be used as input control. 7.5µG of antibody for either N1-methyladenosine(Cat# D345-3 from MBL International) or N6-methyladenosine (Cat# 202 003 from Synaptic systems). was diluted in 200µL PBST, then added to Dynabeads (Dynabeads Protein A, Life Technologies, 10002D) and incubated with 10 rpm rotation for 30 minutes. The tube was then placed on a magnet and gently washed in PBST. Wash was removed from the beads and 900µL of the sample supernatant was added to the antibody-bound beads. This was incubated at room temperature with 10 rpm rotation for 30 minutes. The tube was placed on the magnet and the supernatant was removed followed by 3 200µL washes with lysis wash buffer (100 mM KCl, 5 mM $MgCl_2$, 10 mM HEPES, 0.5% Tween-20, 1 mM DTT, 100 U/mL RNAsin). Samples were then resuspended in 100µL of elution buffer (lysis wash buffer,

0.1% SDS, 30µg Proteinase K) and transferred to a new tube to avoid contamination. Samples were incubated at 50°C for 30 minutes with gentle shaking. Samples were placed on the magnet and supernatant was removed and put in a new tube. One volume of phenol-chloroformisoamyl alcohol mixture was added followed by vortexing, then centrifuging for 1 minute to separate the phases. The upper phase was recovered into a new tube. RNA was precipitated by adding 12µL 3 M sodium acetate, 10µL 20mM Glycogen and 250µL 100% ethanol and placed in -20°C overnight. Next, the samples were spun down at max speed for 30 min and the supernatant was discarded. The visible pellet was further washed with 75% alcohol for 5 min. The pellet was allowed to dry at room temperature for 5 min before resuspending in 10-15µL of RNase free water. The RNA was quantified using a nanodrop (Nanodrop One, Thermo Scientific).

3.12 qPCR

For all RNA IP experiments 200ng of input and IP RNA was used for cDNA synthesis. For gene expression analysis of siRNA mediated knockdown samples or Nuclear and cytoplasmic fractions, 500ng of RNA was reverse transcribed to cDNA using the High Capacity cDNA Reverse Transcription kit (ABI Cat no: 4368814) according to manufacturers' protocol. The PCR program for cDNA synthesis was 25°C for 10 minutes, followed by 37°C for 2 hours. The reaction mixture was heat inactivated at 85°C for 5 minutes. The cDNA samples were stored at -20°C. To study expression levels of target genes real time primers were designed using either the Primer3 plus software or the online Primer depot tool (NCI, NIH). Quantitative analysis of the PCR products was performed using SYBR Green chemistry in Biorad system CFX384.

The master mix per reaction volume (5uL) was prepared as follows:

PowerUp SYBR green reagent (cat no. A25742) 2.5 µL

Forward primer (stock 10µM) 0.075 µL

Reverse primer (stock 10µM) 0.075 µL

Double distilled water 0.35 µL

This reaction mix (3 µL) was pipetted into the Biorad PCR plate wells (cat no.HSP3805) 2 µL of template cDNA was added to each well to make up the

reaction volume to 5 µL. The plate was sealed with an optical cover (1814030) and spun down briefly before placing in the machine. The cycling conditions for the biorad machine are detailed in the following table:

Table 6. qPCR conditions

	Temperature (°C)	Time (min:sec)
Step 1	50	02:00
Step 2	95	10:00
Step 3	95	00:15
Step 4	60	01:00
Step 5	Plate Read	Plate Read
Step 6	Repeat from Step 3 39X	Repeat from Step 3 39X
Step 7 (Melt Curve)	60	00:31
Step 8 (Melt Curve)	60	00:05

All reactions were set up in duplicate or triplicate for each template sample. The endogenous control gene used for normalization were the housekeeping genes Beta actin and HPRT. The CT values obtained from each reaction were used for calculation of relative difference in gene expression using the 2-ΔΔCT method for relative quantification (Pfaffl et al., 2001). Mean CT values were calculated after excluding outliers. The ΔCT value was calculated by subtracting the endogenous gene CT value from the target gene CT. The ΔΔCT was calculated by subtracting the experimental ΔCT from the calibrator ΔCT. The fold change was calculated as the 2-ΔΔCT value. The calculation for enrichment of RNA IP was done using the formula from Sigma Aldrich RIP template. The yield in terms of %Input was calculated as 100*2^-***Δ***Ct where ***Δ***Ct is (Ct[IP]-Ct[Input])

3.13 Cellular Fractionation

Cellular fractionation was performed using a RNA subcellular isolation kit (Active Motif, Cat# 25501). The kit allows for the separation of the nuclear and cytoplasmic RNA. We performed siRNA knockdown for target and then used the kit, followed by cDNA synthesis and qPCR as described in the qPCR section of the methods chapter.

3.14 RNA Scope

RNA Scope is an in-situ hybridization of specifically targeted mRNAs. The probes themselves were designed by and obtained from ACD Bio. The probes ordered were for human RNF168 mRNA, a negative control targeting bacterial mRNA and a positive human control mRNA (POLR2A, PPIB and UBC). RNA Scope works by using 20 sets of two independent probes (double Z probes). These probes have to hybridize to the target sequence in tandem to then be recognized by a secondary amplification probe. The signal is then amplified two more times and then finally a fluorescent probe that recognizes the amplified double Z probes is used. The images are then analyzed within a confocal microscope.

RNA Scope experiments were done by seeding $5x10^3$ cells to roughly 50% confluency in a 8 Chamber Polystyrene Vessel Tissue Culture Treated Glass Slide (Fisher Scientific, 08-774-26). After 48 hours, cells were fixed in 4% paraformaldehyde for 30 minutes. Following fixation cells were dehydrated in steps of 50% EtOH for 5 minutes, 70% EtOH for 5 minutes, 100% EtOH for 5 minutes and a final 100% EtOH for 10 minutes and then stored in -20 until needed. Samples were re-hydrated by applying 70% EtOH for 5 minutes, 50% EtOH for 5 minutes, then PBS for 10 minutes. The top part of the chamber slides were then removed and a hydrophobic barrier was drawn around the wells using a ImmEdge Hydrophobic Barrier PAP Pen (Vector Laboratories, H-4000) and let dry for 1 minute followed by a PBS rinse. 2-4 drops of Pretreat protease III was added to sufficiently cover the wells and incubated for 10 minutes at room temperature inside a humidity control tray followed by an additional PBS rinse. Target probes for control, RNF168 and negative control were applied and left at 40°C for 2 hours. This was followed by two, 5 minute washes with the 1X RNA Scope wash buffer. Primary, secondary, tertiary and fluorescent probes were placed on for 30, 15, 30 and 15 minutes respectively, each with two 5 minute washes in the wash buffer. Cells were then treated with DAPI nuclear stain (part of RNA Scope kit) for 60 seconds followed

immediately by the application of fluoromount and cover slip.

Confocal microscopy in the Olympus FLV1200 under the 60X oil immersion objective. Images were acquired in an unbiased manner with no knowledge of RNA Scope foci but rather selected based on DAPI stain. Images were taken in 3D-stacks of 10-12 images in order to capture the entire volume of the cell with all mRNA foci. Stack images were compiled into a max intensity z-stack single image using ImageJ and then analyzed in Cell Profiler 3.0 in a custom pipeline. The pipeline distinguished between foci located within the nucleus and cytoplasm of each cell and counted the total foci in each condition. Approximately 200 cells were counted for each treatment.

3.15 FLAG-Pulldowns

The ALKBH3-FLAG construct was prepared by cloning ALKBH3 cDNA into a pCR®-Blunt II-TOPO plasmid (NEB). The primers were designed using NEBuilder and were as follows: FWD gcttggtaccgagctATGGAGG-AAAAAAGACGGCGAGC, REV with FLAG, gatgcatgctcgagcTCACTTGTCGTCATCGTCTTTGTAGTCCCAGGGTGCCCC TCG. The Vector backbone was digested with SacI and NotI digestion enzymes (NEB, R3156S and R3189S). The ligation was done according to the HiFi DNA Assembly protocol as provided by NEB.

The FLAG-immunoprecipitation was performed as follows. HEK-293T cells were seeded on a 10cm plate. At 50-60% confluency, cells were transfected with 4µG of FLAG-ALKBH3 using 6.25µL of Lipofectamine LTX transfection reagent (Thermo Fisher, 15338100). 48 hours following transfection, cells were lysed using 1mL FLAG-IP lysis buffer (50 mM Tris HCl pH 7.4, 150 mM NaCl, 1mM EDTA and 1% TRITON X-100) with added protease inhibitor cocktail 1:100. Plates were incubated with lysis buffer for 30 minutes on a shaker in the cold room. Cells were scraped and centrifuged for 10 min at max speed. Supernatant was transferred to a chilled 1.5 mL tube, with 100 µL being taken from that to be used as an input control. The remaining lysate was placed on 40µL of previously washed FLAG-beads and rotated at 4°C overnight. The next day samples were centrifuged at 7000g for 1 min. Supernatant was collected to be used as a Wash sample. The remaining beads were washed 3 times in ice cold TBS (0.5 M Tris HCl pH 7.4, 1.5 M NaCl), with 1 minute centrifuging between each spin. For elution, 9 µL of 3X FLAG peptide was combined with 51 µL TBS (per sample). Elution buffer was added to beads and put on a gentle rocker at 4°C for 30 minutes. Cells were centrifuged at

7000g for 1 minute at 4°C and supernatant was gathered as a Immunoprecipitation sample. All samples were mixed with 2x sample buffer (Thermo Scientific, Cat# NP0007) and 1X DTT reducing agent, then boiled at 95°C for 10 minutes and frozen down at -20°C.

3.16 DNA Repair Cell Line models

The U2OS DR-GFP line use to measure SSA was designed in Maria Jasins lab (Pierce et al., 1999), and donated to us by the Lukas lab at the Center for Protein Research at the University of Copenhagen.

To measure the repair of an I--SceI--generated DSB at the DR-GFP locus, U2OS DR-GFP cells (2 million) were first transfected with the indicated siRNAs, and 24h later cells were transfected with 25 µg of the I--SceI expression vector pCBASce or an empty vector with Lipofectamine LTX (Thermo Fisher Scientific). Cells were harvested 48h post I-SceI transfection and analyzed for GFP positive cells and PCR-based SSA. The GC efficiency was determined by quantifying GFP-positive cells (product of successful GC; normalized to the transfection efficiency) via flow cytometry with a FACS SH800 (Sony Biotechnology).

To investigate SSA, we used a PCR method developed by the Lukas lab (Ochs et al., 2016). 48 h after I-SceI transfections, genomic DNA was prepared using a Qiagen genomic DNA isolation kit. Total DNA was quantified, and 100 ng of each sample was used as a PCR template for two pairs of PCR primers, either f (forward) and SA-r1 (reverse 1), or f and r2 (reverse 2). The primer sequences were as follows:

Table 7. SSA cell line PCR primers

Forward f	TTTGGCAAAGAATTCAGATCC
Reverse r1	CAAATGTGGTATGGCTGATTATG
Reverse r2	ATGACCATGATTACGCCAAG

Conditions for PCR reactions are:

Table 8. SSA Cell line PCR conditions

	Temperature (°C)	Time (min:sec)
Step 1	98	00:45
25 Cycles Of:		
Step 2	95	00:10
Step 3	50	00:30
Step 4	72	00:60
Final extension		
Step 5	72	02:00

Amplification was carried out for 25 cycles and was determined to be in the linear range. After the products were run on a 1% gel, the gel was stained with ethidium bromide, and SSA product bands were quantified with ImageJ.

3.17 Software

Displays of chemical structures were drawn using Chemdraw which is licensed by the University of Iceland. ImageJ, CellProfiler 3.0, R and R Studio are free software available online. The yet to be named (currently Stefan Decoder) application used to identify consensus sites within FASTA sequences was programmed in Java with the assistance of Robert Lowell.

4 Results

4.1 Differential Promoter-Methylation of ALKBH3 gene

4.1.1 ALKBH3 in TCGA

The initial identification of ALKBH3 as a protein of interest arose from my personal interest in combining my background of DNA repair with a curiosity to learn about epigenetics. The project began by cross-referencing a list of DNA repair genes (n = 178) (Kauffman et al., 2008) with a list of genes undergoing differential DNA promoter methylation we identified within The Cancer Genome Atlas (n = 456). The crossover between these lists gave 8 genes. These proteins were ALKBH3, BRCA1, DMC1, EYA2, NABP1, PARP3, RAD51C and TP53I3. The impact of promoter methylation for BRCA1 had been well studied at the time. ALKBH3 had a fairly well defined functional role in DNA alkylation damage repair, but no research had been done on it in relation to epigenetics so we decided to focus on that protein.

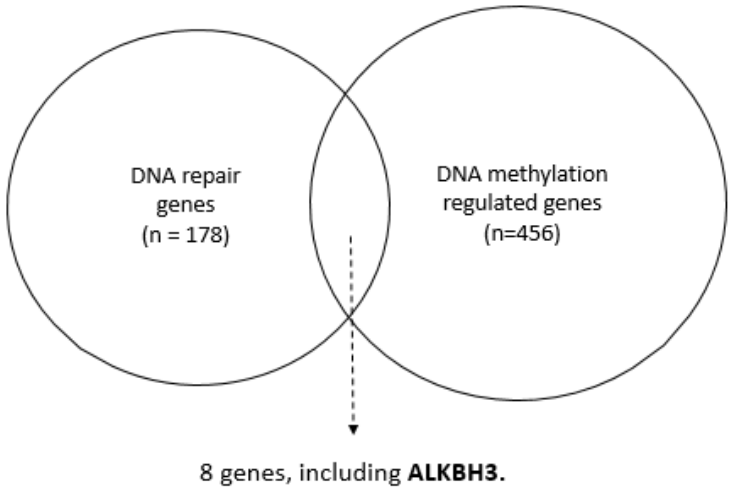


Figure 12. Identifying ALKBH3 as a protein of interest. Chosen out of a cross-referenced list of DNA repair genes and genes regulated by DNA methylation.

Additionally what was interesting was that we found in The Cancer Genome Atlas that ALKBH3 was undergoing differential promoter methylation in ~20% of breast cancer samples available in the database (Figure 13).

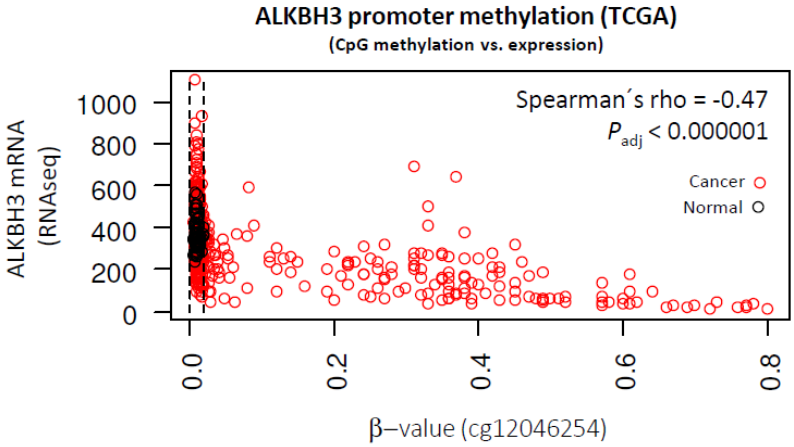


Figure 13. ALKBH3 promoter methylation of cg12046254 compared between normal and cancer tissue. The plot shows the correlation between mRNA expression (y-axis) and the level of promoter methylation (x-axis). The dashed lines (in black) represent the lower and upper 99% confidence limits for the normal breast tissue samples – reflecting the "normal range" of 5-methylcytosine levels for this particular CpG (cg12046254)

We then specifically identified 9 CpGs within the ALKBH3 promoter and checked them for differential methylation. The presence of methylation on that particular CpG would give a positive result, whereas no methylation would be negative. Samples can contain a broad range of methylation levels due to the heterogeneity of the tumors the DNA was isolated from, as well as other factors. 8 of the 9 of the examined CpGs were shown to have differential levels of methylation of the ALKBH3 promoter between cancer and normal tissue (Figure 14).

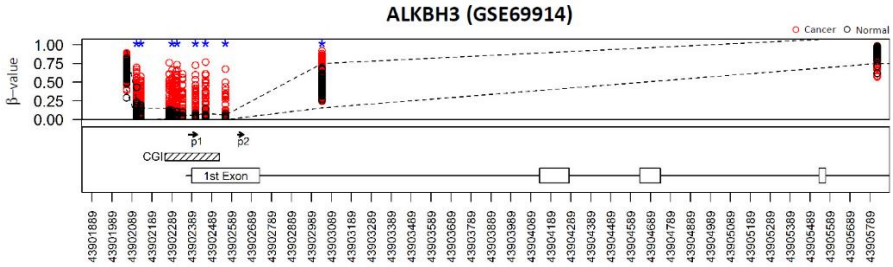


Figure 14. Differential promoter methylation of ALKBH3 gene. Red data points indicate methylation levels in a cancer tissue, black indicate normal tissue. Differentially methylated CpGs between breast cancers and normal breast tissues are indicated by blue asterisk. Major Transcription Start Site (TSS) indicated by p1. The dashed lines (in black)represent the upper and lower 99% confidence intervals for the distribution of β-

values in normal breast tissue samples – thus displaying the “normal range” for each of the CpGs analysed.

ALKBH2, which has been shown to have the same alkylation damage repair targets as ALKBH3, showed no differential methylation levels between tumor and normal tissue (Figure 15)(Aas et al., 2003; Duncan et al., 2002; Dango et al., 2011).

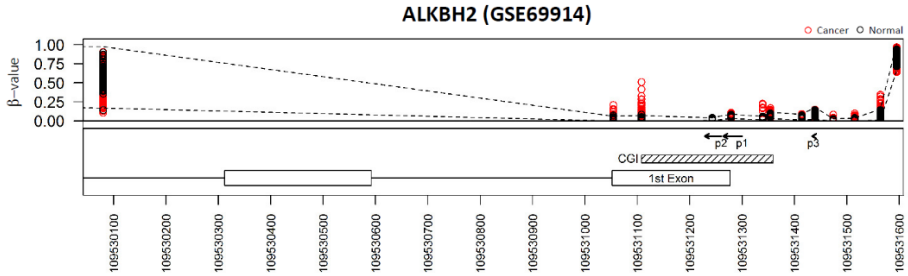


Figure 15. CpG methylation over the promoter of ALKBH2. No significant differential expression of the gene was seen in with the CpG sites identified. Major Transcription Start Site(TSS) indicated by p1. The dashed lines (in black)represent the upper and lower 99% confidence intervals for the distribution of β-values in normal breast tissue samples – thus displaying the “normal range” for each of the CpGs analysed.

The TCGA data suggested that ALKBH3 was undergoing significant promoter methylation that caused differential expression between normal and tumor tissue of the same patients. We looked at the cBioportal (http://www.cbioportal.org/, Cerami et al., 2012), a database of published cancer genomic data, to study how the expression of ALKBH3 would be impacted by promoter methylation. We found that samples with 28.4% promoter methylation showed a 2-fold decrease in the expression level of ALKBH3 whereas samples with 52.5% promoter methylation exhibited a 4-fold decrease in ALKBH3 expression levels.

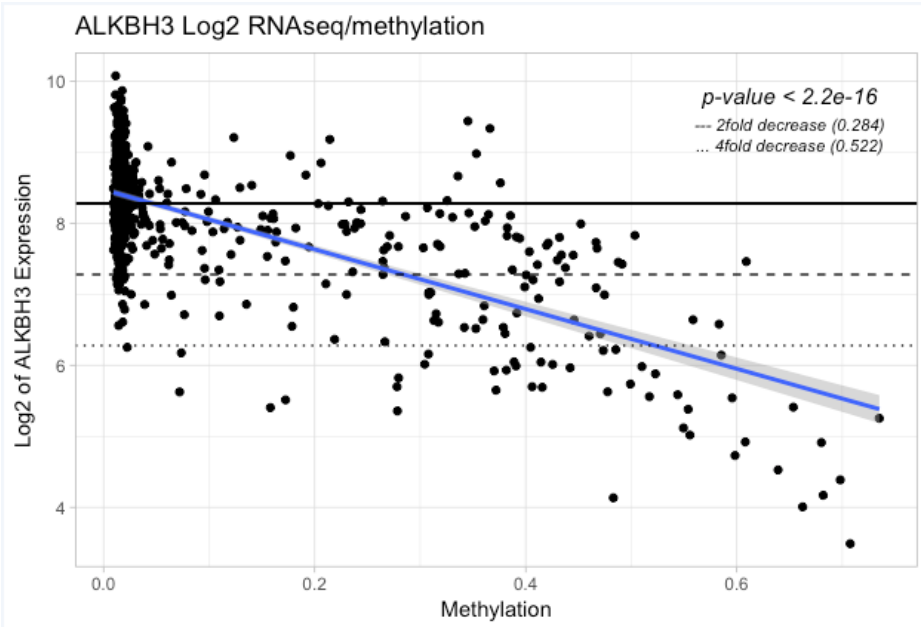


Figure 16. ALKBH3 expression according to promoter methylation levels on a Log2 scale. Higher levels of methylation cause decreased expression of ALKBH3 mRNA. 28.4% promoter methylation correlated with a 2-fold decrease in the expression level of ALKBH3 whereas samples with 52.5% promoter methylation correlated with a 4-fold decrease in ALKBH3 expression levels.

The “level” of promoter methylation is an interesting topic to be reviewed in the discussion, but the trend appears to be that at higher levels of methylation there was lower expression of the ALKBH3 mRNA.

The next step of the project after having identified ALKBH3 promoter methylation within the TCGA was to look into patient samples available to us in Iceland, the cohort of which is described in the methods chapter, and see if we could replicate the phenomenon while also being able to compare the data against available patient data.

4.1.2 ALKBH3 methylated tumors vs normal tissue in Icelandic cohort

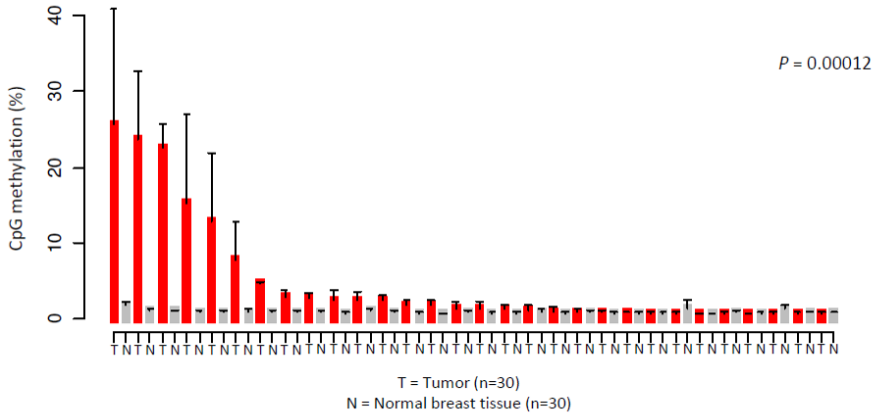


Figure 17. Differential methylation between normal and tumor tissue in paired Icelandic samples. Pyrosequencing for the available paired Icelandic samples confirmed differential promoter methylation between tumor and normal tissue.

We took available tumor samples that had paired normal tissue and compared the promoter methylation of ALKBH3 between the matched samples. Patients who had tumors containing high levels of promoter methylation did not possess significant levels of promoter methylation within their normal tissue (Figure 17). Importantly, this data indicates ALKBH3 promoter methylation is a cancer specific event.

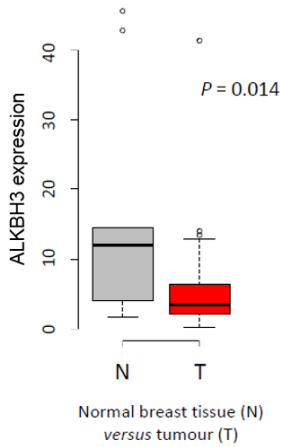


Figure 18. qPCR to measure ALKBH3 mRNA expression in tumors and normal tissue of matched samples. Expression in tumor samples is statistically significantly decreased.

The same paired samples underwent qPCR to analyze the mRNA expression between the tumor and normal tissue. This showed a statistically significant decrease in ALKBH3 mRNA expression in the tumor tissue compared to the normal tissue (Figure 18).

4.1.3 ALKBH3 methylation relates to poor disease outcome

When analyzing survival of Icelandic breast cancer patients, we had to check several different thresholding points of the levels of promoter methylation in order to see a statistically significant impact on survival. Levels of promoter methylation can vary for multiple reasons. Heterogeneity within tumors can account for this variance between patients (Assenov et al., 2018). Additionally, samples were collected during tumor removal surgery and then subjected to DNA extraction. This provides no way of knowing what amount of the sample is actually tumor tissue and what may potentially be normal tissue. 27% (72 of 265) of patients in our samples exhibited some level of ALKBH3 promoter methylation in tumors. In order to see statistically significant impact on patient survival, the tumor sample had to possess ALKBH3 promoter methylation at least at 20%. This left 5% (13 of 265) of patients who had more than 20% methylation present in their tumor samples (Figure 19).

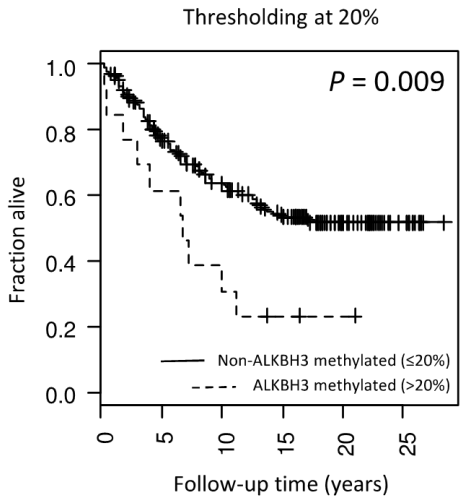


Figure 19. Cox's proportional hazards regression model for breast cancer specific patient survival with respect to level of ALKBH3 promoter methylation. When promoter methylation occurred in greater than 20% of the sample, statistically impacted survival was seen.

When the thresholding of promoter methylation was dropped to lower levels than 20%, we still saw a trend towards decreased survival, yet it was not statistically significant (log-rank hypothesis test) (Figure 20).

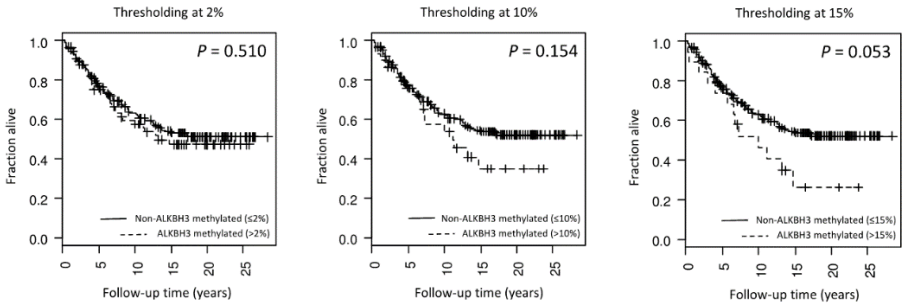


Figure 20. Impact on survival at different levels of promoter methylation. Setting the threshold cutoff for what amount of the sample had promoter methylation is important for getting statistically significant impact on survival. Lower threshold values of 2%, 10% and 15% do not produce a P-value lower than 0.05 (log-rank hypothesis test)

4.1.4 ALKBH3 methylated cell lines

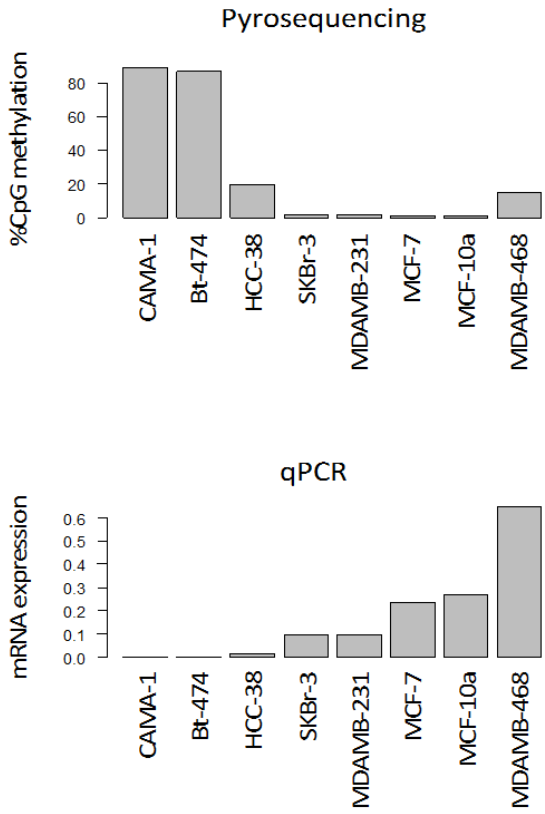


Figure 21. Pyrosequencing and qPCR of cell lines. Pyrosequencing confirmed the existence of multiple cell lines with hyper-methylated ALKBH3 promoters that correlated with loss of mRNA expression.

23 cell lines, 18 breast cancer and 5 ovarian cancer lines, were grown to check for the existence of promoter methylation within a cell line model. This figure (Figure 21) shows a selection of 8 of the breast cell lines, chosen to demonstrate a range of promoter methylation and ALKBH3 mRNA expression. We discovered 2 lines, CAMA-1 and BT-474, to contain very high levels of promoter methylation. These high levels of methylation correlated with a complete lack of ALKBH3 mRNA expression when checked with qPCR. No significant levels of promoter methylation were detected within the ovarian cancer lines we tested (Figure 22).

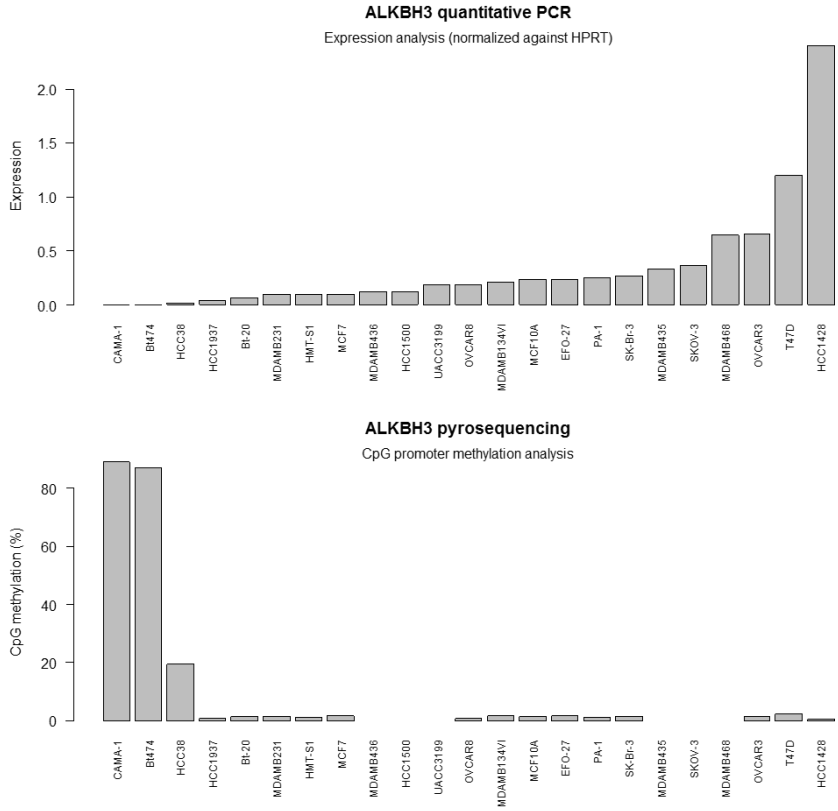


Figure 22. Further pyrosequencing and qPCR of cell lines. Additional cell lies, including 5 ovarian lines were pyrosequenced and had expression checked by qPCR in order to look for examples of promoter methylation corresponding to decreased expression.

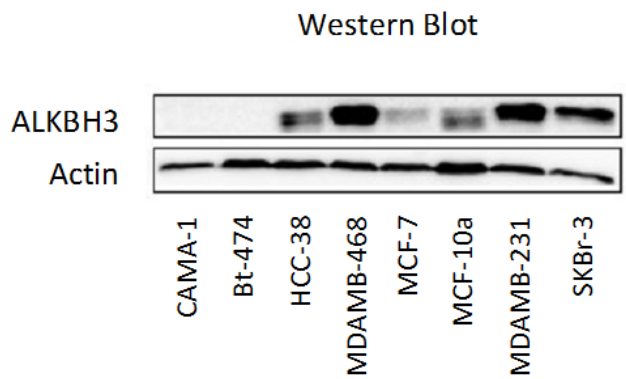


Figure 23. Western blot for cell lines selected to use as examples of varying impact of promoter methylation on expression. Cell lines CAMA-1 and BT-474, which contained high ALKBH3 promoter methylation exhibit no expression of ALKBH3 protein.

Proteins were extracted from selected cell lines and detected by western blot analysis. The cell lines CAMA-1 and BT474 which showed high methylation and no mRNA expression did not express the ALKBH3 protein (Figure 23).

These results confirmed within a cell line model that promoter methylation for the ALKBH3 gene correspond with decreased protein expression. The impact of ALKBH3 being absent did not prove to be lethal to cells.

4.1.5 ALKBH3 absence increases levels of m3C in cells

In order to confirm a functional impact of ALKBH3 promoter methylation, we developed a assay to measure the levels of endogenous m3C in cell lines. Given the reported role of ALKBH3 in m3C repair (Dango et al., 2011), the hypermethylated cell line, CAMA-1, was expected to have higher levels of m3C damage.

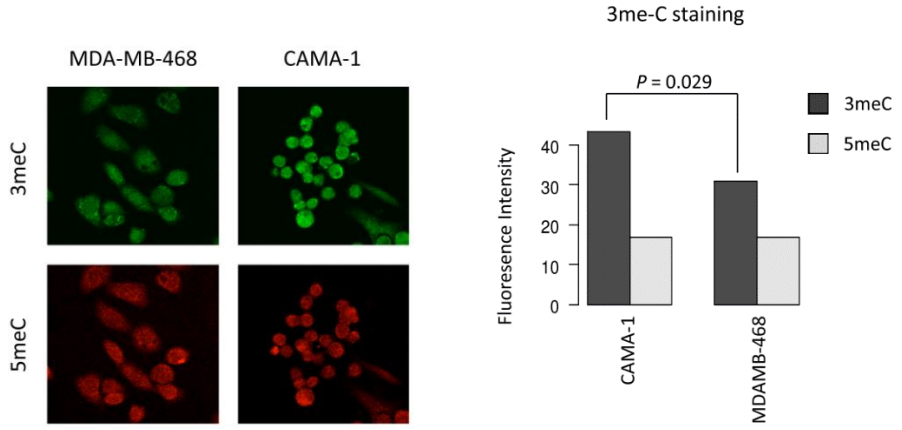


Figure 24. Accumulation of endogenous m3C damage in hyper-methylated and unmethylated cell lines. The confocal images to the left show signal intensity of m3C and m5C for MDAMB-468 (unmethylated ALKBH3 promoter) and CAMA-1 (hypermethylated ALKBH3 promoter). The intensity of the images is quantified to the right, showing that the unmethylated cell line has statistically significant (Wilcoxon's rank sum test) decreased intensity of endogenous m3C while non-significant differences in m5C. 400–600 nuclei (identified by DAPI staining) were analysed for 3meC and 5meC nuclear intensity (mean integrated intensity) under idential conditions.

This was indeed the case. CAMA-1 displayed statistically significant higher levels of m3C when compared to MDAMB-468, while the two cell lines maintained similar levels of m5C (Figure 24). Combining this data with the downregulation of expression we confirmed in the tumors and other methylated cell lines, we can see some functional consequence of ALKBH3 silencing. As mentioned in the introduction, increased m3C can cause inflammation, nucleotide misincorporation, replicative stress, and promote the use of translesion polymerases (Soll et al., 2012; Starcevic et al., 2004; Krokan et al., 2013).

4.2 The Functional Impact of Losing ALKBH3

4.2.1 Reduction of 53BP1 foci

With the knowledge that ALKBH3 is down-regulated in tumor tissue and not normal tissue from the same individual, with one consequence being increased m3C levels in promoter-methylated cells, we aimed to discover any further functional impact of this event.

In the available literature, there was a tandem mass spectrometry analysis after ALKBH3 pulldown (Dango et al., 2011). Among the proteins interacting with ALKBH3 were RIF1 and CHD4, proteins both involved in the regulation of DSB repair through the ubiquitin signalling pathway (Chapman et al., 2011; Larsen et al., 2010; Luijsterburg et al., 2012). As detailed in the introduction, RIF1 is a binding partner of 53BP1 (Chapman et al., 2011). The absence of either 53BP1 or RIF1 causes dysfunction of the complex. Formation of 53BP1 foci indicates that the DSB induced signalling pathway is active. Based on the interaction with RIF1 and CHD4 in the LC-MS/MS, we were curious if the knockdown of ALKBH3 would have any impact on DNA DSB signaling pathway. Next, we were asking whether 53BP1, a key component of the DSB ubiquitin signalling pathway that is downstream of CHD4 and a functional partner of RIF1, would be affected by ALKBH3 depletion. We knocked down ALKBH3 and used Neocarzinostatin to induce DSBs and compared against an untreated control.

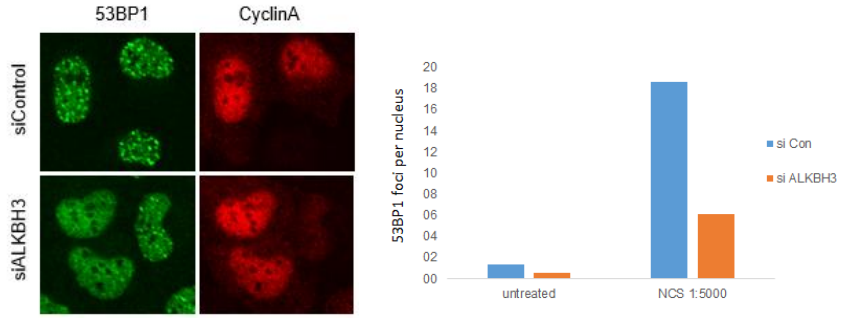


Figure 25. Reduction of 53BP1 foci following siALKBH3. Neocarzinostation (NCS) was used in a concentration of 50 ng/mL for 1 hour to induce DNA DSBs in U2OS cell line. 53BP1 foci formation was dramatically reduced in ALKBH3 knockdown sample. Indication that ALKBH3 plays a role in the regulation of the DSB repair pathway.

Following treatment with neocarzinostatin, there was a dramatic reduction of 53BP1 foci formation in the ALKBH3 knockdown sample (Figure 25). This suggested that ALKBH3 is playing some role in the effective repair of DSBs. We additionally stained for CyclinA, a cell cycle marker, as there was some question as to whether this effect was cell cycle dependent, but this was later determined not to be the case as cells outside of S phase exhibited the same extent of foci formation. The next step was to determine where in the pathway ALKBH3 knockdown was executing its function.

4.2.2 Identifying the cause for downregulation

The next logical step was to check upstream of 53BP1 to see if any of the proteins crucial for its recruitment were affected by ALKBH3 depletion. RNF168 is responsible for placing additional ubiquitin on to the initial ubiquitin placed by RNF8 at sites of DSBs. The absence of the ubiquitin chains created by RNF168 will inhibit 53BP1 recruitment (Doil et al., 2009).

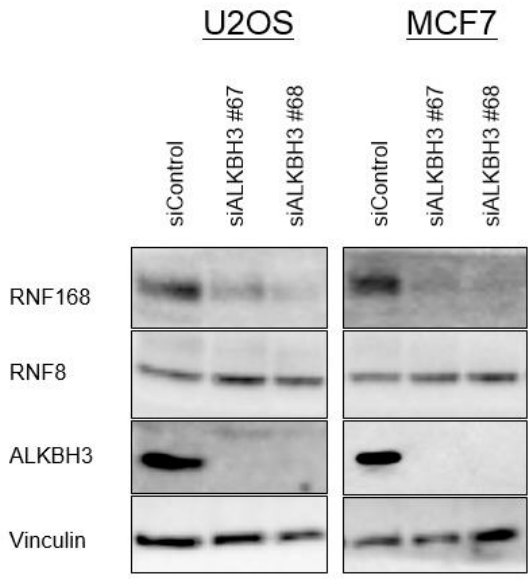


Figure 26. ALKBH3 knockdown causes decreased levels of RNF168. Left hand panel: Knockdown of ALKBH3 decreases RNF168 but not RNF8 protein levels in U2OS. Two different siRNAs for ALKBH3 to ensure it is not an off-target effect. Right hand panel: Same experiment performed in the MCF7 cell line to rule out specific cell line specificity or artifacts.

The knockdown of ALKBH3 resulted in the reduction of RNF168 protein levels. It did however not reduce RNF8, the protein upstream of RNF168 (Figure 26). I therefore focused the research on to how ALKBH3 is regulating RNF168. As mentioned in the introduction, there are many forms of regulation of RNF168. The function of this protein is particularly important in the recruitment of 53BP1 and the subsequent DSB repair pathway determination.

We confirmed by immunofluorescence that ALKBH3 knockdown showed a similar dramatic reduction of 53BP1 foci phenotype similar to RNF168 depletion (Figures 27 & 28).

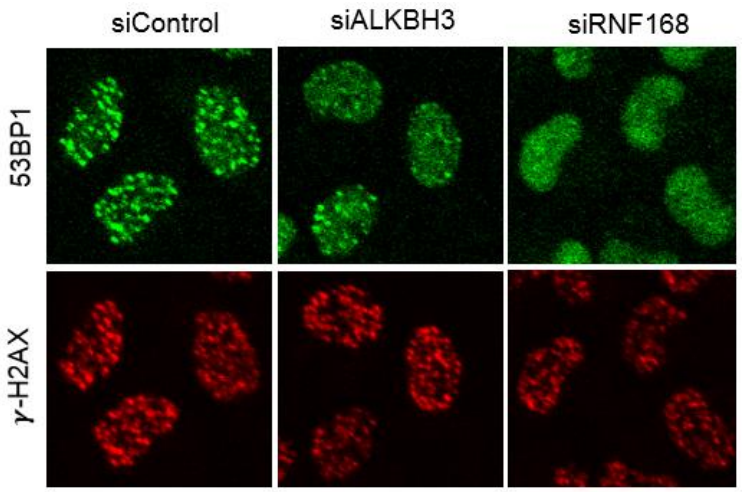


Figure 27. Confirmation of RNF168-like phenotype on 53BP1 foci recruitment. Immunofluorescence images of 53BP1 foci following treatment with NCS for 1 hour (50 ng/mL) in U2OS cells. ALKBH3 knockdown shows similar phenotype as RNF168 knockdown. γ-H2AX is used as a marker of DSBs.

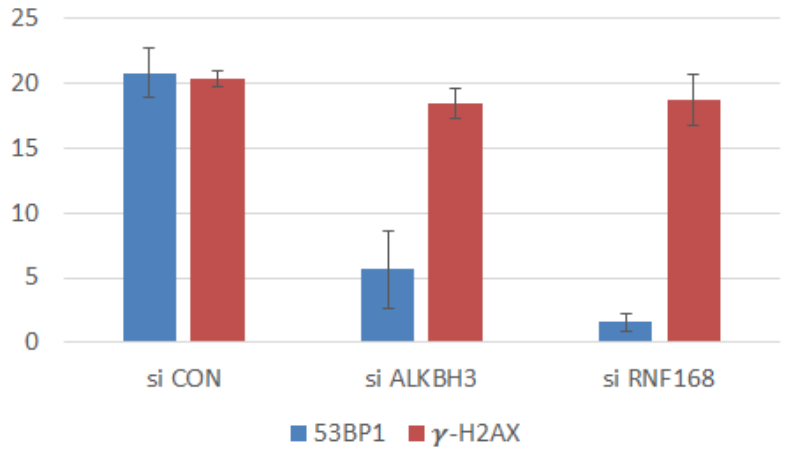


Figure 28. Quantification of immunofluorescence staining for 53BP1 and γ-H2AX foci in U2OS cells. Reduction of 53BP1 foci following ALKBH3 knockdown is similar to that of RNF168 knockdown. The experiment was performed in triplicate and error bars represent standard error of mean.

4.2.3 Separation of function from ASCC3 and ALKBH2

In order to determine if this effect on RNF168 was a result of a known ALKBH3 function or a novel mechanism, we looked at ASCC3 to see if its knockdown would replicate the phenotype. As a reminder, ASCC3 interacts with ALKBH3 and is the helicase responsible for facilitating ALKBH3 access to alkylation damage within the DNA (Dango et al., 2011). We argued, that if ASCC3 would show the same phenotype it may indicate the regulation of RNF168 is somehow linked to the role ALKBH3 has in DNA alkylation repair. ALKBH2 additionally has been shown to act on the same targets within the DNA, so therefore that additionally would provide some insight if this mechanism of regulation is linked with the previously known role of ALKBH3 in alkylation damage repair.

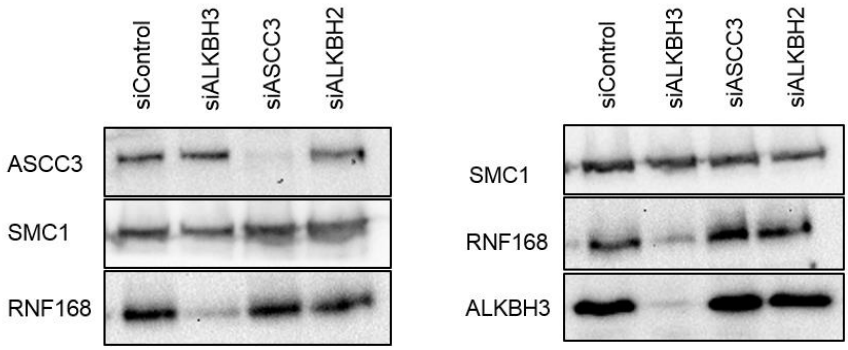


Figure 29. Separation of function from known alkylation repair partners. ASCC3 and ALKBH2 show no impact on the levels of RNF168 protein via western blot. Additionally the knockdown of ASCC3 and ALKBH2 have no impact on the levels of ALKBH3. SMC1 is used as a loading control.

This was however not the case. Knockdown of ASCC3 showed no impact on the levels of RNF168 protein, indicating that this was separate from the role ALKBH3 plays in alkylation damage removal that is dependent on ASCC3 (Figure 29). The same was true for ALKBH2 as there was no decrease in RNF168 protein. These results give additional credence to the idea that the regulation of RNF168 is separate to the role ALKBH3 plays in DNA alkylation repair.

4.2.4 Impact on mRNA expression

An important next step was to check for any impact of ALKBH3 on the

expression of RNF168 mRNA. ALKBH3 had been previously been shown to bind to transcriptional start sites of some highly active promoters, but with no impact on expression of the genes studied (Liefke et al., 2015). Relative quantitative PCR was used to determine the expression levels of RNF168. To ensure that the observed phenotypes were not a consequence of potential siRNA off-target effects, three different siRNAs were used.

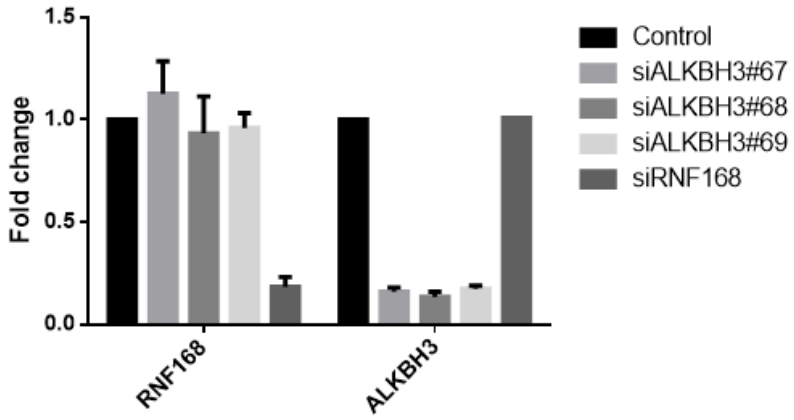


Figure 30. ALKBH3 knockdown has no impact on RNF168 mRNA transcription. qPCR results indicate that ALKBH3 has no effect on RNF168 at a transcriptional level. No significant reduction was seen in any of the three separate ALKBH3 siRNAs checked. Error bars represent standard error of mean.

The knockdown of ALKBH3 showed no impact on the total levels of RNF168 mRNA. Conversely, the knockdown of RNF168 had no impact on the total levels of ALKBH3 mRNA (Figure 30). This data then implies that any regulation ALKBH3 is performing on RNF168 is happening post-transcriptionally.

Additionally, microarray data available from the Liefke et al. 2015 showed no difference in RNF168 mRNA levels following knockdown of ALKBH3 in PC3 cells, providing additional support that ALKBH3 does not affect total levels of RNF168 mRNA (Liefke et al., 2015).

4.2.5 RNF168 Protein Turnover

Following confirmation that the regulation of RNF168 by ALKBH3 is happening post-transcriptionally, the focus of the work shifted to regulation at the protein

level. There is precedent for the protein stability of RNF168 to be regulated by other proteins, for example USP7 (Zhu et al., 2015). Therefore, the 26S proteasome inhibitor MG132 was used to see if protein turnover was affected by ALKBH3 knockdown. By inhibiting the 26S proteasome, protein degradation should be halted. Hypothetically, if ALKBH3 knockdown were to have an effect on turnover, thereby promoting RNF168 protein degradation, treatment with MG132 would see the protein levels of RNF168 return to those seen in the control siRNA of the samples not treated with MG132. If the regulation of RNF168 was unassociated with protein stability and degradation, there would be no return of the protein following treatment with MG132.

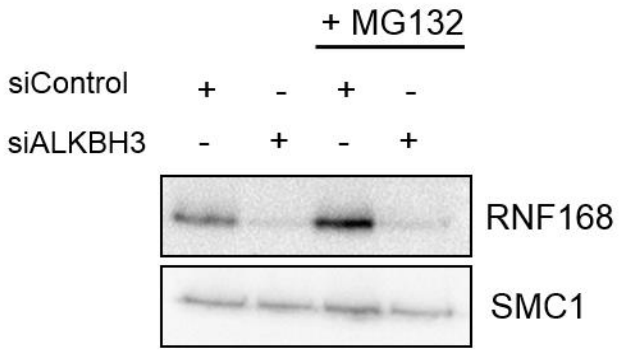


Figure 31. ALKBH3 knockdown does not affect protein turnover. We treated cells with 26S inhibitor MG132 to check for impact on protein turnover. Knockdown of ALKBH3 reduced RNF168 levels in all cases. No return of RNF168 protein was seen in ALKBH3 knockdown sample when treated with MG132, indicating no effect on protein turnover. SMC1 is used as a loading control.

The experiment showed that following treatment with MG132, RNF168 protein levels were not restored in the ALKBH3 knockdown sample (Figure 31). This indicated that the regulation was not based on protein turnover, but did not rule out other forms of protein-protein interaction.

4.2.6 Flag pulldown and protein-protein interaction

In order to check for any potential interaction between RNF168 protein and ALKBH3, a FLAG-tagged version of ALKBH3 was produced in an attempt to pulldown the protein and check for interaction. This however proved to be a

problematic process. ASCC3 is the strongest known interactor of ALKBH3 as shown by Dango et al. (Dango et al., 2011). Dango (2011) was able to show FLAG-ALKBH3 interacting with ASCC3, however this was only done by creating a cell line stably expressing the FLAG-ALKBH3 plasmid.

Additionally this paper showed a silver stain for FLAG-ALKBH3 following separation by a glycerol gradient. The silver stain showed a very small fraction of the ALKBH3 was co-precipitating with ASCC3. While the work of Dango (2011) is certainly robust, this small amount of co-immunoprecipitation could suggest that ALKBH3 has roles outside of the described role in alkylation damage repair involving ASCC3. BioGRID, an interaction database (https://thebiogrid.org/), also does not show any known or inferred interactions with RNF168.

The FLAG-ALKBH3 construct was successfully made and confirmed to be wild-type and full-length by sequencing and western blot analysis, however we were unsuccessful in getting ASCC3 to co-precipitate with with FLAG-ALKBH3 to provide a positive control for the western blots. We were however able to obtain a weak interaction with RIF1 which was also identified as an interactor of ALKBH3 in the Dango LC-MS/MS (Figure 32) (Dango et al., 2011).

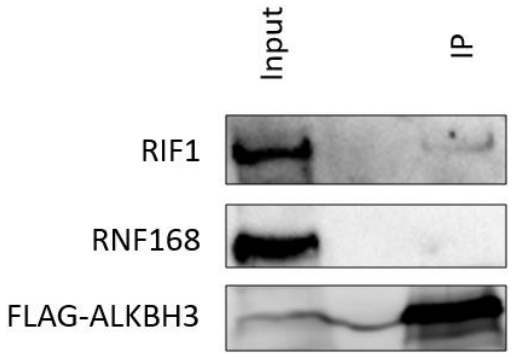


Figure 32. Immunoprecipitation of FLAG-tagged ALKBH3. Faint interaction with RIF1 is seen indicating a successful pulldown. RNF168 shows no interaction with FLAG-ALKBH3.

Despite the lack of ASCC3 interaction as a positive control, we are fairly confident ALKBH3 and RNF168 do not interact at a protein level. RNF168 showed no presence in the immunoprecipitation sample, indicating there is

more than likely no interaction. The group is currently working on a stable U2OS cell line with incorporated and FLAG-ALKBH3 in order to facilitate a more in-depth analysis of this open question.

4.2.7 405nm laser damage and interactions at site of DNA damage

We then examined if ALKBH3 was interacting at the sites of damage themselves. In order to test for this interaction, DNA damage was induced using a confocal microscope by utilizing the high intensity short bursts of a 405 nm laser (Walter et al., 2003). By using laser microirradiation, a specific area can be targeted for DNA damage induction, allowing unambiguous detection of the protein to DNA lesions. The DNA damage is detected as a γ-H2AX-decoreated stripe in the nucleus and if ALKBH3 were to be recruited to the sites of damage it would appear as a stripe as well. In one case we used a GFP-53BP1 cell line and stained for ALKBH3 5, 15, 30, 60 and 240 minutes after laser microirradiation. Another test was staining for γ-H2AX and ALKBH3 using the same time points.

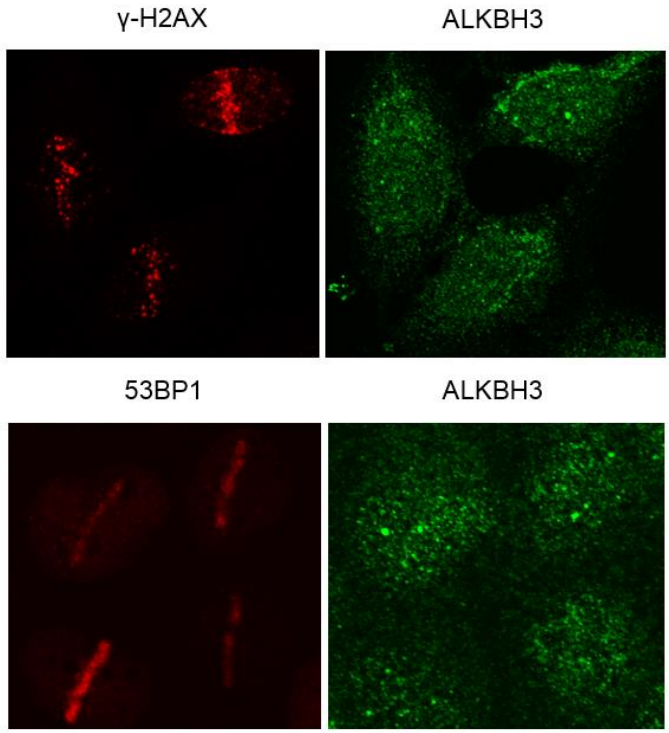


Figure 33. Micro-irradiation of U2OS cells and GFP-53BP1 U2OS cells using a 405nm laser. No co-localization of ALKBH3 was seen after induction of DNA damage to γ-H2AX or 53BP1. Displayed here is 15 minutes after induction of damage.

In all cases, no colocalization of ALKBH3 to the sites of DSBs was seen, despite the clear presence of γ-H2AX and 53BP1 indicating active DSB recognition and repair. This indicated that the regulation of RNF168 was likely not happening at the sites of DNA damage.

4.2.8 RNA methylation as a form of regulation

At this point in the project we were rather perplexed trying to figure out how exactly this regulation was occuring. The classical methods of regulation had all provided negative results, pointing to a potentially more complex mechanism. Around the time we had finished eliminating all the classical methods, a paper in Nature Chemical Biology from Li et al. was published. This paper detailed a type of post-transcriptional modification, N1-methyladenosine (m1A), as being prevalent in human mRNA (Li et al., 2016). Among the mRNAs Li et al. detailed as being modified was RNF168. Furthermore they described ALKBH3 as the RNA demethylase responsible for reversing m1A in RNA. ALKBH3 had previously been described as a RNA demethylase, but it was previously assumed this was for repair purposes and not important for post-transcriptional modification regulation. With few other promising leads, we decided to look into RNA methylation as a form of regulation for RNF168.

4.2.9 RNA Immunoprecipitation

In order to check if RNF168 is modified by N1-methyladenosine, a RNA-immunoprecipitation (RIP or RNA-IP) was carried out. Typically in the past, RIPs have been performed two ways. One method was by crosslinking mRNA binding proteins with mRNA, pulling down the target protein, reversing crosslinks and checking the mRNA bound to the protein by qPCR. The other method was introduced in 2012 which involved fractionating the mRNA and performing the pulldown with a modified nucleoside antibody, then performing sequencing and aligning the sequences to find targets. However, the target we were interested in was a modification on mRNA, not a protein, so the crosslinking method would not work. The sequencing method couldn't be used because we didn't know where exactly the modification was on mRNA sequence and fractionating may cause difficulties identifying the spot due to sensitivity. Antibodies for this modified nucleotide existed, but a method of pulling down a full length modified mRNA had not been established. This

therefore required optimization to insure we were getting total mRNA as well as intact mRNA. The concept that the modified mRNA being checked was undergoing a removal of methylation complicated the analysis, but further complicating the analysis was that several of the typical control genes used for qPCR, such as Beta-actin and HPRT, were also listed as being modified by m1A (Li et al., 2016). To what extent, if any, that ALKBH3 acted on these genes was unknown.

What we expected to see, if ALKBH3 was acting on RNF168, was an increase in RNF168 mRNA pulled down following ALKBH3 knockdown. The knockdown of ALKBH3 would cause decreased demethylation activity, and therefore increased presence of m1A on targeted transcripts. After optimizing the protocol, we successfully got a pulldown of m1A modified RNF168 mRNA.

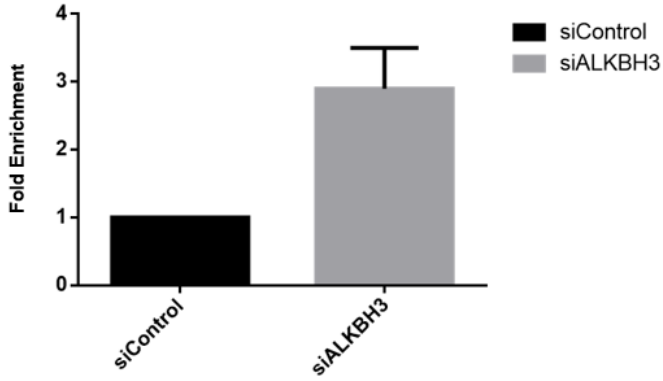


Figure 34. qPCR of RNF168 mRNA pulled down by m1A RNA-IP. RNF168 showed increased levels following ALKBH3 knockdown. Concentrations were normalized against Beta-actin. The error bar represents standard error of mean.

The results showed that indeed following ALKBH3 knockdown, RNF168 mRNA modified with m1A levels were increased. We measured a 2.8 fold increase of RNF168 mRNA following ALKBH3 depletion. This indicated that there was perhaps regulation happening to RNF168 through the RNA demethylase activity of ALKBH3. The concept of protein regulation through these RNA modifications was quite new at the time, with very few examples of specific proteins being regulated. The regulation of RNF168 through RNA modifications itself was completely novel.

4.2.10 FTO phenocopies ALKBH3

With confirmation that RNF168 is modified by m1A, and this modification is regulated by ALKBH3, we went back into the literature on demethylases. Interestingly, two of the 9 AlkB members, ALKBH5 and FTO, had been implicated in RNA demethylase activity (Jia et al., 2011; Zheng et al., 2013). We performed a knockdown of ALKBH5 and FTO to see if there was any additional action by RNA demethylases on RNF168 protein levels.

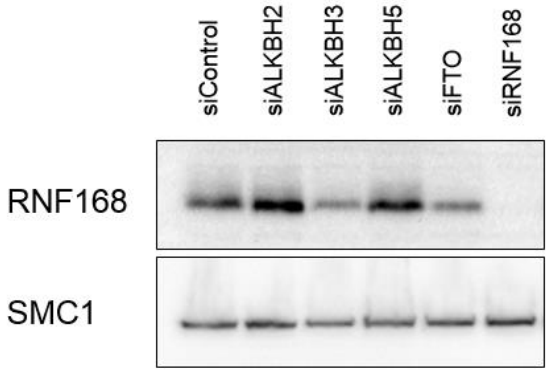


Figure 35. siALKBH3 and siFTO decrease RNF168 protein levels. FTO showed similar RNF168 protein depletion phenotype as ALKBH3. siALKBH5 had no effect on RNF168 protein levels. SMC1 is shown as a loading control.

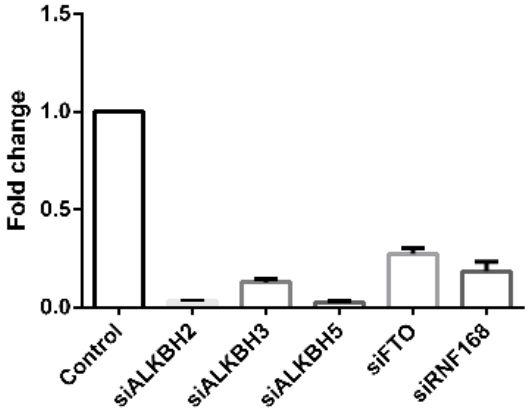


Figure 36. qPCR for knockdown efficiency of siRNAs. All siRNAs knocked down their targets efficiently. Error bars represent standard error of mean.

Interestingly, FTO knockdown showed the same phenotype as ALKBH3 (Figure 35). However, ALKBH5 knockdown showed no impact on

RNF168 levels. While this result was intriguing because it confirmed an additional RNA demethylase has an impact on RNF168, it also complicated our theory due to the differing demethylase targets of ALKBH3 and FTO. As mentioned in the introduction, the established targets of ALKBH3 are m1A and m3C, the former of which has been implicated in mRNA regulation. FTO, at the time of this experiment, had been established as a m6A demethylase. ALKBH5 had also been established as a m6A which made seeing no effect after ALKBH5 knockdown perplexing.

While this work was in progress, a paper was published detailing that FTO was preferentially targeting N6-2'-O-dimethyladenosine over N6-methyladenosine (Mauer et al., 2016). This importantly illustrated that the demethylase targets of these proteins have not been completely defined as of yet. Work is still being done to specify which modifications are being acted on by each demethylase, and that there are potentially modifications that have not yet been considered as targets. An additional important takeaway from this paper was the separation of targets between ALKBH5, which targets m6A, and FTO, now shown to target m6Am rather than m6A. This information provides a potential explanation as to why we saw the reduction of RNF168 phenotype in FTO and not ALKBH5.

4.2.11 ALKBH3 and FTO

Having now seen a replication of our phenotype using FTO, we became curious about any possible association between FTO and ALKBH3. There were multiple possibilities to consider as to how the regulation was occuring. One possibility was that ALKBH3 and FTO were in a complex, and were acting either in unison or having one act as an adaptor protein to demethylate the target mRNA methylation. Or one of these proteins could be impacting the expression of the other and cause the reduction of the other when knocked down. One could also be affected localization, either within the cellular compartments or to the direct mRNA site. Another possibility was looking at the methylation targets. Could RNF168 mRNA be targeted by both m1A and m6A and both were being removed as part of the regulatory process. Another option could be, as shown by the Mauer et al. paper, that the targets of ALKBH3 and FTO had not been well enough defined and both proteins were acting on the same modification.

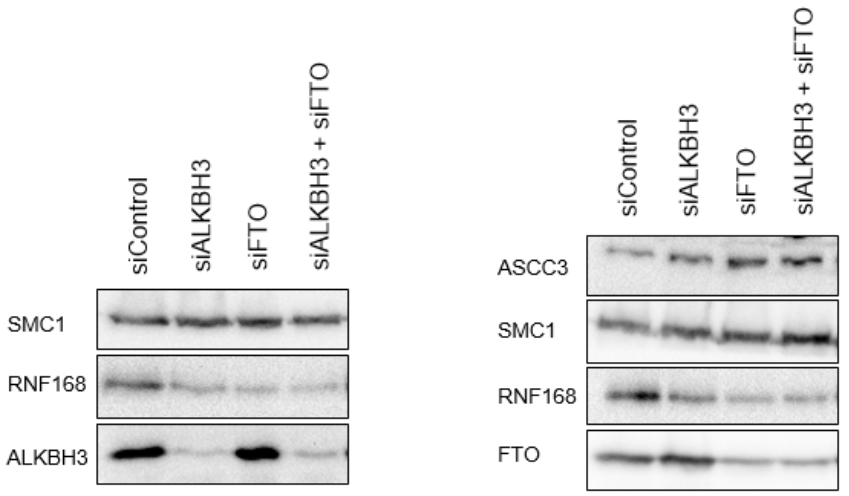


Figure 37. ALKBH3 and FTO have no impact on each other's protein levels. No depletion of ALKBH3 protein was seen following FTO depletion and the same was true for ALKBH3 depletion on FTO levels. Additionally no additive effect on RNF168 protein levels. SMC1 is used as a loading control.

The results of the western were interesting as they eliminated some theories as to the interaction between ALKBH3 and FTO. There was no reduction of FTO expression when knocking down ALKBH3, and vice versa. This eliminated the idea that perhaps one was affecting the other's expression or stability. It however does not rule out them acting in a complex to achieve their goal. It is worth noting that none of the published mass-spectrometry experiments for ALKBH3 have listed FTO as an interactor (Dango et al., 2011, Zhao et al., 2015). However, in our FLAG-pulldown experiments we have seen FTO being pulled down along with FLAG-ALKBH3 (Figure 38).

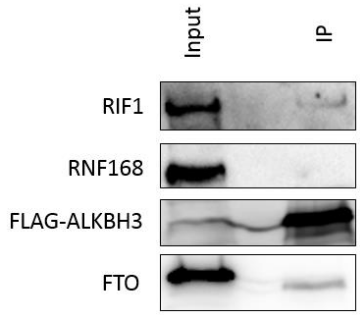


Figure 38. Western blot of FLAG-ALKBH3 immunoprecipitation. FTO is shown to be pulled down with ALKBH3. RIF1 is shown as a positive control.

I still consider these results preliminary and I am hesitant to state that there is an interaction between ALKBH3 and FTO as the FLAG-IP issues are still being remedied by the lab as mentioned previously. This interaction has been seen multiple times in our hands but has no external confirmation.

Another interesting result was the co-depletion of ALKBH3 and FTO showing no additive effect on the reduction of RNF168 levels. This points more towards them acting on the same modified nucleotide, or different modifications on different RNF168 transcripts rather than the action of both of them on a single RNF168 mRNA to achieve the phenotype.

As an independent confirmation that this phenotype is not an off-target effect of the siRNAs, we performed a western using siRNAs for ALKBH3 and FTO.

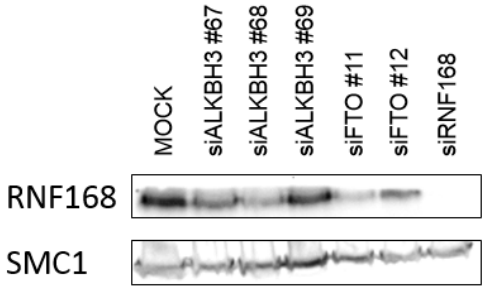


Figure 39. ALKBH3 and FTO knockdown with differing siRNAs. Reduction of RNF168 is seen in all cases. We selected siALKBH3 #68 and siFTO #11 to conduct most experiments with as they had the most efficient protein knockdown. SMC1 is used as a loading control.

4.2.12 Compensation of ALKBH3 by FTO expression

An idea arisen by the overlap in phenotype of ALKBH3 and FTO depletion was that there was potentially some compensatory role between the two proteins. The idea was that in the absence of ALKBH3 expression due to promoter methylation, FTO expression would increase in order to compensate. Conversely if the expression of ALKBH3 was high then FTO expression would decrease due to the potentially overlapping role. We checked the mRNA expression of ALKBH3 and FTO within all the breast cancer lines we had available and saw no such trend.

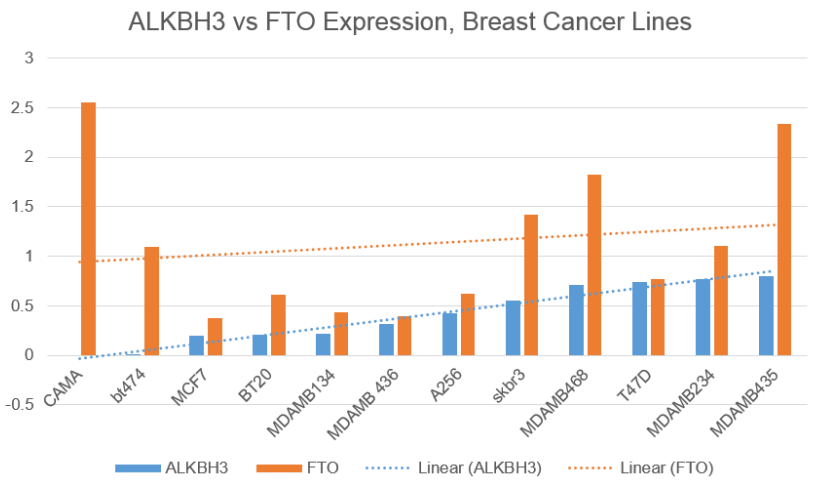


Figure 40. Expression of ALKBH3 and FTO mRNA in breast cancer cell lines, ranked by ascending order of ALKBH3 expression. qPCR analysis of our available cell line mRNA for ALKBH3 and FTO expression.

Additionally we checked within the Broad Institute's online database (https://www.broadinstitute.org/) of cell line mRNA expression and confirmed there was likely no compensatory increase in expression of FTO in the absence of ALKBH3 expression. It is however not completely ruled out due to not being able to directly compare expression between cell lines. The ALKBH3 promoter methylated cell lines, CAMA-1 and BT-474, do show relatively high expression of FTO. Also the compensation might be on the protein level, which you would not see by looking at mRNA levels.

4.2.13 Cellular Fractionation

The next step was to determine how methylation on mRNA is regulating the expression of RNF168. As mentioned in the introduction, there were a few established ways that methylation had been tied to regulation. First, we considered splicing, specifically “Alternative Last Exon Splicing”. We briefly checked this by comparing RNF168 primers that spanned exon 4 to 5 against primers that spanned 5 to 6 but saw no difference with and without ALKBH3 knockdown (data not shown). The other two known methods of regulation were nuclear export and translational control. We first looked into nuclear export due to the fact that if there is a defect there, there will be less mRNA to be translated. The first test was to perform a cellular RNA fractionation, and then check with qPCR the RNF168 levels in the nuclear and cytoplasmic fraction.

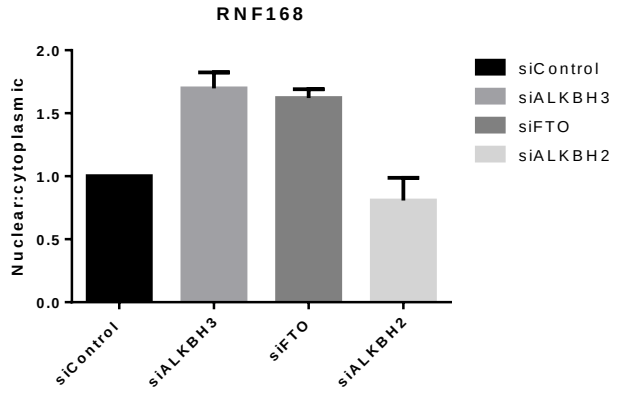


Figure 41. Nuclear:Cytoplasmic ratio of RNF168 mRNA following knockdown of ALKBH3, FTO and ALKBH2. Performed fractionation and measured RNF168 via qPCR in nuclear and cytoplasmic samples. Increased ratio of nuclear:cytoplasmic in ALKBH3 and FTO knockdown samples indicates nuclear export deficiency.

The fractionation showed an increased nuclear retention of RNF168 mRNA following knockdown of ALKBH3 and FTO, but not ALKBH2. This result provided additional evidence for the separate functions of ALKBH3 and ALKBH2, but also the redundancy of the mechanism for ALKBH3 and FTO. Combined with the previous result of unchanged total levels of mRNA, this pointed towards mRNA nuclear export as the area in which the methylation was regulating RNF168.

4.2.14 RNA Scope

In order to provide additional confirmation of the fractionation result, and in an effort to visualize the phenotype we performed RNA scope. RNA scope is a technique that utilizes multiple probes specific for a target mRNA and visualization by microscopy. We designed probes specifically for RNF168 mRNA, as well as standard control genes and performed the method with knockdown of ALKBH3 and FTO.

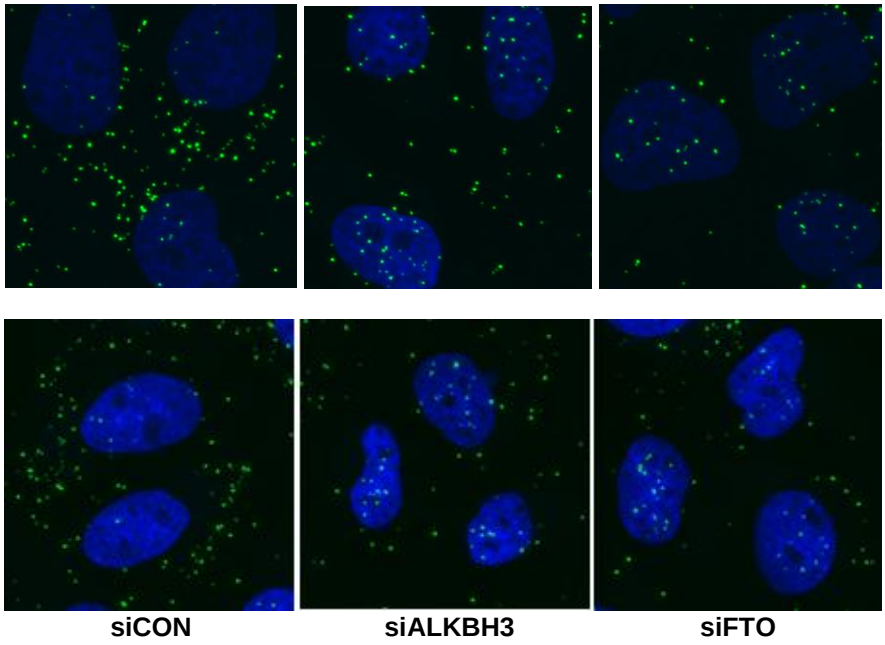
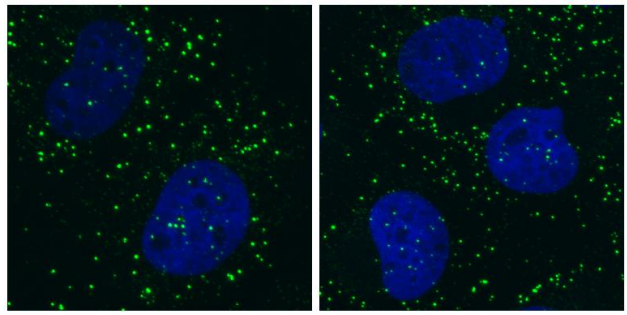
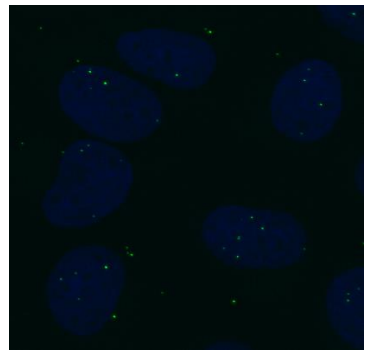


Figure 42. RNA Scope for RNF168 mRNA. In situ hybridization specifically targeted for RNF168 mRNA. Blue is DAPI staining for nucleus, Green is RNF168 mRNA. Nuclear retention of RNF168 foci seen in siALKBH3 and siFTO samples. Images were acquired by confocal microscopy and displayed as maximum-intensity projections.



siCON **siALKBH3**

Figure 43. RNA Scope for control gene PPIB. siControl and siALKBH3 showed no quantitative difference in nuclear export ratio. Images were acquired by confocal microscopy and displayed as maximum-intensity projections.



siRNF168

Figure 44. RNA scope for siRNF168. RNF168 foci almost entirely ablated following RNF168 knockdown. Images were acquired by confocal microscopy and displayed as maximum-intensity projections.

The RNA Scope showed that there were indeed RNF168 mRNA foci being retained in the nucleus following the knockdown of ALKBH3 and FTO. The quantification of the foci provided the same results as the fractionation, showing a 2-fold and greater increase in the nuclear to cytoplasmic ratio foci when compared to the control siRNA. Within the targeted control gene (PPIB) there was no notable change in the nuclear to cytoplasmic ratio.

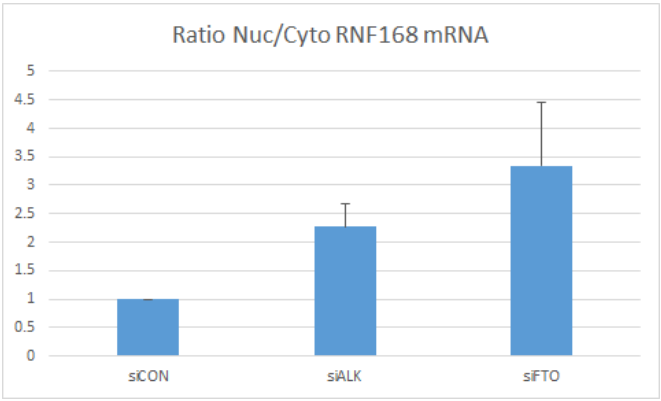


Figure 45. Quantification of RNA scope results. Data put into ratio of nuclear:cytoplasmic with control siRNA set as 1. This was performed in triplicate with error bars representing the standard deviation within the samples.

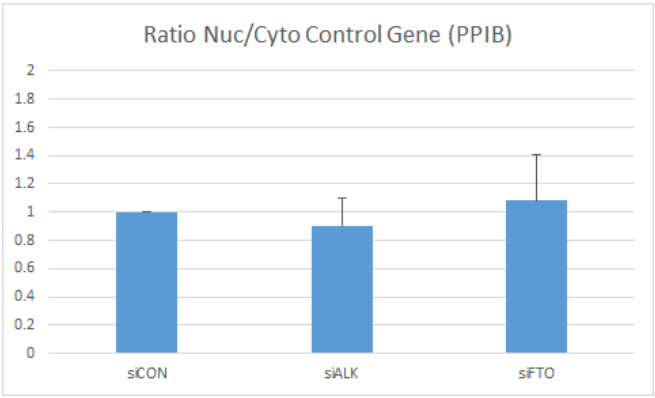


Figure 46. Ratio of Nuclear to Cytoplasmic for mRNA control gene (PPIB) of RNA Scope. Control genes for RNA scope showed no change in mRNA export efficiency. This was performed in duplicate with error bars indicating standard deviation.

With confirmation of the fractionation results by RNA Scope, we had narrowed our model towards nuclear export. There were several existing models for RNA export being affected by methylation. These models however differed in that they typically required methylation for recognition by proteins and were then subsequently exported (Roundtree et al., 2016; Dominissini et al., 2016) but our model would have to show decreased export with increased levels of methylation.

4.2.15 The impact of ALKBH3 knockdown on DSB repair

We used cell line reporter models in order to examine if the downregulation of RNF168 by ALKBH3 mimics the DSB repair consequences found in literature. We used a DSB repair reporter cell line (DR-GFP) to measure the use of gene conversion (GC), the accurate form of homology directed repair, and SSA, the deleterious form of homology directed repair. This model uses a plasmid with a I-Sce1 restriction site that can have a DSB induced. There are primer pairs and different points on a plasmid to measure the type of repair used. If the cell uses GC, there will be limited resection and provide a PCR product with the first reverse primer. Additionally the use of GC in the DR-GFP line will restore a GFP coding sequence, meaning the cells using GC to repair the DSB can be sorted by FACS based off of a GFP signal. If SSA is being used, the cell will resect the DNA until the more distant homologous sequences are linked and provide a PCR product with the second reverse primer.

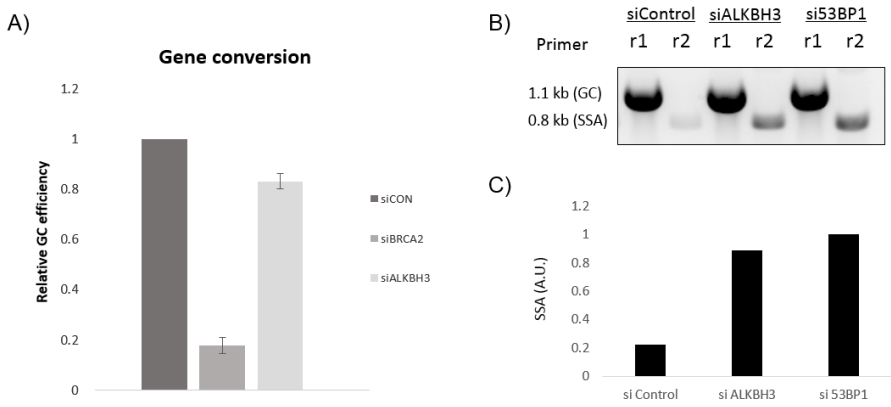


Figure 47. The impact of ALKBH3 knockdown on GC and SSA efficiency. Panel A shows GC efficiency. ALKBH3 confers a 20% reduction in GC efficiency. BRCA2 is used as a control for GC reduction. Panel B shows a PCR for the SSA assay. The lower band (0.8 kb) represents the amount of SSA occuring within the sample. Reverse primers will provide an indication of which form of HDR is used within the sample. 53BP1 is used as a control by promoting the use of SSA. Panel C is a quantification of panel B's SSA bands with A.U., arbitrary units.

Our test yielded results that indicated ALKBH3 does in fact have an impact on DSB repair pathway choice. There was a 20% reduction in GC efficiency within the ALKBH3 knockdown sample (Figure 47 A). The SSA assay indicated there is a increase in use of SSA repair comparable to that of 53BP1 (Figure

47 B&C). The SSA results are however still preliminary as we have only seen this result once. Despite being preliminary, this result importantly indicates that the absence of ALKBH3 decreases, but does not eliminate gene conversion and promotes the use of a mutagenic form of DSB repair. The absence of ALKBH3 is comparable with that of 53BP1 depletion seen in Ochs (2013), which showed the lack of 53BP1 reduces GC while promoting the use of SSA (Ochs et al., 2013).

4.3 Deciphering the Mechanism of Regulation

4.3.1 Consensus Sites and site identifying program

For a mRNA to be demethylated, it must contain methylation. Methyltransferases are the proteins responsible for the addition of methylation. These proteins have long been studied, but recently the specificity in which they act has been expanded on. Several of these proteins have had consensus sites or specific motifs in which they bind to. We chose to look at the proteins responsible for the deposition of methylation we were interested in due to the known demethylase activity of ALKBH3 and FTO. For m1A, there was a known methyltransferase complex containing TRMT6, TRMT61 and TRMT10C. This complex was identified to have preference for the nucleotide motif GUUCRA (R representing a purine) (Li et al., 2017). This complex had also been identified as recognizing tRNA-like loops in mRNA (Safra et al., 2017). In order to check for this and other consensus sites, we developed a custom tool that would derive all possible permutations of a motif and identify them within a FASTA sequence. We began by checking RNF168 mRNA (ID = NM_152617.3) to see if it contained the GUUCRA consensus site.

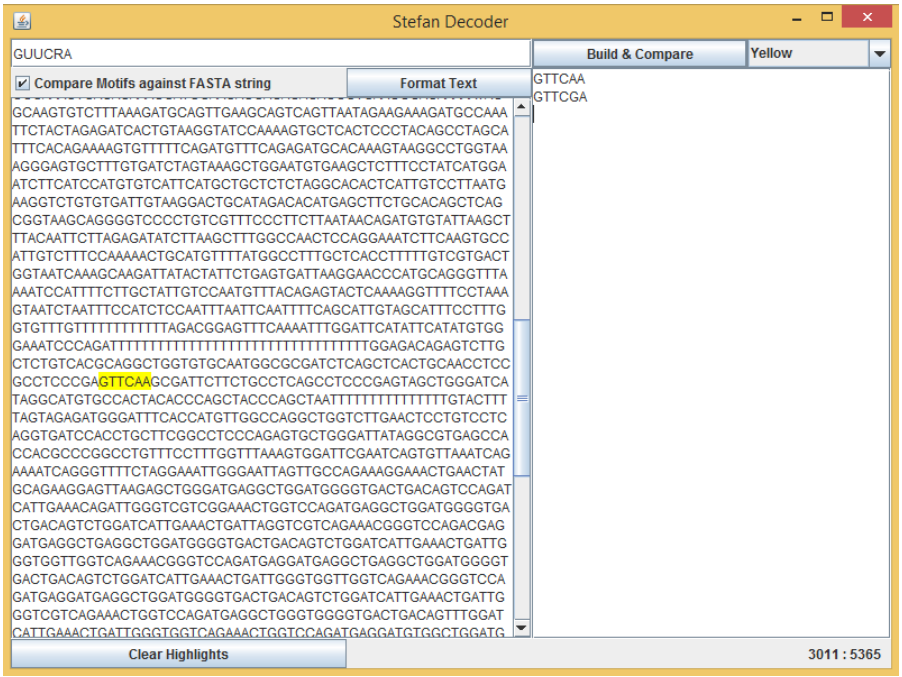


Figure 48. Identification of GUUCRA consensus motif location(yellow) within RNF168 mRNA FASTA sequence (NM_152617.3). GUUCRA motif gives the permutations GTTCAA and GTTCGA.

The RNF168 mRNA sequence did in fact have one GUUCRA motif starting at nucleotide 3011 (Figure 48). This placed the motif in the 3' UTR of the mRNA perhaps indicating that if this was indeed the location of the modification it would serve a regulatory role prior to the translation. Meyer et al. demonstrated that m6A residues are enriched within the 3' UTR and near stop codons and have a regulatory effect, so the placement of the GUUCRA motif in the 3' UTR is not out of line with existing research (Meyer et al., 2012). We checked for a potential presence of a tRNA-like loop using predictive mRNA folding software from The RNA Institute from the State University of New York at Albany (http://unafold.rna.albany.edu/?q=mfold%2Frna-folding-form), but the software did not indicate the presence of a tRNA motif at that location.

The next motif to look at was the one responsible for m6A. The methyltransferase complex responsible for the deposition of m6A is the METTL3-METTL14-WTAP complex described in the introduction. The motif identified for this complex has been established as DRACH (where D

represents A, G or U, and H represents A, C or U) (Grozhik et al.; 2015, Meyer et al., 2012).

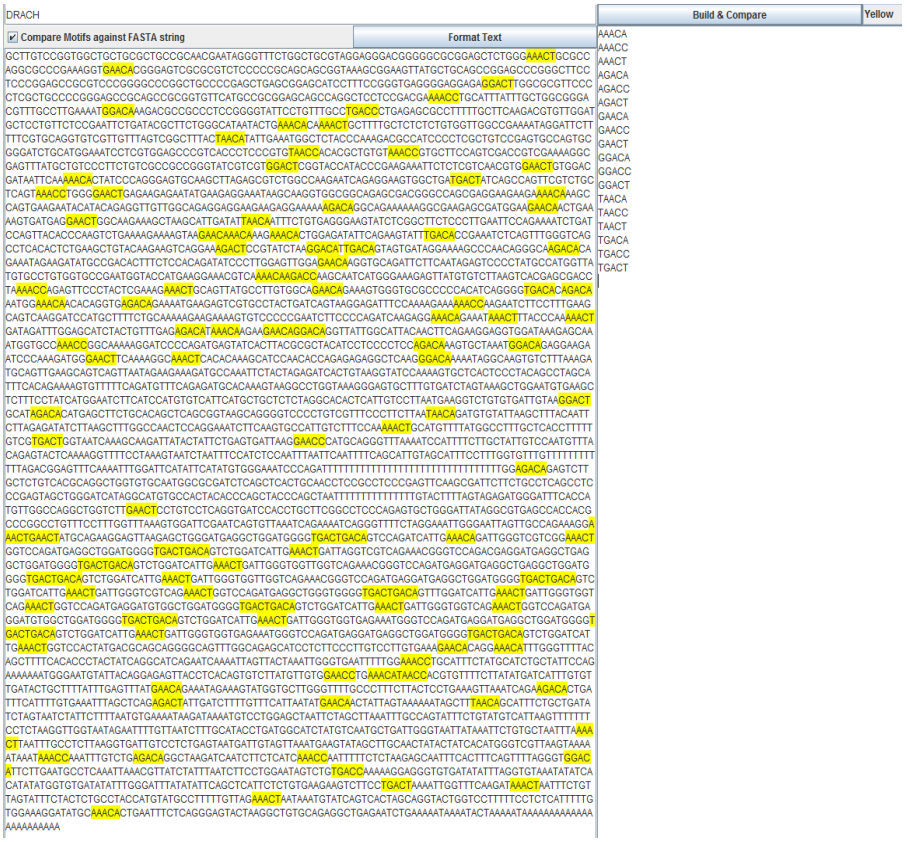


Figure 49. DRACH motifs present in the RNF168 mRNA. DRACH motifs exist for the METTL3/METTL14/WTAP complex as a targeted nucleotide motif for the deposition of methylation at the N6 position.

Clearly the existence of DRACH motifs within the RNF168 mRNA is far more common (Figure 49). It is thought that m6A occurs around 3 times per every transcript in the genome so the presence of DRACH motifs in abundance comes as no surprise. The motif gives the possibility of 18 different permutations of 5 nucleotides. These permutations are found in all components of the the mRNA and show only that it is possible that m6A is being placed on one of these locations. The only indication of a specifically targeted locus may be at nucleotide 900 as identified by a methylome created by Dominissini et al. (Dominissini et al., 2012). This would place the m6A modification in the second exon of RNF168 mRNA.

While m6A potentially and likely has a presence on RNF168 mRNA, its impact can only be speculative at this point. The interest in m6A stems from FTO displaying the same phenotype as ALKBH3, but FTO has since been shown to preferentially target m6Am over m6A. Current methods for RNA-immunoprecipitation make it difficult to distinguish between m6A and m6Am due to the non-specificity of antibodies. Distinguishing between these two modifications can currently only be done with Thin Layer Chromatography, or possibly LC/MS (Mauer et al., 2016). These techniques are not currently established within our lab and neither has been shown to be possible to use when targeting a single mRNA.

4.3.2 Methyltransferases depletion

The presence of methyltransferase binding motifs made us wonder what the impact of knocking down the methyltransferases would have on RNF168. Our initial idea was that this modification, whether it be m1A or m6A, was perhaps responsible for the nuclear export defect, and ALKBH3 removing the modification would allow for the normal export. Therefore co-depleting the methyltransferase and ALKBH3 would rescue the observed phenotype and restore the RNF168 protein.

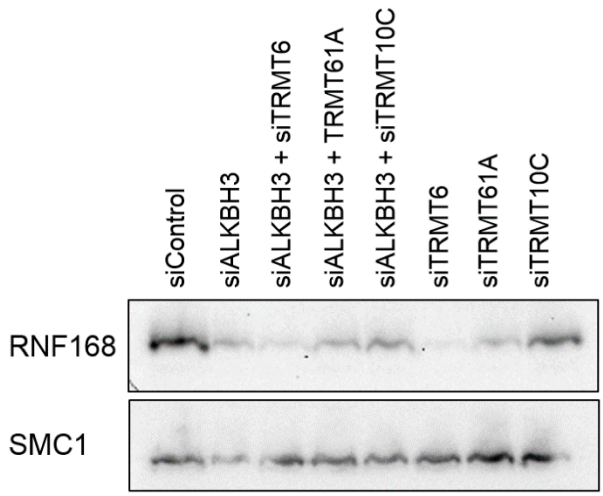


Figure 50. Co-depletion of ALKBH3 and m1A depositing complex members TRMT6, TRMT61A and TRMT10C. RNF168 protein levels go down in all cases, including TRMT complex individual knockdowns. SMC1 is used as a loading control.

Interestingly, but disappointingly at the time, this was not the case. The knockdowns showed no rescue following co-depletion, and the TRMT proteins themselves caused a dramatic reduction of RNF168 levels (Figure 50). My hypothesis had to change, as this pointed towards the methylation being a necessary intermediate step in the process between transcription and nuclear export.

We additionally tested the same idea for the m6A methyltransferase complex with a similar result.

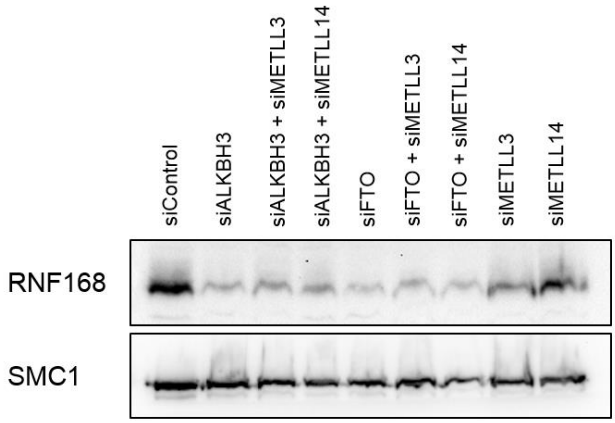


Figure 51. Co-depletion of METTL3 and METTL14 with FTO and ALKBH3. No rescue of RNF168 protein seen with co-depletion. SMC1 is used as a loading control.

The m6A methyltransferase proteins additionally showed no indication of a rescue following co-depletion of the methyltransferases and demethylases (Figure 51). We therefore reexamined whether the knockdown of the methyltransferase complex members alone would have any impact on RNF168 expression.

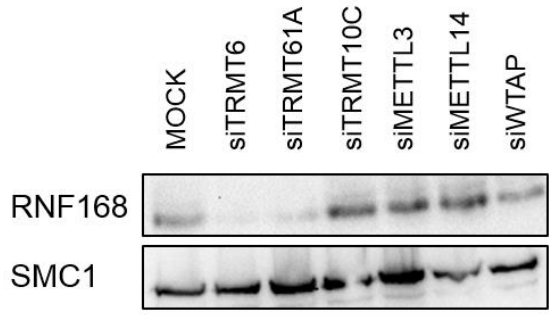


Figure 52. Western for methyltransferase depletion. Knockdown of TRMT6 and TRMT61 show knockdown of RNF168. TRMT10C, METTL3, METTL14 and WTAP appear to not affect RNF168 expression. SMC1 is used as a loading control.

The knockdown of TRMT6 and TRMT61A causing knockdown of RNF168 protein levels gave additional credence to the idea that RNA methylation is the culprit for our phenotype, and specifically m1A. By having an impact on the RNF168 protein levels, it can be inferred that the m1A methyltransferases are somehow involved in the regulation, possibly as an intermediate step between transcription and nuclear export. My theory as to how these methyltransferases are involved in the regulation is detailed later in the discussion chapter. The lack of a phenotype with TRMT10C is in line with the reported role of TRMT10C being present at the mitochondria, in the sense that it would likely not be involved in our nuclear export defect (Safra et al., 2017).

4.3.3 SRSF and YTH Domain proteins

The decoder tool allowed for the quick identification of proteins with potential binding sites with the RNF168 mRNA. With confirmation that motif-recognizing methyltransferases proteins indeed had some effect on RNF168, we began to look for additional proteins with established consensus motifs. A 2009 study from Long and Caceres identified consensus motifs for a majority of the SR (Serine and Arginine rich) protein family. Of the 10 SR proteins, SRSF1, SRSF3 and SRSF7 had consensus motifs with matching sequences within the RNF168 mRNA.

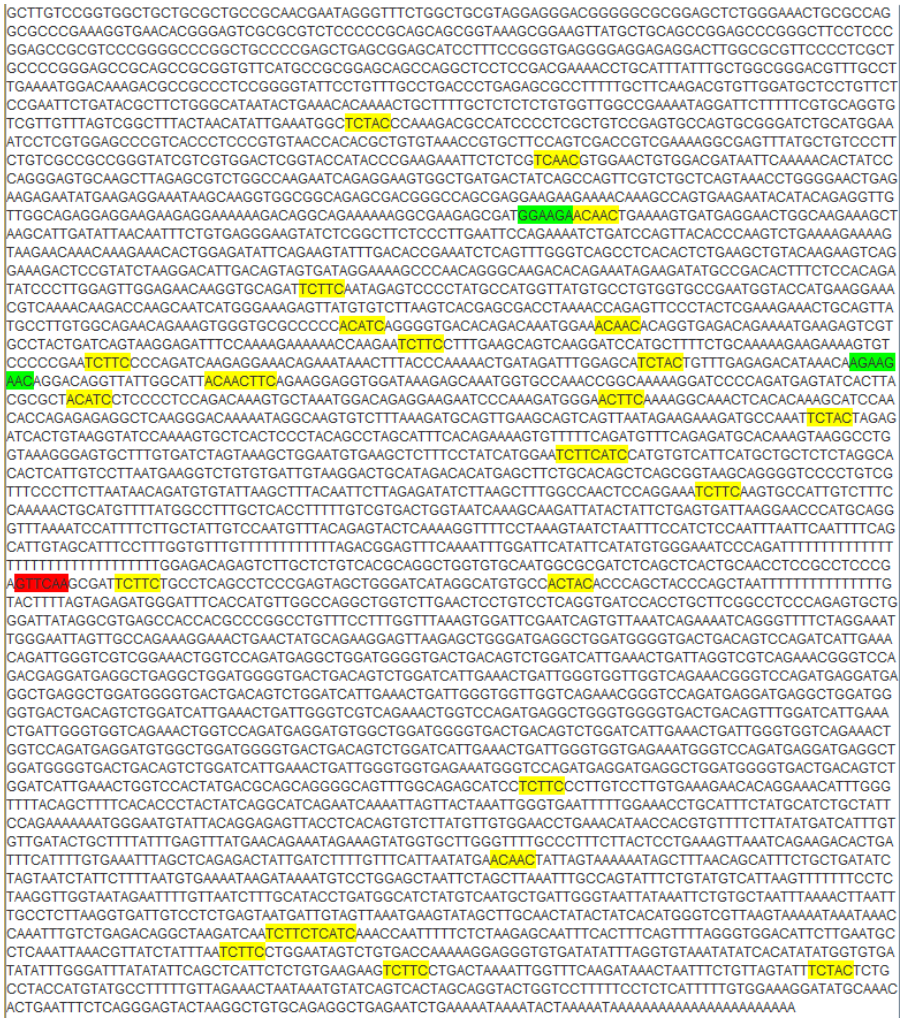


Figure 53. Consensus motifs for SRSF proteins. Identification of consensus motifs for SRSF1 (RGAAGAAC), highlighted in green and SRSF3 (WCWWC), highlighted in Yellow. Red highlight indicates positioning of theoretical m1A site based on GUUCRA motif.

Interestingly, SRSF3 had already been implicated in nuclear export (Roundtree et al., 2016). SRSF3 also had a binding motif just 5 base pairs away from the GUUCRA motif, the TRMT recognition motif, we had identified. Additionally, the Roundtree paper describing the role of SRSF3 in export also detailed it binding to YTHDC1 and NXF1. The YTH domain proteins had already been of interest due to their published recognition of m6A within mRNA. YTHDF1, YTHDF2, YTHDF3 and YTHDC1 had also been recently

revealed to recognize m1A as well as m6A (Dai et al., 2018). We had discounted looking at the YTHDF1-3 proteins as their roles have been established as exclusively cytoplasmic, and not been considering YTHDC1 until the information about it recognizing m1A was published. We decided to knockdown SRSF3, YTHDC1 and hnRNP K. hnRNP K was one of the proteins pulled down in the ALKBH3 LC-MS/MS by Dango (2011) and has a described role in mRNA transport. Additionally we took a look at NXF1 as it was published as being recruited by SRSF3 and is a described nuclear export protein (Roundtree et al., 2016).

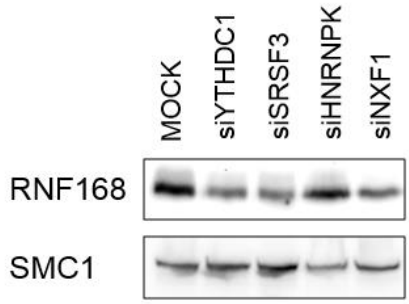


Figure 54. Western blot for m1A interactors and export proteins. Western blot showing reduction of RNF168 protein levels following YTHDC1, SRSF3, and NXF1 knockdown. HNRNPK showed minimal to no reduction. SMC1 is used as a loading control.

The SRSF3 protein identified by the decoder, it's interactor YTHDC1 and NXF1 all showed decreased levels of RNF168 protein. HNRNPK showed little to no knockdown of RNF168. This indicated that several proteins that have thus far been implicated in nuclear export and shown to recognize m1A are also potentially interacting with RNF168 mRNA. This also provides additional credence to the methylation dependent mRNA nuclear export idea.

4.4 Summary

The functional consequences of lacking ALKBH3 likely extend beyond increased levels of DNA alkylation damage. This project has shown that the knockdown of ALKBH3 affects the protein expression of RNF168, a key protein in DNA DSB repair pathway choice (Figures 26, 30). In the process of understanding the regulatory pathway, we eliminated many classical forms of regulation and determined nuclear export to be the targeted pathway of

regulation (Figures 30-33).

Work on proving the export defect model is still ongoing within our lab. The fractionation and RNA scope provide strong evidence that a nuclear export deficiency is responsible for the decreased levels of RNF168 protein (Figures 20-25). The RNA-immunoprecipitation gave an indication that RNA methylation is responsible for this export defect (Figure 34). The work went on to show multiple other proteins with established roles in RNA methylation pathways to have an affect on RNF168 protein expression (Figures 48-54). This included proteins considered readers and writers of RNA methylation. As the work in our and other labs continues, the exact mechanism through which this regulation occurs should become clearer.

5 Discussion

The DNA damage response is crucial to maintaining the integrity of DNA and the health of a cell. Unrepaired lesions within the DNA can lead to genomic instability and potentially aid in the formation of diseases such as cancer. Some cancers possess dysfunctional DNA repair and chemotherapeutic treatments may aim to exploit this weakness that distinguishes cancer cells from normal healthy cells. ALKBH3 is a DNA repair protein involved in the repair of alkylation damage (Aas et al., 2003; Sedgwick et al., 2006; Dango et al., 2013). Within The Cancer Genome Atlas we found that ALKBH3 contained a hyper-methylated promoter in 20% of breast cancers. This hyper-methylation, a form of epigenetic regulation, lead to a dramatic reduction of ALKBH3 mRNA expression and therefore a decrease in total ALKBH3 protein levels (Stefansson et al., 2017).

In the first part of this project, we examined this event within a sample of Icelandic breast tumors and found the incidence of ALKBH3 promoter methylation to be present in a subset of cancers, ranging from 5-27%, based on the parameters. This indicated that ALKBH3 promoter methylation within breast cancers is likely a global phenomenon. Importantly, this methylation occurred only within the tumor tissue, but not the normal tissue of the same patients. This therefore identified a differentiating characteristic for cancer tissue based on the status of a DNA repair protein. Additionally, tumors with higher levels of ALKBH3 promoter methylation exhibited decreased survival rates. With this knowledge, the next step of the project was to elucidate what the functional impact of lacking ALKBH3 would be.

We discovered that the knockdown of ALKBH3 caused a decrease in protein levels of RNF168, a protein crucial in ubiquitin signaling and effective DNA double-strand break repair. We methodically ruled out various forms of regulation including transcription, protein stability and interaction at the site of damage. After some review of the literature, we decided to examine if RNF168 is regulated by ALKBH3 through RNA methylation, a form of regulation known as epitranscriptomics. Epitranscriptomics is the study of biochemical marks on RNA. Similar to epigenetics, epitranscriptomics involves modifications to nucleotides that may alter their functions, but has no impact on the sequence of nucleotides. A form of these modifications occur as a methyl group. The addition and removal of these methyl groups from nucleotides within mRNA has been implicated in having roles within alternative splicing (Zhao et al., 2014), mRNA export (Wickramasinghe et al., 2015), and translational efficiency (Lin et al., 2016; Slobodin et al., 2017), thus causing differing

expression of proteins and affecting cells accordingly. Our data indicates that ALKBH3 is acting on N1-methyladenosine (m1A), a known ALKBH3 target, on the RNF168 mRNA. The knockdown of ALKBH3 allowed for increased levels of RNF168 mRNA to be pulled down by RNA-immunoprecipitation due to increased levels of m1A modified mRNA.

We were then able to show ALKBH3 is regulating cellular localization of RNF168 mRNA by impacting nuclear export. Following ALKBH3 knockdown, we saw nuclear retention of RNF168 mRNA in the nucleus. This was confirmed by using both cellular fractionation as well as a RNA FISH method called RNA Scope.

The function of ALKBH3 was ensured to be separate from its known alkylation repair function by checking if ASCC3 or ALKBH2 shared the phenotype, which they did not. With indications that RNF168 is being modified by m1A, and this modification is being regulated by ALKBH3, we went back into the literature to study demethylases. Interestingly, two of the 9 AlkB members, ALKBH5 and FTO, had been implicated in RNA demethylase activity (Jia et al., 2011, Zheng et al., 2013). We performed a knockdown of ALKBH5 and FTO to see if there was any additional action by RNA demethylases on RNF168 protein levels.

Interestingly, FTO knockdown showed the same phenotype as ALKBH3. However, ALKBH5 knockdown showed no impact on RNF168 levels. FTO additionally showed the same nuclear export defect phenotype when checked with fractionation and RNA Scope. While this result was intriguing because it confirmed an additional RNA demethylase as having an impact on RNF168, it also complicated the project due to the differing demethylase targets of ALKBH3 and FTO. As mentioned in the introduction, the established targets of ALKBH3 are m1A and m3C, the former of which has been implicated in mRNA regulation. FTO, at the time of this experiment, had been established as a m6A demethylase. ALKBH5 had also been established as a m6A demethylase which made seeing no effect after ALKBH5 knockdown more confusing.

Around the time we got this result, a paper was published detailing that FTO was preferentially targeting N6-2'-O-dimethyladenosine (m6Am) over N6-methyladenosine (m6A) (Mauer et al., 2016). This importantly illustrated that the demethylase targets of these proteins have not been completely defined as of yet. Work is still being done to specify which modifications are being acted on by each demethylase, and that there are potentially modifications that have not yet been considered as targets. An additional important takeaway from this

paper was the separation of targets between ALKBH5, which targets m6A, and FTO, now shown to target m6Am rather than m6A. This information gave some explanation as to why we saw the reduction of RNF168 phenotype in one and not the other.

The regulation of RNF168 by ALKBH3 is important for numerous reasons. RNF168 being regulated in an epitranscriptomic manner is a novel method of regulation for a protein that is already heavily regulated (Bartocci et al., 2013). This is also the first example of DNA repair being regulated by epitranscriptomics, additionally by a protein that is known as a DNA repair protein. The DNA repair field has perhaps previously downplayed or not thoroughly examined the importance of mRNA when examining repair defects. Most commonly research is conducted on mutations within the genetic code that then cause defective protein function and any regulation on the mRNA level is ignored or simply not considered. The field of DNA repair may have to reconsider some of its approaches in the future. Even outside of DNA repair, epitranscriptomics is likely to become a recognized and important form of regulation to be considered in future research and disease treatment.

5.1 General Discussion

5.1.1 Variation in promoter methylation levels

A point of discussion regarding the epigenetic aspect is the varying level of promoter methylation within samples. Why does this occur and when does this occur in the cancer development? A more obvious explanation into why variance is seen is due to the heterogeneity of the sample. Tumors can be extremely heterogeneous in composition and therefore not contain the same genotype nor display the exact same phenotype (Shipitsin 2007). This obviously greatly complicates treatment when attempting to use precision medicine. If only a percentage of the tumor has a defect in a particular repair gene, that means treatment that exploits this weakness may only affect the clones with the defect, allowing the rest of the tumor to persist and potentially causing further heterogeneity due to secondary tumor creation by the chemotherapy.

An additional source of heterogeneity in samples that must be considered is the method in which the samples are acquired. All samples used in this thesis come from surgery removing primary breast tumors and adjacent normal tissue. When collecting tumor tissue it is difficult to insure that purely tumor tissue is being collected for the sample. When the DNA and RNA is isolated from the tissue it removes any capability to perform a stain to check

for the tumor status of the tissue prior to analysis. This may result in samples that vary greatly in the amount of tumor tissue contained in the isolated DNA/RNA.

A point that is then raised from heterogeneity is when ALKBH3 promoter methylation occurs during cancer development. It is extremely difficult, likely impossible with current techniques to determine this, but our data may give us clues into where it may occur. Our data indicates that tumors with greater than 20% promoter methylation have a statistically impacted survival, yet it is difficult to ascertain if the ALKBH3 promoter methylation was responsible for the decreased survival of these patients. Overall, we saw 27% of patients exhibit some level of promoter methylation. The mechanisms in which promoter methylation is regulated are currently poorly understood, so it is difficult to rule out hyper-methylation being undone at some point in the progression of the tumor. Rather than look at the epigenetic angle of ALKBH3, it is perhaps more relevant to look at the functional impact of not having ALKBH3.

5.1.2 The functional impact of ALKBH3 silencing in DNA repair

ALKBH3 has a well established role in the repair of alkylation damage (Aas et al., 2003; Sedgwick et al., 2006; Dango et al., 2013). In our paper we showed that knockdown of ALKBH3 allowed for increased levels of endogenous m3C (Stefansson et al., 2017). Increased m3C can cause decreased transcription efficiency and replication fork stalling, as well as promote the use of translesion polymerases such as POLQ that promote the use of error prone methods of DNA repair (Wood et al., 2016). The combination of these events may allow for increased formation of mutations within the DNA and therefore provide some of the “hits” required for a cell to become carcinogenic.

We also have to consider the implications of the regulation of RNF168. As seen in this thesis, the knockdown of ALKBH3 dramatically reduces, but does not fully deplete RNF168 protein levels. RNF168 has a clear role in regulating efficient and accurate DSB repair, and the disruption of the RNF168-53BP1 pathway has been shown to promote the use of the highly mutagenic SSA repair pathway (Ochs et al., 2013). Ochs (2013) furthermore showed that the combination of BRCA1 and 53BP1 knockout causes cells to depend on RAD52, thereby necessitating the use of SSA for cell survival (Ochs et al., 2013). The use of SSA will confer a resistance to PARP inhibitors, which are a common form of treatment for breast cancer that are HR deficient (Lord and Ashworth, 2012). However the loss of a single BRCA1 allele and RNF168 will

cause sensitivity to PARP inhibitors, with RNF168 having been shown to be crucial for promoting RAD51 loading in BRCA1 +/- cells (Zong et al., 2019). Cancer formation and cancer resistance to drugs is often dependent on more than one dysfunctional or disregulated protein. The absence of BRCA1 alone will not cause a cell to be resistant to PARP inhibitors, in fact it hypersensitizes it to that treatment (Farmer et al., 2005). Epigenetically silencing ALKBH3, which may then downregulate RNF168, which then inhibits 53BP1 recruitment, could theoretically be the second hit to a BRCA1 mutated cell that will cause it to turn to SSA for repair in times of stress. Using SSA could mean this cell could be rapidly acquiring new mutations, potentially in oncogenes, therefore promoting the formation of a new cancer. Or in case the cell is already considered an identified cancer, it could mean resistance to treatment with chemotherapeutic drugs. Our preliminary DNA repair assay results indicate that ALKBH3 depletion does in fact promote the use of SSA. This may then imply that ALKBH3 promoter methylation could be used as a marker when deciding treatment options, particularly in cancers that are HR deficient.

Because ALKBH3 does not fully deplete RNF168, it may allow for the accurate DSB repair pathways to still occur that allow the cell to survive, but decrease them to a point where mutations are allowed to “slip through” and aid in the formation of a disease. This resembles in some sense a haplo-insufficiency type case, such as those seen as a potential explanation of BRCA1 and BRCA2 related cancers (Pathania et al., 2014; Stefansson et al., 2011; Skoulidis et al., 2010).

Even in cases where there are no additional mutations or epigenetic events impairing DNA repair, the downregulation of RNF168 could still have consequences on the ability of the cell to perform accurate repair. As detailed in the introduction, RNF168 is key to the recruitment of 53BP1 and proteins downstream of 53BP1 such as PTIP, SHLD1, SHLD2, SHLD3, and REV7. These proteins all play a crucial role in accurate DNA repair. Preventing recruitment of factors important to both NHEJ and HR could be especially problematic in cells that are non-dividing and therefore do not have HR available. This again would force the cell to deal with genotoxic stress with forms of repair caused by hyper-resection that are mutagenic and potentially conducive to the formation of a disease.

When taken together, my impression would be that the use of mutagenic forms of DNA repair may aid in the formation of a mutator phenotype within the DNA. This could potentially point to ALKBH3 promoter methylation being perhaps an early event in carcinogenesis. It is not

completely disabling a cell's ability to perform effective repair, but weakening to an extent that may eventually disrupt the normal function of other genes.

5.1.3 ALKBH3 regulation in cancer and tissue specificity

Although this thesis has entirely focused on the absence of ALKBH3, there exists cases where the opposite is occurring. ALKBH3 has been shown to be overexpressed in various cancers, including lung and prostate cancer (Tasaki et al., 2011; Yamato et al., 2012). This is interesting in contrast to the paper our group published showing that the ALKBH3 promoter is hypermethylated, and down-regulated, within a portion of breast cancers. Our work was later validated by another group showing the same epigenetic silencing of ALKBH3 in breast cancer (Knijnenberg et al., 2018). Yet another paper was published showing ALKBH3 epigenetic silencing occurs in multiple cancers including high levels of silencing in endometrial, stomach and cervical cancers (Saghafinia et al., 2018). This importantly demonstrates not only that ALKBH3 silencing is not breast cancer specific, but ALKBH3 regulation can vary from tissue to tissue.

Given that ALKBH3 has a variety of roles in protein regulation, perhaps largely through mRNA methylation control, it can then be assumed dysregulation of ALKBH3 may have other consequences. Akin to hyper- and hypomethylation within promoters of genes that have been associated with disease, variation of the normal levels of methylation and regulation by ALKBH3 could be conducive to the formation of disease.

The 2012 Yamato study showed knocking down of ALKBH3 in pancreatic tumor cells that overexpressed ALKBH3 would cause a decrease in tumorigenesis and increase in apoptosis. While they don't specifically designate a role that ALKBH3 is playing in these cancers, they bring up a potential link in the regulation of VEGF. Interestingly, VEGF is listed as containing a m1A site in the Dai et al., 2018 paper discussing m1A being a target of the YTH family proteins (Dai et al., 2018). VEGF itself is known to serve a role in angiogenesis, and angiogenesis is key in the progression of cancers by providing blood supply to growing tumors (Carmeliet 2005). Yamato saw a decrease in VEGF expression following ALKBH3 knockdown. However in this case mRNA expression is decreased, likely causing to the decreased protein levels. Together with the known m1A site, this could implicate ALKBH3 in having a role in the regulation of VEGF expression through epitranscriptomic means. This would imply that the mechanism of ALKBH3 regulating m1A in mRNA is not simply an export defect, but could

have varied consequences and actions depending on the location of the methylation in the mRNA transcript and the tissue in which the transcript is being acted on. It is wrong to assume that one mechanism applies to all RNA affected by m1A. We can not simply apply the nuclear export defect phenotype to VEGF without repeating many of the same experiments conducted in this thesis. Looking at the example of m6A, there is a wide variety of ways that methylation marks can regulate protein expression. Many of these variances come from positioning of the methylation on the RNA. ALKBH3 may in fact be regulating a large number of m1A sites within RNA, but the result of that regulation is likely to vary quite a bit based on the location of the modification. Furthermore these mechanisms could vary from tissue to tissue.

5.1.4 Methylation as a form of regulation

Methylation of RNA has at this point clearly been demonstrated as a form of regulation. In the future, as the exact ways in which RNA methylation is controlled and impacts cells are discovered, RNA methylation regulation may be thought of similar to other forms of regulation in the cell such as protein ubiquitylation or other types of protein modifications.

The are many potential purposes of regulation by RNA methylation. It could be a form of regulation used to ensure less waste of nucleotides. Producing proteins when unnecessary is wasteful and “expensive” for the cells. It would be interesting to look at this type of regulation during starving conditions to see if there is an upregulation of RNA being controlled by methylation in order for cells to not waste energy on non-essential proteins. Li (2016) touched somewhat on the idea, showing that UV treated cells had a differing abundance of m1A modified mRNAs (Li et al., 2016).

As proposed by my theoretical model, perhaps methylation is a necessary signal for mRNA splicing. As seen in Roundtree et al. 2016, the absence of methylation will caused decreased binding of YTHDC1. Without YTHDC1, SRSF3 will not be recruited. SRSF10 will be recruited to the mRNA and promote exon-excluding splicing. RNF168 mRNA was shown to be a significant target of YTHDC1 indicating this pathway is likely with some merit (Roundtree et al., 2017). This same model could be applied to many different mRNAs, some with necessary differential splicing. Methylation writers are therefore potentially very important as in their absence there will be either incorrect or no splicing. When there are no erasers, many forms of expression are dysregulated. An example being the pathway seen in this thesis. The absence of a methylation eraser, ALKBH3 or FTO, can cause correct nuclear

export to be inhibited and therefore decrease protein production.

5.1.5 Cellular proliferation decrease in absence of ALKBH3

It has been shown that cellular proliferation decreases when ALKBH3 is knocked down or absent. It has been suggested that increased levels of alkylation damage may be responsible for the decreased proliferative ability (Johannessen et al., 2013). There is certainly merit to this analysis due to cells usually needing to repair damage in order to clear cell cycle checkpoints to proceed through the phases of the cell cycle. ALKBH2, which is believed to function solely in the DNA alkylation repair role, likely exhibits this phenotype for this reason when knocked down. As Ringvoll (2006) showed, *AlkBh2* knockout mice have increased levels of m1A in their genomic DNA as well as showing increased sensitivity to MMS. This was however not the case in *Alkbh3* knockout mice (Ringvoll et al., 2006). This may therefore imply the main role of ALKBH3 is more complex, and could quite possibly be the regulation of methylation on RNA. As mentioned earlier, proteins involved in proliferation like VEGF, appear to have their mRNAs modified by m1A and could be targeted by ALKBH3.

5.1.6 Rethinking Silent Mutations and Non-Coding Variants

One aspect raised by the specificity in which methylation on mRNA is regulated is the impact of silent mutations. Silent mutations are considered such because although they are recognized by a different tRNA, the mutation will not affect the resulting amino acid due to the wobble hypothesis. A base pair substitution mutation can occur in the DNA and not affect the resulting protein. However if consensus and binding sites of methyltransferases and other proteins such as splicing factors are as specific as suggested, silent mutations may have a greater impact on the regulation of a protein than expected. The consensus motifs are typically 5-8 base-pairs, and disrupting one of these bases may cause a protein to not recognize the site. Additionally for bases that are modified, such as adenosine becoming N1-methyladenosine, a mutation would cause no modification to be added and disrupt and potential regulatory mechanism requiring the modification. While silent mutations typically only apply to mutations within the coding sequence, the same principle would be appropriate to consider for mutations within the 5’ and 3’ UTR. The methylation considered in this thesis are largely found in the regions outside the CDS (Safra et al., 2017; Mauer et al., 2017).

This would also increase the importance of considering mutations that lay outside the CDS. Currently genome-wide association studies often look at loss of function mutations to explain why decreases of expression might occur. This is often attributed to protein truncation leading to improper protein folding and subsequent degradation of the misfolded or dysfunctional proteins. This type of study requires to perform functional analysis of the protein to confirm folding or functional defects which is labor intensive and sometimes inconclusive. There are likely cases of protein expression decreasing due to mutations that disrupt regulatory elements controlled by the addition and removal of methylation marks, but have not been considered due to the field still being somewhat in its infancy.

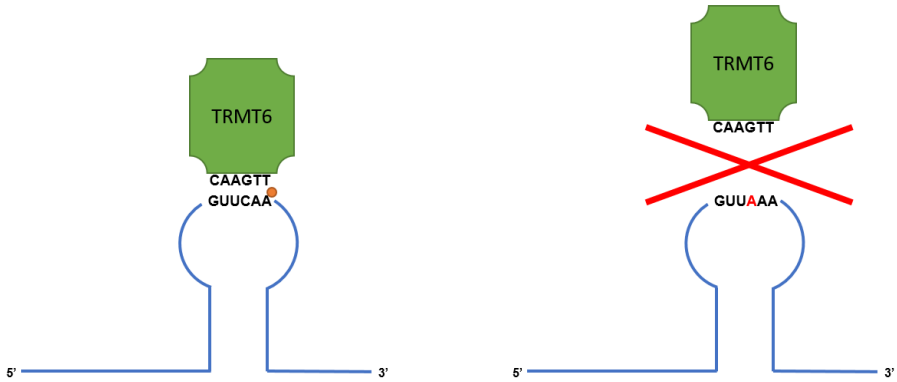


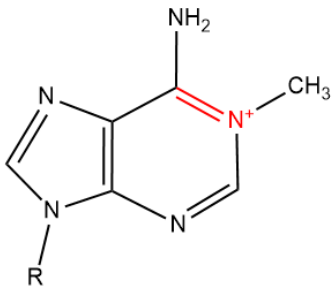
Figure 55. Theoretical example of silent and non-coding mutations disrupting recognition of short motifs in mRNA. Here a C>A mutation disrupts TRMT6 methyltransferase recognition and leaves typically methylated Adenosine unmethylated.

5.2 Thoughts on the Field of Epitranscriptomics

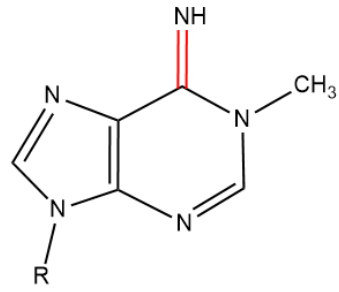
5.2.1 The Differing Depictions of N1-methyladenosine

There is currently a disparity in depictions of m1A depending on the publication. One version of the modification has a double bond to the 6' nitrogen, whereas the other depiction of m1A has a double bond between the 1' and 6' carbons of the of base, with a positive charge indicated at the 1' carbon bound by the methyl group. These differing depictions of m1A appear in modomics, pubchem, and typical depictions of m1A in dimroth rearrangement (Boccaletto et al., 2018; Kim et al., 2015; Segal et al., 1979). Biochemically, a double bond going up to the 6' nitrogen is a less chemically

stable structure than the other version, which forms a more stable benzene ring. It is possible that both of these forms potentially exist as resonance structures.



Dimroth Depiction



Modomics/Pubchem

Figure 56. The varying depictions of N1-methyladenosine. The primary different is the location of the double bond and the positive charge within one of the structures.

The less chemically stable neutral version of m1A is likely more prone to undergoing Dimroth rearrangement due to not containing a benzene ring in its structure. This would mean m6A existing at the same site as m1A and could be a potential explanation for FTO and the replication of the phenotype seen with ALKBH3.

Another potential importance of this difference occurs when m1A is depicted with a positive charge. The positively charged adenosine could potentially repel SRSF proteins involved in splicing and export. SRSF proteins, as implied by their name are Serine-Arginine rich. Serine and Arginine are positively charged amino acids and may therefore be repelled by the m1A if it contains a positive charge. If both versions of m1A, positive and neutral charge versions exist, this could cause different regulatory elements to be applied to the different versions. Transcripts containing more neutrally charged m1A would allow for the SRSF proteins to bind next to it unperturbed and carry out their normal splicing and export function. Positively charged m1A may skip the SRSF proteins, undergo a different export process and be more involved with

the cytoplasmic YTH Domain Family proteins which are involved in mRNA translation and degradation and recognize m1A.

5.2.2 Unreported methylation targets of ALKBH3 and FTO

A major qualm I have with my own results is why both ALKBH3 and FTO express the same phenotype when they have been reported to have different targets. There are indications from my own data that FTO is demethylating m1A, which would explain the overlap. This is additionally supported from a recent report by the He group that indicated m1A is being removed from tRNAs T-loop structure (Wei et al., 2018). This might provide additional credence to the idea that the GUUCRA site contained by the RNF168 mRNA is located on a T-loop secondary structure within the 3' UTR.

If FTO and ALKBH3 are indeed interacting as seen in my preliminary FLAG data that could mean that they are potentially in a complex. How they affect each others function is hard to decipher without further experiments. It could however be something similar to the methyltransferase complexes which have overlapping functions but are active under different conditions, an example potentially being varying mRNA folding structures.

Another potential explanation for the overlapping targets could be that ALKBH3 and FTO are acting on a separate modification altogether that has not been looked at yet. Perhaps both FTO and ALKBH3 are acting on the methyl mark at 2'-O-methyladenosine. As is the case with N6,2′-O-dimethyladenosine (m6Am), the m6A antibody recognized both m6a and m6Am. Perhaps this is also true for a modification called 1,2′-O-dimethyladenosine (m1Am). This modification has no known research conducted, but perhaps ALKBH3 recognizes it and is removing the modification from the 2'-O-methyladenosine position. It would be necessary to perform an assay similar to the one used in the Mauer paper in which they use Thin Layer Chromatography to check for the activity of FTO repairing m6Am rather than m6A. Apply the same method to ALKBH3 with m1A and m1Am, and perhaps even 2'-O on it's own (Mauer et al., 2016).

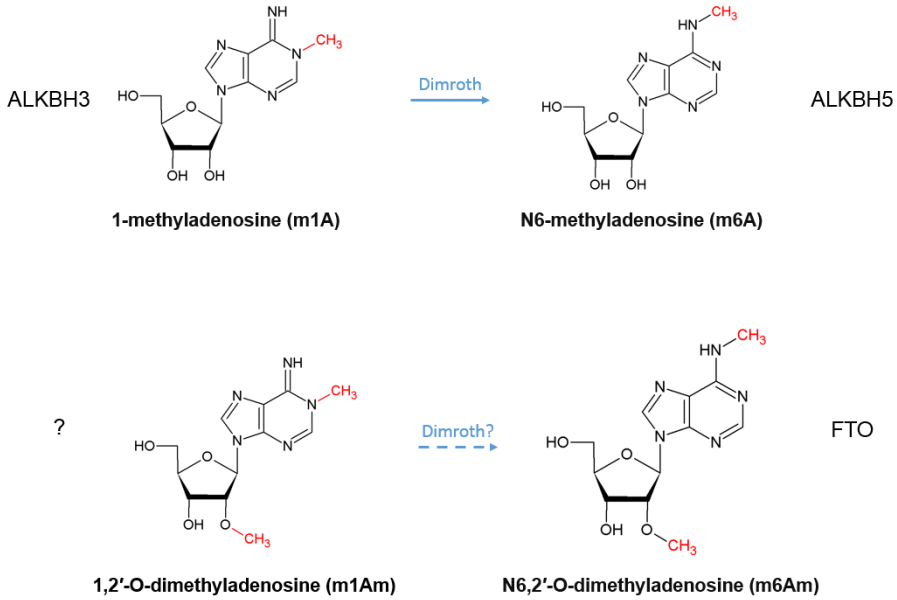
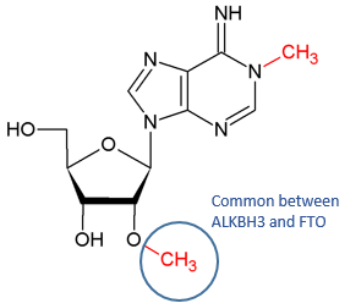


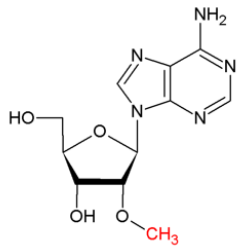
Figure 57. Structures of N1-methyladenosine, N6-methyladenosine, N1,2'-O-dimethyladenosine, and N6,2'-O-dimethyladenosine.

To the best of my knowledge I don't believe much, if any, research has been conducted on N1,2'-O-dimethyladenosine. It has certainly not yet been shown to exist in mRNA, though this is possibly due to LC/MS methods lacking a proper standard and not looking for this modification. While this is purely speculative, it is possible that this is an alternate modified nucleotide that explains the overlap in function we see with ALKBH3 and FTO. Though the 2018 publication from Dai et al. indicated FTO is capable of demethylating m1A from tRNA, it is therefore more likely that FTO has an ability to demethylate m1A on T-loop structures within mRNA.



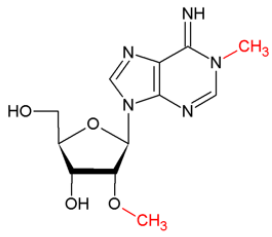
1,2′-O-dimethyladenosine (m1Am)

Figure 58. A common modification site on N1,2'-O-dimethyladenosine hypothetically linking ALKBH3 and FTO.

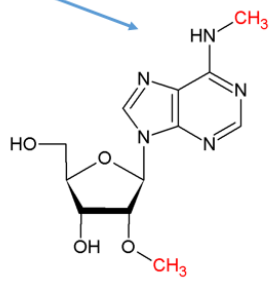


2′-O-methyladenosine

precursor



1,2′-O-dimethyladenosine (m1Am)



N6,2′-O-dimethyladenosine (m6Am)

Figure 59. 2'-O-methyladenosine as a precursor modified nucleotide prior to the addition of the methyl group to the the N1 or N6 positions.

5.2.3 Interaction of demethylases with methylation targets

ALKBH3 did not appear in a proteomics study of binding interactors of m1A, pointing in the direction that the interaction between ALKBH3 and m1A

modified mRNA is too transient to be detected in a proteomics study (Dai et al., 2018). This could have something to do with the method in which ALKBH3 removes methyl groups. ALKBH3 is a α-ketoglutarate dependent dioxygenase, which means that the repair mechanism involves oxidizing the methyl adduct, causing it to be destabilized and released along with formaldehyde (Sedgwick et al., 2004). ALKBH3 does not do direct reversal of alkylation damage by interacting with the modification in the way that proteins like MGMT do (Soll et al., 2017). This is likely why it is difficult to identify ALKBH3 as a binding interactor of m1A, because it is simply not binding to the modification.

5.2.4 Need for mRNA Folding Research

An obvious point I take from my own data is the need for further understanding of mRNA folding and increased ability to identify secondary structures. There are examples of proteins thought to target different types of RNA such as tRNA, that end up targeting mRNA due to the secondary structures that mimic another RNA subtype (Safra et al., 2017; Li et al., 2017). Recognition of consensus motifs could also be dependent on the secondary structures as seen with the GUUCRA motif and T-loops. This is likely more prevalent than we currently understand and both mRNA and DNA folding likely has a important role in allowing for these modifications to be added but also preventing them from occurring in certain situations.

5.2.5 Methylation Patterns as a Signal

Something I would be curious to see as the field develops further is whether or not there are specific patterns of methylation on mRNA that lead to specific functional mechanisms to be used. There is a biological precedent for such types of things seen in mechanisms like promoter methylation, where specific sites being methylated have more impact on mRNA expression. An example being the CpGs we looked at in this thesis and our paper (Stefansson et al., 2017). The idea of patterns for signals is not novel within in epitranscriptomics. The He group is looking into translational activation based on m6A profiles (Roundtree et al., 2017). The work is indicative of patterns being responsible for use of specific mechanisms, but still requires further understanding of how methylation is placed to be truly understood. This is also seen in epigenetics with histone acetylation and methylation having impacts on transcription rates (Kouzarides et al., 2007). When the field develops to a point where it can reliably map out modifications within a transcript it will be interesting to take

together the modifications and look for a pattern.

5.2.6 A Variation of meRIPseq to Identify Targeted Methylation Sites

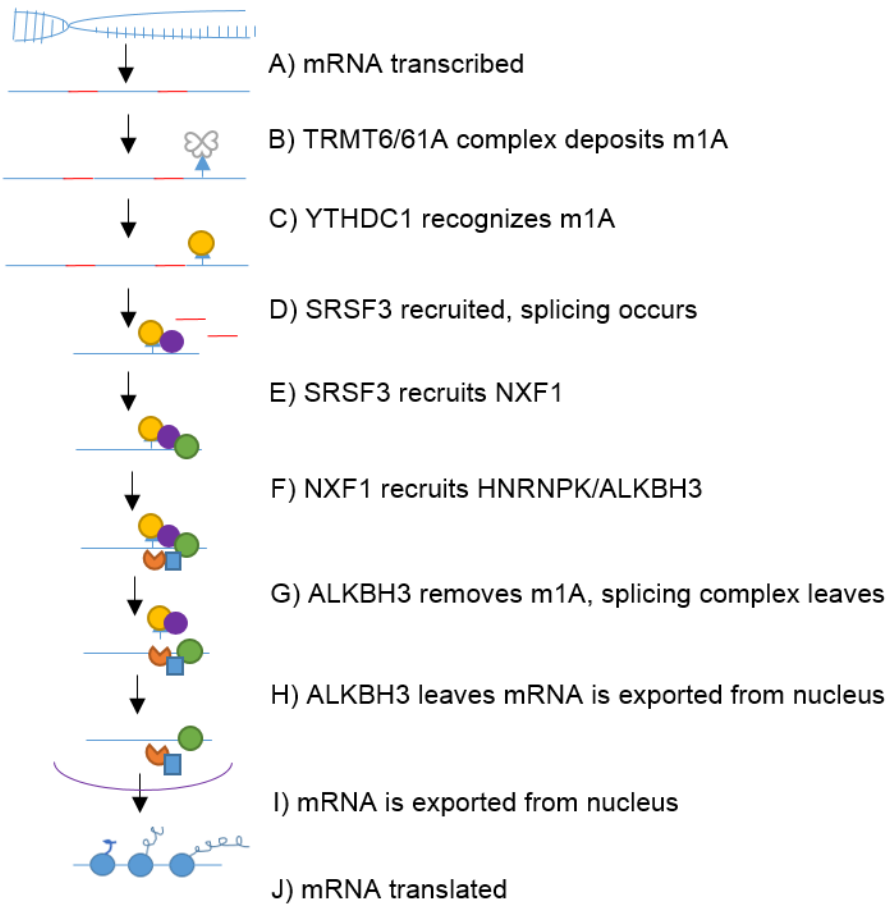
To understand what the consequences of these changes of expression might be we would have to identify additional targets for ALKBH3. Currently the field is refining down which RNAs are considered to be modified by m1A (Safra et al., 2017; Li et al., 2017; Li et al., 2016) However, RNF168 has not appeared on any of these lists, and our data shows a clear enrichment of RNF168 mRNA containing m1A following ALKBH3 depletion. This indicates the current methods the field is using to identify targets still need to be refined for increased sensitivity, as low expressing mRNAs such as RNF168 could be being missed in their current models. It is also likely that low-expressing mRNAs could be more sensitive to this type of regulation as any type of change in an already small pool of mRNA would have a more drastic impact on that genes expression.

A potential method to help identify targets of ALKBH3 would be to combine RNA-IP and RNA sequencing. This was already established in a sense with meRIP-seq. However, this would differ slightly from that protocol in the procedures and analysis. This would first involve preparing a large amount of lysate from a siControl sample and a siALKBH3 sample. The samples would then undergo fractionation to break apart the mRNA into smaller pieces. RNA-IP would then be performed on the fractionated samples. We would then use RNA sequencing on the fragments pulled down, which would give us an amount of reads for all the immunoprecipitated fragments and then align them to identify which mRNA is being pulled down, and given the randomness of the fragmentation we would have an idea where in the mRNA the modification is located. If we then compare the amounts of reads, we may potentially see an increase in certain fragments/mRNAs. Should there be an increase in reads of a particular location, that would be an indication that this is a site of a methylation mark targeted by ALKBH3. This method could also be applied to FTO, ALKBH5 and other demethylases targeting mRNA modifications, and expanded to other modifications. While untested, I believe variation of the meRIP-seq method could be a way to simultaneously identify specifically targeted sites at a high resolution as well as targeted mRNAs with increased sensitivity over mass sequencing all mRNA resulting from an RNA-IP.

5.3 Future Directions

5.3.1 A Potential Model for the Regulation of RNF168 mRNA by ALKBH3

I have thought of a potential way to fit all of the data beyond the export data together into a potential model. Much of the model is based off of educated guesses based on available data from other publications combined with preliminary data from my own project. If the project were to continue, I would likely perform experiments to check if my potential model has any merit. I will first show a representation of the model and then explain what indications there are to potentially explain each step.



A) There are no indications that ALKBH3 knockdown have any effect on the transcription of RNF168. Our own data indicates there are no changes to total levels of mRNA and this data is additionally backed up from a published microarray that indicated shRNA for ALKBH3 had no impact on mRNA levels of RNF168 (Liefke et al., 2015).

B) The TRMT6/TRMT61A/TRMT10C methyltransferase complex has been shown to be a methyl-transferase responsible for depositing m1A onto tRNAs as well as select mRNAs (Safra et al., 2017). We identified the established GUUCRA binding motif of this complex as being present in RNF168 and located in the 3' UTR. When we depleted the TRMT members we saw a decrease in the levels of RNF168 protein, but not SMC1 suggesting this is not a globally repressive event. This would provide some indication that the deposition of a methyl group is a crucial first step in the proper processing of pre-mRNA or mRNA of RNF168. This step has the most uncertainty due to not being 100% confident which modification is responsible for the phenotype due to the developing nature of the epitranscriptomic field. It is possible it is a seperate modification, likely m6A or m6Am in this step.

C) Steps B and C are dependent on each other. We saw that following knockdown of YTHDC1, there was a depletion of RNF168. YTHDC1 has been established as recognizing methylation marks, both m1A and m6A, as well as having a role in mediating nuclear export of mRNAs (Xu et al., 2014; Roundtree et al., 2017). More specifically RNF168 mRNA has been listed as a target of YTHDC1, providing additional support to this idea (Roundtree et al., 2017). The recognition of RNF168 mRNA would likely depend on the successful deposition of a methyl group from the prior step.

D) YTHDC1 has been shown to regulate splicing by the recognition of a methyl group by recruiting SRSF3 when present to promote exon-inclusive splicing and in its absence SRSF10 is recruited to promote exon skipping (Xiao et al., 2016). The Xiao paper identifies that the SRSF3 protein interacts with YTHDC1. Additionally, they identify that SRSF3 has a binding site nearby the modification site. We see a binding site corresponding to the established SRSF3 motif just downstream of the GUUCRA (Long et al., 2009; Li et al., 2017). The knockdown of SRSF3 additionally showed a decrease in RNF168 suggesting this protein is somehow needed, likely in the proper splicing of the mRNA.

E) The Xiao paper that identifies SRSF3 and YTHDC1 as interactors additionally identifies SRSF3 and NXF1 (Nuclear Export Factor 1) as interactors (Xiao et al., 2016). Interestingly there is no interaction between YTHDC1 and NXF1 despite the shared intermediate SRSF3. We have seen depletion of NXF1 and SRSF3 causes a reduction of RNF168 protein. A next step would be to perform RNA Scope following knockdown of these proteins to see if they replicate nuclear export phenotype.

F) This step is the most hypothetical of the model. NXF1 has been shown to be a interactor of HNRNPK by proteomics data available online from BioGRID (Chatr-aryamontri et al., 2017; Havugimana et al., 2010). HNRNPK is being singled out as it was a published interactor of ALKBH3 in the mass-spec performed by Dango et al. (Dango et al., 2011). HNRNPK is a RNA binding protein that has been implicated in nuclear export that has a consensus motif for binding of YCYYSCCM (Y = C or T, S = G or C, M = A or C) which is present on several locations with the RNF168 mRNA (Mikula et al., 2013). This step suggests that HNRNPK is bringing ALKBH3 along with it or recruiting it to the RNF168 mRNA. My results showed there is no reduction of RNF168 after HNRNPK knockdown, so this could suggest either HNRNPK is not essential or potentially part of a larger complex responsible for ALKBH3 recruitment that can compensate for the loss of HNRNPK. Other members of the HNRNP family have been shown to alter RNA structure in order to allow for methylation readers and writers to access the methylation (Liu et al., 2017). It is possible HNRNPK is acting in conjunction with other HNRNP proteins to perform such a function.

G) ALKBH3 has been shown to cleave m1A from mRNA. Our own data shows increased levels of m1A on RNF168 mRNA following ALKBH3 knockdown. The latter part of this step is theoretical but can be tested. I'm suggesting that the cleaving of m1A by ALKBH3 allows YTHDC1 to be released from the mRNA due to no longer having a binding substrate. The experiment to do this would be doing a Crosslinking-immunoprecipitation (CLIP) for YTHDC1 with and without knockdown of ALKBH3. We have ordered a FLAG-HA tagged YTHDC1 that can be pulled down to test this. If my theory is correct then following ALKBH3 knockdown we would expect to see increased levels of RNF168 mRNA bound to YTHDC1 indicating that it is still bound to the m1A. This step could potentially happen before NXF1 is recruited

which may explain why there is no interaction seen between YTHDC1 and NXF1 directly while they both bind SRSF3.

H) ALKBH3 has been shown to be 90% present in the nucleus, so it likely does not aid in the shuttling of the mRNA out of the nucleus. Also as an enzyme that performs its role through oxidative demethylation, there is no indication that ALKBH3 has anything more than a transient interaction with mRNAs. It could be possible to pull-down HNRNPK to see if we get binding of ALKBH3 and RNF168 mRNA at the same time, but I suspect this process is very quick. NXF1 likely remains on the mRNA to complete the export process.

I) The inhibition of nuclear export following ALKBH3 or FTO depletion has been the primary finding of this thesis. We saw both with RNA Scope and nuclear fractionation that this process was inhibited following ALKBH3 and FTO knockdown. Assuming ALKBH3 or FTO are unchanged nuclear export should proceed as normal. However to prove NXF1 is involved in this process knockdown of NXF1 and confirmation of replicated phenotype needs to be confirmed through western blot, RNA Scope and fractionation.

J) Provided this model and this thesis have some validity, mRNA should be translated regularly following export. There is a possibility of some involvement of the YTH Domain Family proteins given their recognition of m1A. If in some cases m1A persists on the mRNA this may cause increased nuclear decay by P-bodies if recognized by YTHDF2, or potentially increase translation if recognized by YTHDF1 or YTHDF3. It would be interesting to test if P-body inhibitors had any impact on the levels of RNF168 both with and without ALKBH3/FTO knockdown.

5.3.2 Intermediate protein between ALKBH3 and RNF168 being regulated by m1A control

One thing I've been curious about is if ALKBH3 is acting solely on RNF168 mRNA or if there is an additional intermediate or tertiary protein controlled by m1A regulation that is involved in the process. Using the Stefan Decoder program I created along with a friend I was able to quickly check if mRNA of various proteins contained the TRMT6/61A consensus motif of GUUCRA. Two proteins involved in mRNA export, SRSF3 and YTHDC1 both contained the consensus motif. Similar to the RNF168 mRNA the motif (or motifs in the case of SRSF3) was found in the final exon and within the 3' UTR. What is

additionally interested about these proteins is that they were determined by Dai to be m1A binding proteins (Dai et al., 2018). Perhaps future investigations into this regulation should examine if there is any impact of the levels of these proteins in the presence/absence of ALKBH3 or FTO. I've included them as part of my potential model because I believe there is evidence in the literature for their interaction with the RNF168 mRNA and when combined with my data there is a logical connection between RNF168 regulation by m1A.

Another protein that I think is worth checking as an intermediary is SRSF6. This protein has come up very often in m1A papers as being modified and targeted by proteins such as YTHDC1. SRSF6 itself has been implied in the determination of alternative splicing and could possibly have a role in the regulation of RNF168. There is however no current indication that it is involved in nuclear export so it could just be a side player in the process described in this thesis.

5.3.3 Control of m1A Levels Following Induction of Stress

An obvious next step for our research group would be to look at changes seen after the induction of DNA damage. There is great reason to believe that methylation and demethylation events happen in a very dynamic manner, so it would be interesting to look at a protein such a RNF168 which involved in DNA damage repair. Treatment with a damage inducing reagent such as neocarzinostatin and repeating the experiments done in this thesis could provide insight into just how dynamic this regulation is. The question still remains if the methylation mark is helpful or detrimental for increased expression. My data suggests that increased methylation, particularly for RNF168, has a repressive effect on the protein levels. However this is only observing it in normal conditions, and perhaps other elements of the pathway are altered to be more or less altered in the event of something like DNA damage.

5.3.4 CRISPR for ALKBH3

A future experiment would be performing CRISPr to knockout ALKBH3 and see if we get total depletion of RNF168. With siRNAs we never saw a complete depletion of RNF168 protein , but siRNAs are never 100% efficient in their knockdown capabilities. Disabling the gene would give a clearer impact into what the absence of ALKBH3 means for the cell. It would be additionally

interesting to knockout both ALKBH3 and FTO in the same cell line. We never saw any additive effect of the co-depletion so it is uncertain if this would have any additional impact on RNF168. It would be interesting to see how, if at all, viable the ALKBH3 -/- cell line would be. Our tumor and cell line data suggests that cell lines not expressing ALKBH3 are indeed viable.

Another interesting CRISPR to perform would be mutating our proposed GUUCRA binding site for TRMT6/TRMT61A in the RNF168 mRNA 3' UTR. Our preliminary data suggests that the TRMT complex has some function in the regulation of RNF168 and my theoretical model proposes that initial deposition of a methylation mark is required for successful splicing and then export. CRISPR for many of the binding motifs identified would be an interesting series of experiments but would likely require more preliminary data prior to performing the CRISPR.

6 Conclusions

ALKBH3 is doing something interesting. This thesis is likely only beginning to scratch the surface of the complex regulation of DNA, RNA, and proteins that the ALKB family enzymes are involved in. While epigenetics as a field is further along, the vast majority of epitranscriptomic modifications are currently poorly characterized and their impacts not well understood. The need for creativity and originality while maintaining academic diligence and rigor make this a very exiciting time to be a researcher. I believe that as the understanding of epigenetic and epitranscriptomic regulation grows, it will be accompanied with improved identification, comprehension, and treatment of diseases. The results of this thesis suggest that the downregulation of ALKBH3 will affect RNF168 expression, and thereby affect a cell‘s ability to perform efficient and accurate DNA repair, an event with many implications in the formation of a disease. I am personally excited to see what the future of this field will be.

References

Aas, P. A., Otterlei, M., Falnes, P. O., Vågbø, C. B., Skorpen, F., Akbari, M., … Krokan, H. E. (2003). Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature*, *421*(6925), 859–863. http://doi.org/10.1038/nature01363

Alarcón, C. R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., & Tavazoie, S. F. (2015). HNRNPA2B1 Is a Mediator of m6A-Dependent Nuclear RNA Processing Events. *Cell*, *162*(6), 1299–1308. http://doi.org/10.1016/j.cell.2015.08.011

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., … Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature*, *500*, 415–422. http://doi.org/10.1038/nature12477

Arnoult, N., & Karlseder, J. (2015). Complex interactions between the DNA-damage response and mammalian telomeres. Nature Structural and Molecular Biology, 22(11), 859–866. http://doi.org/10.1038/nsmb.3092

Ashworth, A. (2008). A synthetic lethal therapeutic approach: Poly(ADP) ribose polymerase inhibitors for the treatment of cancers deficient in DNA double-strand break repair. *Journal of Clinical Oncology*, *26*(22), 3785–3790. http://doi.org/10.1200/JCO.2008.16.0812

Assenov, Y., Brocks, D., & Gerhäuser, C. (2018). Intratumor heterogeneity in epigenetic patterns. *Seminars in Cancer Biology*, *51*(September 2017), 12–21. http://doi.org/10.1016/j.semcancer.2018.01.010

Bartocci, C., & Denchi, E. L. (2013). Put a RING on it: Regulation and inhibition of RNF8 and RNF168 RING finger E3 ligases at DNA damage sites. *Frontiers in Genetics*, *4*(JUL), 1–12. http://doi.org/10.3389/fgene.2013.00128

Baylin, S. B., & Jones, P. A. (2011). T H E N E X T 1 0 Y E A R S — T I M E L I N E A decade of exploring the cancer epigenome — biological and translational implications. *Nature Reviews Cancer*. Retrieved from http://dx.doi.org/10.1038/nrc3130

Bennardo, N., Cheng, A., Huang, N., & Stark, J. M. (2008). Alternative-NHEJ Is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair, *4*(6). http://doi.org/10.1371/journal.pgen.1000110

Beucher, A., Birraux, J., Tchouandong, L., Barton, O., Shibata, A., Conrad, S., … Lobrich, M. (2009). ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. The EMBO Journal, 28(August), 3413–3427. http://doi.org/10.1038/emboj.2009.276

Boccaletto, P., Machnicka, M. A., Purta, E., Wirecki, T. K., Ross, R., Limbach,

A., … Bujnicki, J. M. (2018). MODOMICS : a database of RNA modification pathways . 2017 update. *Nucleic Acids Research*, *46*(November 2017), 303–307. http://doi.org/10.1093/nar/gkx1030

Bodell, W. J., & Singer, B. (1979). Influence of Hydrogen Bonding in DNA and Polynucleotides on Reaction of Nitrogens and Oxygens toward Ethylnitrosourea. *Biochemistry*, *18*(13), 2860–2863. http://doi.org/10.1021/bi00580a029

Boersma, V., Moatti, N., Segura-Bayona, S., Peuscher, M. H., Van Der Torre, J., Wevers, B. A., … Jacobs, J. J. L. (2015). MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5′ end resection. *Nature*, *521*(7553), 537–540. http://doi.org/10.1038/nature14216

Bohgaki, T., Bohgaki, M., Cardoso, R., Panier, S., Zeegers, D., Li, L., … Hakem, R. (2011). Genomic Instability, Defective Spermatogenesis, Immunodeficiency, and Cancer in a Mouse Model of the RIDDLE Syndrome. *PLoS Genetics*, *7*(4), e1001381. http://doi.org/10.1371/journal.pgen.1001381

Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., … Helleday, T. (2005). Specific killing of BRCA2-deficient tumours with inhibitors of poly (ADP-ribose) polymerase. Nature, 434(May), 913–917. http://doi.org/10.1038/nature05789

Bunting, S. F., Callén, E., Wong, N., Chen, H. T., Polato, F., Gunn, A., … Nussenzweig, A. (2010). 53BP1 inhibits homologous recombination in brca1-deficient cells by blocking resection of DNA breaks. *Cell*, *141*(2), 243–254. http://doi.org/10.1016/j.cell.2010.03.012

Callen, E., Di Virgilio, M., Kruhlak, M. J., Nieto-Soler, M., Wong, N., Chen, H. T., … Nussenzweig, A. (2013). 53BP1 mediates productive and mutagenic DNA repair through distinct phosphoprotein interactions. *Cell*, *153*(6), 1266–1280. http://doi.org/10.1016/j.cell.2013.05.023

Carmeliet, P. (2005). VEGF as a Key Mediator of Angiogenesis in Cancer. *Oncology*, *69*(suppl 3), 4–10. http://doi.org/10.1159/000088478

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Arman, B., … Schultz, N. (2012). The cBio Cancer Genomics Portal : An Open Platform for Exploring Multidimensional Cancer Genomics Data. American Association for Cancer Research, (May). http://doi.org/10.1158/2159-8290.CD-12-0095

Celeste, A., Petersen, S., Romanienko, P. J., Fernandez-capetillo, O., Chen, H. T., Livak, F., … Bonner, W. M. (2002). Genomic Instability in Mice Lacking Histone H2AX. *Science*, *296*(May), 922–928.

Cells, H., Desrosiers, R., Friderici, K., & Rottman, F. (1974). Identification of Methylated Nucleosides in Messenger RNA from Novikoff Hepatoma

Cells. *Proceedings of the National Academy of Sciences*, *71*(10), 3971–3975.

Chang, H. H. Y., Pannunzio, N. R., Adachi, N., & Lieber, M. R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology*. http://doi.org/10.1038/nrm.2017.48

Chapman, J. R., Barral, P., Vannier, J., Steger, M., Tomas-loba, A., Sartori, A. A., … Boulton, S. J. (2013). Article RIF1 Is Essential for 53BP1-Dependent Nonhomologous End Joining and Suppression of DNA Double-Strand Break Resection. *Cell*, (49), 858–871. http://doi.org/10.1016/j.molcel.2013.01.002

Chapman, J. R., Taylor, M. R. G., & Boulton, S. J. (2012). Playing the End Game : DNA Double-Strand Break Repair Pathway Choice. Cell, 47(4), 497–510. http://doi.org/10.1016/j.molcel.2012.07.029

Chasseval, D., Barraud, A., Fondane, M., Buck, D., Malivert, L., Hufnagel, M., … Ste, J. (2006). Cernunnos , a Novel Nonhomologous End-Joining Factor , Is Mutated in Human Immunodeficiency with Microcephaly. Cell, 287–299. http://doi.org/10.1016/j.cell.2005.12.030

Chatr-aryamontri, A., Oughtred, R., Boucher, L., Rust, J., Chang, C., Kolas, N. K., … Montr, D. (2017). The BioGRID interaction database : 2017 update. *Nucleic Acids Research*, 1–11. http://doi.org/10.1093/nar/gkw1102

Chen, H., Lisby, M., & Symington, L. S. (2013). Article RPA Coordinates DNA End Resection and Prevents Formation of DNA Hairpins. *Molecular Cell*, *50*(4), 589–600. http://doi.org/10.1016/j.molcel.2013.04.032

Choi, M., Kipps, T., & Kurzrock, R. (2016). ATM Mutations in Cancer: Therapeutic Implications. *Molecular Cancer Therapeutics*, *15*(8), 1781–1791. http://doi.org/10.1158/1535-7163.MCT-15-0945

Chu, X., Erdman, R., Susek, M., Gerst, H., Derr, K., Al-Agha, M., … Gerhard, G. S. (2008). Association of Morbid Obesity With FTO and INSIG2 Allelic Variants. *JAMA Surgery*, *143*(3), 235–240.

Collins, N., & Wooster, R. (1997). Absence of methylation of CpG dinucleotides within the promoter of the breast cancer susceptibility gene BRCA2 in normal tissues and in breast and ovarian cancers. *British Journal of Cancer*, *76*(9), 1150–1156.

Coots, R. A., Liu, X. M., Mao, Y., Dong, L., Zhou, J., Wan, J., … Qian, S. B. (2017). M6A Facilitates eIF4F-Independent mRNA Translation. *Molecular Cell*, 1–11. http://doi.org/10.1016/j.molcel.2017.10.002

Couch, F. J., Nathanson, K. L., & Offit, K. (2014). Two decades after BRCA: Setting paradigms in personalized cancer care and prevention. *Science*, *343*(6178), 1466–1470. http://doi.org/10.1126/science.1251827

Cunningham, J. M., Cicek, M. S., Larson, N. B., Davila, J., Wang, C., Larson, M. C., … Goode, E. L. (2014). Clinical characteristics of ovarian cancer classified by BRCA1, BRCA2, and RAD51C status. *Scientific Reports*, *4*, 1–7. http://doi.org/10.1038/srep04026

Cutter, A. R., & Hayes, J. J. (2015). A brief review of nucleosome structure. *FEBS Letters*, *589*(20), 2914–2922. http://doi.org/10.1016/j.febslet.2015.05.016

Dai, X., Wang, T., Gonzalez, G., & Wang, Y. (2018). Identification of YTH Domain-Containing Proteins as the Readers for N1-Methyladenosine in RNA. *Analytical Chemistry*. http://doi.org/10.1021/acs.analchem.8b01703

Daley, J. M., & Sung, P. (2013). Previews RIF1 in DNA Break Repair Pathway Choice. *MOLCEL*, *49*(5), 840–841. http://doi.org/10.1016/j.molcel.2013.02.019

Daley, J. M., & Sung, P. (2014). 53BP1, BRCA1, and the Choice between Recombination and End Joining at DNA Double-Strand Breaks. *Molecular and Cellular Biology*, *34*(8), 1380–1388. http://doi.org/10.1128/MCB.01639-13

Dango, S., Mosammaparast, N., Sowa, M. E., Xiong, L. J., Wu, F., Park, K., … Shi, Y. (2011). DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. *Molecular Cell*, *44*(3), 373–384. http://doi.org/10.1016/j.molcel.2011.08.039

Dedeurwaerder, S., Desmedt, C., Calonne, E., Singhal, S. K., Haibe-Kains, B., Defrance, M., … Fuks, F. (2011). DNA methylation profiling reveals a predominant immune component in breast cancers. EMBO Molecular Medicine, 3(12), 726–741. http://doi.org/10.1002/emmm.201100801

Devgan, S. S., Sanal, O., Doil, C., Nakamura, K., Nahas, S. A., Pettijohn, K., … Gatti, R. A. (2011). Homozygous deficiency of ubiquitin-ligase ring-finger protein RNF168 mimics the radiosensitivity syndrome of ataxia-telangiectasia. *Cell Death and Differentiation*, *18*(9), 1500–1506. http://doi.org/10.1038/cdd.2011.18

Dinglay, S., Trewick, S. C., Lindahl, T., & Sedgwick, B. (2000). Defective processing of methylated single-stranded DNA by E. coli alkB mutants. *Genes and Development*, *14*(16), 2097–2105. http://doi.org/10.1101/gad.14.16.2097

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D. H., Pepperkok, R., … Lukas, C. (2009). RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. *Cell*, *136*(3), 435–446. http://doi.org/10.1016/j.cell.2008.12.041

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., … Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, *485*(7397), 201–206. http://doi.org/10.1038/nature11112

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., … He, C. (2016). The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature*, *530*(7591), 441–6. http://doi.org/10.1038/nature16998

Dominissini, D., & Rechavi, G. (2017). Loud and Clear Epitranscriptomic m1A Signals: Now in Single-Base Resolution. *Molecular Cell*, *68*(5), 825–826. http://doi.org/10.1016/j.molcel.2017.11.029

Drabløs, F., Feyzi, E., Aas, P. A., Vaagbø, C. B., Kavli, B., Bratlie, M. S., … Krokan, H. E. (2004). Alkylation damage in DNA and RNA — repair mechanisms and medical significance. *DNA Repair*, *3*, 1389–1407. http://doi.org/10.1016/j.dnarep.2004.05.004

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., … Wu, L. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nature Communications*, *7*, 1–11. http://doi.org/10.1038/ncomms12626

Duncan, T., Trewick, S. C., Koivisto, P., Bates, P. A., Lindahl, T., & Sedgwick, B. (2002). Reversal of DNA alkylation damage by two human dioxygenases. *PNAS*.

Edo, K., & Koide, Y. (1997). Neocarzinostatin Chromophore: Structure and Mechanism of DNA Cleavage. In *Neocarzinostatin* (pp. 23–45).

Engel, M., & Chen, A. (2017). The emerging role of mRNA methylation in normal and pathological behavior. *Genes, Brain and Behavior*, (October), 1–9. http://doi.org/10.1111/gbb.12428

Esteller, M., Garcia-Foncillas, J., Andion, E., Goodman, S. N., Hidalgo, O. F., Vanaclocha, V., … Herman, J. G. (2000). INACTIVATION OF THE DNA-REPAIR GENE MGMT AND THE CLINICAL RESPONSE OF GLIOMAS TO ALKYLATING AGENTS. *The New England Journal of Medicine*, *343*(19), 1350–1354.

Farmer, H., Mccabe, N., Lord, C. J., Tutt, A., Johnson, D., Richardson, T., … Ashworth, A. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature, 239(1991), 236–239.

Feng, L., Fong, K. W., Wang, J., Wang, W., & Chen, J. (2013). RIF1 counteracts BRCA1-mediated end resection during DNA repair. Journal of Biological Chemistry, 288(16), 11135–11143. http://doi.org/10.1074/jbc.M113.457440

Fumagalli, C., Pruneri, G., Possanzini, P., Manzotti, M., Barile, M., Feroce, I., … Bonanni, B. (2012). Methylation of O 6 -methylguanine-DNA

methyltransferase (MGMT) promoter gene in triple-negative breast cancer patients. *Breast Cancer Res Treat*, 131–137. http://doi.org/10.1007/s10549-011-1945-9

Greenberg, R. A. (2018). Assembling a protective shield. Nature Cell Biology, 68(July), 13–14.

Grozhik, A. V., & Jaffrey, S. R. (2017). Epitranscriptomics: Shrinking maps of RNA modifications. *Nature*. http://doi.org/10.1038/nature24156

Grozhik, A. V, & Jaffrey, S. R. (2018). Distinguishing RNA modifications from noise in epitranscriptome maps. *Nature Chemical Biology*, *14*(3), 215–225. http://doi.org/10.1038/nchembio.2546

Gudjonsson, T., Altmeyer, M., Savic, V., Toledo, L., Dinant, C., Grøfte, M., … Lukas, C. (2014). Spreading of Chromatin Ubiquitylation at Damaged Chromosomes. *Cell*, *159*(6), 1476–1477. http://doi.org/10.1016/j.cell.2014.11.027

Harper, J. W., & Elledge, S. J. (2007). The DNA Damage Response : Ten Years After. Cell, 28, 739–745. http://doi.org/10.1016/j.molcel.2007.11.015

Harrigan, J. A., Belotserkovskaya, R., Coates, J., Dimitrova, D. S., Polo, S. E., Bradshaw, C. R., … Jackson, S. P. (2011). Replication stress induces 53BP1-containing OPT domains in G1 cells. Journal of Cell Biology, 193(1), 97–108. http://doi.org/10.1083/jcb.201011083

Haussmann, I. U., Bodi, Z., Sanchez-Moran, E., Mongan, N. P., Archer, N., Fray, R. G., & Soller, M. (2016). M6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. *Nature*, *540*(7632), 301–304. http://doi.org/10.1038/nature20577

Havugimana, P. C., Hart, G. T., Nepusz, T., Yang, H., Turinsky, A. L., Li, Z., … Emili, A. (2010). Resource A Census of Human Soluble Protein Complexes. *Cell*, (150), 1068–1081. http://doi.org/10.1016/j.cell.2012.08.011

Hustedt, N., & Durocher, D. (2017). The control of DNA repair by the cell cycle. *Nat Cell Biol*, *19*(1), 1–9. http://doi.org/10.1038/ncb3452

Janssens, S., & Tschopp, J. (2006). Signals from within: The DNA-damage-induced NF-κB response. *Cell Death and Differentiation*, *13*(5), 773–784. http://doi.org/10.1038/sj.cdd.4401843

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., … He, C. (2011). N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chemical Biology*, *7*(12), 885–887. http://doi.org/10.1038/nchembio.687

Johannessen, T. A., Prestegarden, L., Grudic, A., Hegi, M. E., Tysnes, B. B., & Bjerkvig, R. (2013). The DNA repair protein ALKBH2 mediates

temozolomide resistance in human. *Neuro-Oncology*, *15*(3), 269–278.

Jonasson, J. G., Stefansson, O. A., Johannsson, O. T., Sigurdsson, H., Agnarsson, B. A., Olafsdottir, G. H., … Tryggvadottir, L. (2016). Oestrogen receptor status, treatment and breast cancer prognosis in Icelandic BRCA2 mutation carriers. *British Journal of Cancer*, *115*(7), 776–783. http://doi.org/10.1038/bjc.2016.249

Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, *13*(7), 484–492. http://doi.org/10.1038/nrg3230

Fernandez-vidal, A., Vignard, J., & Mirey, G. (2017). Around and beyond 53BP1 Nuclear Bodies. International Journal of Molecular Sciences, 18(2611), 1–20. http://doi.org/10.3390/ijms18122611

Kass, E. M., & Jasin, M. (2010). Collaboration and competition between DNA double-strand break repair pathways. *FEBS Letters*, *584*(17), 3703–3708. http://doi.org/10.1016/j.febslet.2010.07.057

Kauffmann, A., Rosselli, F., Lazar, V., Winnepenninckx, V., Mansuet-Lupo, A., Dessen, P., … Sarasin, A. (2008). High expression of DNA repair pathways is associated with metastasis in melanoma patients. *Oncogene*, *27*(5), 565–573. http://doi.org/10.1038/sj.onc.1210700

Knijnenburg, T. A., Wang, L., Zimmermann, M. T., Chambwe, N., Gao, G. F., Cherniack, A. D., … Wang, C. (2018). Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. Cell Reports, 23(1), 239–254.e6. http://doi.org/10.1016/j.celrep.2018.03.076

Koboldt, D. C., Fulton, R. S., McLellan, M. D., Schmidt, H., Kalicki-Veizer, J., McMichael, J. F., … Palchik, J. D. (2012). Comprehensive molecular portraits of human breast tumours. Nature, 490(7418), 61–70. http://doi.org/10.1038/nature11412

Kong, X., Mohanty, S. K., Stephens, J., Heale, J. T., Gomez-Godinez, V., Shi, L. Z., … Berns, M. W. (2009). Comparative analysis of different laser systems to study cellular responses to DNA damage in mammalian cells. *Nucleic Acids Research*, *37*(9). http://doi.org/10.1093/nar/gkp221

Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell*, *128*(4), 693–705. http://doi.org/10.1016/j.cell.2007.02.005

Kozbial, P. Z., & Mushegian, A. R. (2005). Natural history of S-adenosylmethionine-binding proteins. *BMC Structural Biology*, *5*. http://doi.org/10.1186/1472-6807-5-19

Krokan, H. E., & Bjoras, M. (2013). Base Excision Repair. *Cold Spring Harbor Perspectives in Biology*, *6*, 6–11. http://doi.org/10.1016/S0065-3233(04)69001-2

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., … Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. Nature, 518(7539), 317–329. http://doi.org/10.1038/nature14248

Lalloo, F., & Dg, E. (2012). Familial breast cancer. *Clinical Genetics*, *1994*(2), 105–114. http://doi.org/10.1111/j.1399-0004.2012.01859.x

Larsen, D. H., Poinsignon, C., Gudjonsson, T., Dinant, C., Payne, M. R., Hari, F. J., … Lukas, J. (2010). The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage. *Journal of Cell Biology*, *190*(5), 731–740. http://doi.org/10.1083/jcb.200912135

Lee, D. H., & Goldberg, A. (1998). Proteasome inhibitors: Valuable new tools for cell biologists. *Trends in Cell Biology, 8*(10), 397–403. http://doi.org/10.1016/S0962-8924(98)01346-4

Lee, S. E. (2014). Single-Strand Annealing. *Molecular Life Sciences*, *10*, 3–6. http://doi.org/10.1007/978-1-4614-6436-5

Li, M., Zhao, X., Wang, W., Shi, H., Pan, Q., Lu, Z., … Klungland, A. (2018). Ythdf2-mediated m6A mRNA clearance modulates neural development in mice. *Genome Biology*, *19*(1), 1–16. http://doi.org/10.1186/s13059-018-1436-y

Li, X., Peng, J., Yi, C., & N, T. M. (2017). Transcriptome-Wide Mapping of N1-Methyladenosine Methylome. In A. Lusser (Ed.), *RNA Methylation: Methods and Protocols* (Vol. 1562, pp. 245–255). New York, NY: Springer New York. http://doi.org/10.1007/978-1-4939-6807-7_16

Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., & Yi, C. (2016). Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nature Chemical Biology*, *12*(5), 311–6. http://doi.org/10.1038/nchembio.2040

Liao, S., Sun, H., & Xu, C. (2018). YTH Domain: A Family of N6-methyladenosine (m6A) Readers. *Genomics, Proteomics and Bioinformatics*, *16*(2), 99–107. http://doi.org/10.1016/j.gpb.2018.04.002

Liefke, R., Windhof-Jaidhauser, I. M., Gaedcke, J., Salinas-Riester, G., Wu, F., Ghadimi, M., & Dango, S. (2015). The oxidative demethylase ALKBH3 marks hyperactive gene promoters in human cancer cells. *Genome Medicine*, *7*(1), 66. http://doi.org/10.1186/s13073-015-0180-0

Linder, B., Grozhik, A. V, Olarerin-george, A. O., Meydan, C., Mason, C. E., & Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nature Methods*, (May). http://doi.org/10.1038/nmeth.3453

Linder, B., Grozhik, A. V, Olarerin-george, A. O., Meydan, C., Mason, C. E., & Jaffrey, S. R. (2016). Single-nucleotide resolution mapping of m6A and

m6Am throughout the transcriptome. *Nature Methods*, *12*(8), 767–772. http://doi.org/10.1038/nmeth.3453.Single-nucleotide

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., … He, C. (n.d.). A METTL3-METTL14 complex mediates mammalian nuclear RNA N 6 - adenosine methylation. http://doi.org/10.1038/nchembio.1432

Liu, N., Zhou, K. I., Parisien, M., Dai, Q., Diatchenko, L., & Pan, T. (2017). N 6 -methyladenosine alters RNA structure to regulate binding of a low-complexity protein. Nucleic Acids Research, 45(10), 6051–6063. http://doi.org/10.1093/nar/gkx141

Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., … Kawaji, H. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biology*, *16*(1), 1–14. http://doi.org/10.1186/s13059-014-0560-6

Lord, C. J., & Ashworth, A. (2012). The DNA damage response and cancer therapy. *Nature*, *481*(7381), 287–294. http://doi.org/10.1038/nature10760

Lord, C. J., & Ashworth, A. (2016). BRCAness revisited. *Nature Publishing Group*, *16*(2), 110–120. http://doi.org/10.1038/nrc.2015.21

Luijsterburg, M. S., & Attikum, H. Van. (n.d.). Close encounters of the RNF8th kind : when chromatin meets DNA repair. *Current Opinion in Cell Biology*, *24*(3), 439–447. http://doi.org/10.1016/j.ceb.2012.03.008

Lukas, C., Savic, V., Bekker-jensen, S., Doil, C., Neumann, B., Pedersen, R. S., … Lukas, J. (2011). 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nature Cell Biology, 13(3), 243–253. http://doi.org/10.1038/ncb2201

Lukas, J., Lukas, C., & Bartek, J. (2011). More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance. *Nature Cell Biology*, *13*(10), 1161–1169. http://doi.org/10.1038/ncb2344

Macon, J. B., & Wolfenden, R. (1968). 1-Methyladenosine. Dimroth Rearrangement and Reversible Reduction. *Biochemistry*, *7*(10), 3453–3458. http://doi.org/10.1021/bi00850a021

Mailand, N., Bekker-jensen, S., Faustrup, H., Melander, F., Bartek, J., & Lukas, C. (2007). RNF8 Ubiquitylates Histones at DNA Double-Strand Breaks and Promotes Assembly of Repair Proteins. *Cell*, (131), 887–900. http://doi.org/10.1016/j.cell.2007.09.040

Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., … Jaffrey, S. R. (2016). Reversible methylation of m6Am in the 5′ cap controls mRNA stability. *Nature*, *541*(7637), 371–375. http://doi.org/10.1038/nature21022

Mcvey, M., & Lee, S. E. (2008). MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. Trends Genet, 24(11), 529–538. http://doi.org/10.1016/j.tig.2008.08.007.MMEJ

Meyer, K. D., & Jaffrey, S. R. (2017). Rethinking m A Readers , Writers , and Erasers. *Annual Review of Cell and Developmental Biology*, *33*(July), 1–24.

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., & Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell*, *149*(7), 1635–1646. http://doi.org/10.1016/j.cell.2012.05.003

Mikula, M., Bomsztyk, K., Goryca, K., Chojnowski, K., & Ostrowski, J. (2013). Heterogeneous Nuclear Ribonucleoprotein (HnRNP) K Genome-wide Binding Survey Reveals Its Role in Regulating 3 □ -End RNA Processing and Transcription Termination at the Early Growth Response 1 (EGR1) Gene through XRN2. *The Journal of Biological Chemistry*, *288*(34), 24788–24798. http://doi.org/10.1074/jbc.M113.496679

Muggerud, A. A., Rønneberg, J. A., Wärnberg, F., Botling, J., Busato, F., Jovanovic, J., … Tost, J. (2010). FOXC1 , PPP2R2B and PTEN in ductal carcinoma in situ and early invasive breast cancer, 1–10.

Noordermeer, S. M., Adam, S., Setiaputra, D., Barazas, M., Pettitt, S. J., Ling, A. K., … Durocher, D. (2018). The shieldin complex mediates 53BP1-dependent DNA repair. *Nature*, *560*(7716), 117–121. http://doi.org/10.1038/s41586-018-0340-7

O'Kane, G. M., Connor, A. A., & Gallinger, S. (2017). Characterization, Detection, and Treatment Approaches for Homologous Recombination Deficiency in Cancer. *Trends in Molecular Medicine*, *23*(12), 1121–1137. http://doi.org/10.1016/j.molmed.2017.10.007

Ochs, F., Somyajit, K., Altmeyer, M., Rask, M.-B., Lukas, J., & Lukas, C. (2016). 53BP1 fosters fidelity of homology-directed DNA repair. *Nature Structural & Molecular Biology*, *23*(8), 714–21. http://doi.org/10.1038/nsmb.3251

Okamoto, K., Bartocci, C., Ouzounov, I., Diedrich, J. K., Iii, J. R. Y., & Denchi, E. L. (2013). A two-step mechanism for TRF2-mediated chromosome-end protection. Nature, 494(7438), 502–505. http://doi.org/10.1038/nature11873

Panier, S., & Boulton, S. J. (2013). Double-strand break repair: 53BP1 comes into focus. Nature, 15(1), 7–18. http://doi.org/10.1038/nrm3719

Pathania, S., Bade, S., Guillou, M. Le, Burke, K., Reed, R., Bowman-colin, C., … Livingston, D. M. (2014). BRCA1 haploinsufficiency for replication stress suppression in primary cells. *Nature Communications*, *5*, 1–15. http://doi.org/10.1038/ncomms6496

Patrono, C., Sterpone, S., Testa, A., & Cozzi, R. (2014). Polymorphisms in base excision repair genes: Breast cancer risk and individual radiosensitivity. *World Journal of Clinical Oncology*, *5*(5), 874. http://doi.org/10.5306/wjco.v5.i5.874

Paull, T. T., & Lee, J.-H. (2005). The Mre11/Rad50/Nbs1 Complex and Its Role as a DNA Double-Strand Break Sensor for ATM. *Cell Cycle*, *4*(6), 737–740. http://doi.org/10.4161/cc.4.6.1715

Pedersen, R. S., Karemore, G., Gudjonsson, T., Rask, M.-B., Neumann, B., Heriche, J., … Lukas, C. (2016). Profiling DNA damage response following mitotic perturbations. Nature Communications, 7. http://doi.org/10.1038/ncomms13887

Petrovics, G., Liu, A., Shaheduzzaman, S., Furusato, B., Sun, C., Chen, Y., … Srivastava, S. (2007). Frequent overexpression of ETS -related gene-1 (ERG1) in prostate cancer transcriptome Optimal primer design using the novel primer design program : MSPprimer provides accurate methylation analysis of the ATM promoter. *Oncogene*, *1*(March), 1210433. http://doi.org/10.1038/sj.onc.1210745

Peuscher, M. H., & Jacobs, J. J. L. (2011). DNA-damage response and repair activities at uncapped telomeres depend on RNF8. Nature Cell Biology, 13(9), 1139–1145. http://doi.org/10.1038/ncb2326

Pierce, A. J., Johnson, R. D., Thompson, L. H., Pierce, A. J., Johnson, R. D., Thompson, L. H., & Jasin, M. (1999). {XRCC}3 promotes homology-directed repair of {DNA} damage in mammalian cells service {XRCC}3 promotes homology-directed repair of {DNA} damage in mammalian cells, 2633–2638.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT – PCR. *Nucleic Acids Research*, *29*(9), 16–21.

Pietrucha, B., Heropolitan, E., Geffers, R., EnBen, J., Wieland, B., Bogdanova, N. V., & Dörk, T. (2017). Clinical and Biological Manifestation of RNF168 Deficiency in Two Polish Siblings. *Frontiers in Immunology*, *8*(December), 1–8. http://doi.org/10.3389/fimmu.2017.01683

Polo, S. E., & Jackson, S. P. (2011). Dynamics of DNA damage response proteins at DNA breaks : a focus on protein modifications. *Genes and Development*, *25*, 409–433. http://doi.org/10.1101/gad.2021311.critical

Prakash, R., Zhang, Y., Feng, W., & Jasin, M. (2015). Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. *Cold Spring Harbor Perspectives in Biology*, 1–27.

Ringvoll, J., Moen, M. N., Nordstrand, L. M., Meira, L. B., Pang, B., Bekkelund, A., … Klungland, A. (2008). AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA. Cancer Research, 68(11), 4142–4149. http://doi.org/10.1158/0008-5472.CAN-08-0796

Rogakou, E. P., Boon, C., Redon, C., & Bonner, W. M. (1999). Megabase Chromatin Domains Involved in DNA Double-Strand Breaks In Vivo. *The Journal of Cell Biology*, *146*(5), 905–915.

Roundtree, I. A., Evans, M. E., Pan, T., & Chuan He. (2017). Dynamic RNA Modifications in Gene Expression Regulation. *Cell*, *169*(7), 1187. http://doi.org/10.1016/j.cell.2017.05.045

Roundtree, I. A., Luo, G. Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., … He, C. (2017). YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *ELife*, *6*, 1–28. http://doi.org/10.7554/eLife.31311

Saal, L. H., Gruvberger-saal, S. K., Persson, C., Jumppanen, M., Lo, K., Pires, M. M., … Murty, V. V. V. S. (2008). Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. *Nature Genetics*, *40*(1), 102–107. http://doi.org/10.1038/ng.2007.39

Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., … Schwartz, S. (2017). The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature*. http://doi.org/10.1038/nature24456

Saghafinia, S., Mina, M., Riggi, N., Hanahan, D., & Ciriello, G. (2018). Pan-Cancer Landscape of Aberrant DNA Methylation across Human Tumors Resource Pan-Cancer Landscape of Aberrant DNA Methylation across Human Tumors. Cell Reports, 25(4), 1066–1080.e8. http://doi.org/10.1016/j.celrep.2018.09.082

Sedgwick, B. (2004). Repairing DNA-methylation damage. *Nature Reviews Molecular Cell Biology*, *5*(2), 148–157. http://doi.org/10.1038/nrm1312

Sedgwick, B., Bates, P. a., Paik, J., Jacobs, S. C., & Lindahl, T. (2007). Repair of alkylated DNA: Recent advances. *DNA Repair*, *6*(4), 429–442. http://doi.org/10.1016/j.dnarep.2006.10.005

Shanbhag, N. M., Rafalska-metcalf, I. U., Balane-bolivar, C., Janicki, S. M., & Greenberg, R. A. (2010). ATM-Dependent Chromatin Changes Silence Transcription In cis to DNA Double-Strand Breaks. Cell, 141(6), 970–981. http://doi.org/10.1016/j.cell.2010.04.038

Sheth, U., & Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science*, *300*(5620), 805–808. http://doi.org/10.1126/science.1082320

Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., … He, C. (2017). YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Research*, *27*(3), 315–328. http://doi.org/10.1038/cr.2017.15

Shivange, G., Monisha, M., Nigam, R., Kodipelli, N., & Anindya, R. (2016). RecA stimulates AlkB-mediated direct repair of DNA adducts. *Nucleic*

Acids Research, gkw611. http://doi.org/10.1093/nar/gkw611

Simó-Riudalbas, L., & Esteller, M. (2014). Cancer genomics identifies disrupted epigenetic genes. *Human Genetics*, *133*(6), 713–725. http://doi.org/10.1007/s00439-013-1373-5

Skoulidis, F., Cassidy, L. D., Pisupati, V., Jonasson, J. G., Bjarnason, H., Eyfjord, J. E., … Venkitaraman, A. R. (2010). Article Germline Brca2 Heterozygosity Promotes in a Murine Model of Familial Pancreatic Cancer. *Cancer Cell*, *18*(5), 499–509. http://doi.org/10.1016/j.ccr.2010.10.015

Śledź, P., & Jinek, M. (2016). Structural insights into the molecular mechanism of the m6A writer complex. *ELife*, *5*(September), 1–16. http://doi.org/10.7554/eLife.18434

Slobodin, B., Han, R., Calderone, V., Vrielink, J. A. F. O., Loayza-puch, F., Elkon, R., … Methylation, N. (2017). Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Article Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional. *Cell*, *169*(2), 326–337.e12. http://doi.org/10.1016/j.cell.2017.03.031

Soll, J. M., Sobol, R. W., & Mosammaparast, N. (2016). Regulation of DNA Alkylation Damage Repair: Lessons and Therapeutic Opportunities. *Trends in Biochemical Sciences*, *xx*, 1–13. http://doi.org/10.1016/j.tibs.2016.10.001

Spitzwieser, M., Holzweber, E., Pfeiler, G., Hacker, S., & Cichna-markl, M. (2015). Applicability of HIN-1 , MGMT and RASSF1A promoter methylation as biomarkers for detecting field cancerization in breast cancer. *Breast Cancer Research*, 1–13. http://doi.org/10.1186/s13058-015-0637-5

Starcevic, D., Dalal, S., & Sweasy, J. B. (2004). Is there a link between DNA polymerase β and cancer? *Cell Cycle*, *3*(8), 998–1001. http://doi.org/10.4161/cc.3.8.1062

Stefansson, O. A., Jonasson, J. G., Olafsdottir, K., Hilmarsdottir, H., Olafsdottir, G., Esteller, M., … Eyfjord, J. E. (2011). CpG island hypermethylation of BRCA1 and loss of pRb as co-occurring events in basal/triple-negative breast cancer. *Epigenetics*, *6*(5), 638–649. http://doi.org/10.4161/epi.6.5.15667

Stefansson, O. A., Moran, S., Gomez, A., Sayols, S., Arribas-Jorba, C., Sandoval, J., … Esteller, M. (2015). A DNA methylation-based definition of biologically distinct breast cancer subtypes. Molecular Oncology, 9(3), 555–568. http://doi.org/10.1016/j.molonc.2014.10.012

Stewart, G. S., Panier, S., Townsend, K., Al-Hakim, A. K., Kolas, N. K., Miller, E. S., … Durocher, D. (2009). The RIDDLE Syndrome Protein Mediates

a Ubiquitin-Dependent Signaling Cascade at Sites of DNA Damage. *Cell*, *136*(3), 420–434. http://doi.org/10.1016/j.cell.2008.12.042

Stoilov, P., Rafalska, I., & Stamm, S. (2002). YTH: A new domain in nuclear proteins. *Trends in Biochemical Sciences*, *27*(10), 495–497. http://doi.org/10.1016/S0968-0004(02)02189-8

Stoltzfus, C. M., & Dane, R. W. (1982). Accumulation of spliced avian retrovirus mRNA is inhibited in S-adenosylmethionine-depleted chicken embryo fibroblasts. *Journal of Virology*, *42*(3), 918–931.

Stucki, M., Clapperton, J. A., Mohammad, D., Yaffe, M. B., Smerdon, S. J., & Jackson, S. P. (2005). MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks. Cell, 1213–1226. http://doi.org/10.1016/j.cell.2005.09.038

Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., … Yan, W. (2017). ALKBH5-dependent m6A demethylation controls splicing and stability of long 3′-UTR mRNAs in male germ cells. *Proceedings of the National Academy of Sciences*, 201717794. http://doi.org/10.1073/pnas.1717794115

Tasaki, M., Shimada, K., Kimura, H., Tsujikawa, K., & Konishi, N. (2011). ALKBH3, a human AlkB homologue, contributes to cell survival in human non-small-cell lung cancer. *British Journal of Cancer*, *104*(4), 700–706. http://doi.org/10.1038/sj.bjc.6606012

Teschendorff, A. E., Gao, Y., Jones, A., Ruebner, M., Beckmann, M. W., Wachter, D. L., … Widschwendter, M. (2016). DNA methylation outliers in normal breast tissue identify field defects that are enriched in cancer. Nature Communications, 7, 1–12. http://doi.org/10.1038/ncomms10478

Thorleifsson, G., Walters, G. B., Gudbjartsson, D. F., Steinthorsdottir, V., Sulem, P., Helgadottir, A., … Stefansson, K. (2009). Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nature Genetics*, *41*(1), 18–24. http://doi.org/10.1038/ng.274

Tomasetti, C., & Vogelstein, B. (2015). Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science, 347(6217).

Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., & Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature*, *419*(September). http://doi.org/10.1038/nature01047.1.

Tsai, C. J., & Chu, G. (2013). Cooperative Assembly of a Protein-DNA Filament for Nonhomologous End Joining. The Journal of Biological Chemistry, 288(25), 18110–18120. http://doi.org/10.1074/jbc.M113.464115

Tulinius, H., Olafsdottir, G. H., Sigvaldason, H., Arason, a, Barkardottir, R. B.,

Egilsson, V., … Eyfjord, J. E. (2002). The effect of a single BRCA2 mutation on cancer in Iceland. *Journal of Medical Genetics*, *39*(7), 457–462. http://doi.org/10.1136/jmg.39.7.457

Tutt, A., Tovey, H., Cheang, M. C. U., Kernaghan, S., Kilburn, L., Gazinska, P., … Bliss, J. M. (2018). Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: The TNT Trial. *Nature Medicine*, *24*(5), 628–637. http://doi.org/10.1038/s41591-018-0009-7

Ukai, A., Maruyama, T., Mochizuki, S., Ouchida, R., & Masuda, K. (2006). Role of DNA polymerase θ in tolerance of endogenous and exogenous DNA damage in mouse B cells. *Genes to Cells*, *11*, 111–121. http://doi.org/10.1111/j.1365-2443.2006.00922.x

Vijg, J., & Suh, Y. (2013). Genome Instability and Aging. *Annual Review of Physiology*, *75*(1), 645–668. http://doi.org/10.1146/annurev-physiol-030212-183715

Walter, J., Cremer, T., Miyagawa, K., & Tashiro, S. (2003). A new system for laser-UVA-microirradiation of living cells. *Journal of Microscopy*, *209*(February), 71–75.

Wang, P., Doxtader, K. A., & Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Molecular Cell*, *63*(2), 306–317. http://doi.org/10.1016/j.molcel.2016.05.041

Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., … He, C. (2014). N 6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*, *505*(7481), 117–120. http://doi.org/10.1038/nature12730

Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., … He, C. (2015). N6-methyladenosine modulates messenger RNA translation efficiency. *Cell*, *161*(6), 1388–1399. http://doi.org/10.1016/j.cell.2015.05.014

Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., He, P. C., … He, C. (2018). Differential m6A, m6Am, and m1A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. *Molecular Cell*, *0*(0), 1–13. http://doi.org/10.1016/j.molcel.2018.08.011

Wickramasinghe, V. O., & Laskey, R. A. (2015). Control of mammalian gene expression by selective mRNA export. *Nature Publishing Group*, (June), 1–12. http://doi.org/10.1038/nrm4010

Willett, C. G., Boucher, Y., Di Tomaso, E., Duda, D. G., Munn, L. L., Tong, R. T., … Jain, R. K. (2004). Direct evidence that the VEGF-specific antibody bevacizumab has antivascular effects in human rectal cancer. *Nature Medicine*, *10*(2), 145–147. http://doi.org/10.1038/nm988

Wilson, J. H., & Elledge, S. J. (2002). BRCA2 Enters the Fray. *Science*, *62*(September), 1822–1823.

Wyatt, M. D., & Pittman, D. L. (2008). Methylating agents and DNA repair responses: methylated bases and sources of strand breaks. *Chemical Research in Toxicology*, *19*(12), 1580–1594. http://doi.org/10.1021/tx060164e.Methylating

Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., … Yang, Y.-G. (2016). Nuclear m 6 A Reader YTHDC1 Regulates mRNA Splicing. *Molecular Cell*, *61*(4), 507–519. http://doi.org/10.1016/j.molcel.2016.01.012

Xu, G., Chapman, J. R., Brandsma, I., Yuan, J., Mistrik, M., Bartkova, J., … Nieuwland, M. (2015). REV7 counteracts DNA double-strand break resection and impacts PARP inhibition. *Nature*, *521*(7553), 541–544. http://doi.org/10.1016/j.nbd.2016.12.013

Yamato, I., Sho, M., Shimada, K., Hotta, K., Ueda, Y., Yasuda, S., … Nakajima, Y. (2012). PCA-1/ALKBH3 contributes to pancreatic cancer by supporting apoptotic resistance and angiogenesis. *Cancer Research*, *72*(18), 4829–4839. http://doi.org/10.1158/0008-5472.CAN-12-0328

Yang, X., Yang, Y., Sun, B. F., Chen, Y. S., Xu, J. W., Lai, W. Y., … Yang, Y. G. (2017). 5-methylcytosine promotes mRNA export-NSUN2 as the methyltransferase and ALYREF as an m 5 C reader. *Cell Research*, *27*(5), 606–625. http://doi.org/10.1038/cr.2017.55

Zack, T. I., Schumacher, S. E., Carter, S. L., Cherniack, A. D., Saksena, G., Tabak, B., … Getz, G. (2013). Pan-cancer patterns of somatic copy number alteration. Nature Genetics, 45(10). http://doi.org/10.1038/ng.2760

Zhao, B. S., Roundtree, I. A., & He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol*, *18*(1), 31–42. http://doi.org/10.1001/jamasurg.2014.1086.Feasibility

Zhao, X., Yang, Y., Sun, B.-F., Shi, Y., Yang, X., Xiao, W., … Yang, Y.-G. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Research, 24(12), 1403–1419. http://doi.org/10.1038/cr.2014.151

Zhao, Y., Majid, M. C., Soll, J. M., Brickner, J. R., & Dango, S. (2015). Noncanonical regulation of alkylation damage resistance by the OTUD 4 deubiquitinase. The EMBO Journal, 34(12), 1687–1703.

Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., … He, C. (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell*, *49*(1), 18–29. http://doi.org/10.1016/j.molcel.2012.10.015

Zhou, J., Wan, J., Shu, X. E., Mao, Y., Liu, X.-M., Yuan, X., … Qian, S.-B. (2018). N 6 -Methyladenosine Guides mRNA Alternative Translation during Integrated Stress Response. *Molecular Cell*, 1–12.

http://doi.org/10.1016/j.molcel.2018.01.019

Zhu, Q., Sharma, N., He, J., Wani, G., & Wani, A. a. (2015). USP7 deubiquitinase promotes ubiquitin-dependent DNA damage signaling by stabilizing RNF168*. *Cell Cycle*, *14*(9), 1413–1425. http://doi.org/10.1080/15384101.2015.1007785

Zong, D., Adam, S., Wang, Y., Sasanuma, H., Callen, E., Johnson, N., … Nussenzweig, A. (2019). BRCA1 Haploinsufficiency Is Masked by RNF168- Article BRCA1 Haploinsufficiency Is Masked by RNF168-Mediated Chromatin Ubiquitylation. Molecular Cell, 73, 1–15. http://doi.org/10.1016/j.molcel.2018.12.010

Zou, L., & Elledge, S. J. (2003). Sensing DNA Damage Through ATRIP Recognition of RPA-ssDNA. *Science*, *300*(June), 1542–1549.

Stefansson et al. BMC Cancer (2017) 17:469
DOI 10.1186/s12885-017-3453-8

BMC Cancer

RESEARCH ARTICLE Open Access

CpG promoter methylation of the ALKBH3 alkylation repair gene in breast cancer

Olafur Andri Stefansson[1,2], Stefan Hermanowicz[2,3], Jasper van der Horst[1,2], Holmfridur Hilmarsdottir[1,2], Zuzanna Staszczak[1], Jon Gunnlaugur Jonasson[2,4,5], Laufey Tryggvadottir[2,4], Thorkell Gudjonsson[2,3] and Stefan Sigurdsson[2,3*]

Abstract

Background: DNA repair of alkylation damage is defective in various cancers. This occurs through somatically acquired inactivation of the MGMT gene in various cancer types, including breast cancers. In addition to MGMT, the two *E. coli* AlkB homologs ALKBH2 and ALKBH3 have also been linked to direct reversal of alkylation damage. However, it is currently unknown whether ALKBH2 or ALKBH3 are found inactivated in cancer.

Methods: Methylome datasets (GSE52865, GSE20713, GSE69914), available through Omnibus, were used to determine whether ALKBH2 or ALKBH3 are found inactivated by CpG promoter methylation. TCGA dataset enabled us to then assess the impact of CpG promoter methylation on mRNA expression for both ALKBH2 and ALKBH3. DNA methylation analysis for the ALKBH3 promoter region was carried out by pyrosequencing (PyroMark Q24) in 265 primary breast tumours and 30 proximal normal breast tissue samples along with 8 breast-derived cell lines. ALKBH3 mRNA and protein expression were analysed in cell lines using RT-PCR and Western blotting, respectively. DNA alkylation damage assay was carried out in cell lines based on immunofluorescence and confocal imaging. Data on clinical parameters and survival outcomes in patients were obtained and assessed in relation to ALKBH3 promoter methylation.

Results: The ALKBH3 gene, but not ALKBH2, undergoes CpG promoter methylation and transcriptional silencing in breast cancer. We developed a quantitative alkylation DNA damage assay based on immunofluorescence and confocal imaging revealing higher levels of alkylation damage in association with epigenetic inactivation of the ALKBH3 gene ($P = 0.029$). In our cohort of 265 primary breast cancer, we found 72 cases showing aberrantly high CpG promoter methylation over the ALKBH3 promoter (27%; 72 out of 265). We further show that increasingly higher degree of ALKBH3 promoter methylation is associated with reduced breast-cancer specific survival times in patients. In this analysis, ALKBH3 promoter methylation at >20% CpG methylation was found to be statistically significantly associated with reduced survival (HR = 2.3; $P = 0.012$). By thresholding at the clinically relevant CpG methylation level (>20%), we find the incidence of ALKBH3 promoter methylation to be 5% (13 out of 265).

Conclusions: ALKBH3 is a novel addition to the catalogue of DNA repair genes found inactivated in breast cancer. Our results underscore a link between defective alkylation repair and breast cancer which, additionally, is found in association with poor disease outcome.

Keywords: Breast cancer, ALKBH3, Epigenetics, DNA methylation, DNA repair, Alkylation, Prognosis

* Correspondence: stefsi@hi.is
[2]Faculty of Medicine, University of Iceland, Vatnsmyrarvegur 16 (4th floor), 101 Reykjavik, Iceland
[3]Department of Biochemistry and Molecular Biology, Biomedical Center, Vatnsmyrarvegur 16 (5th floor), 101 Reykjavik, Iceland
Full list of author information is available at the end of the article

Background

Epigenetics can be described as the study on mechanisms by which changes in phenotype are established and stably maintained following cellular divisions without involving any changes in genotype. The fundamental unit of chromatin is the nucleosome representing a short stretch of DNA wrapped around a protein complex consisting of histone variants arranged as octamers [1]. Nucleosome occupancy over regulatory regions in DNA associates with transcriptional activity as densely occupied regions are poorly accessible to transcription factors [2]. The degree of nucleosome occupancy, or chromatin packaging, is regulated by the use of so-called epigenetic marks of which the best studied is undoubtedly DNA methylation involving the addition of methyl groups to the 5-position on cytosine bases (5-meC) where cytosine is followed by guanine, i.e. so-called CpG sites [3]. Methyl-Binding Domain-containing proteins recognize and bind to methylated CpGs and recruit histone modifiers to mediate or maintain repressed chromatin structure [3].

In cancer cells, the epigenome is frequently disrupted - characterized by global loss in repressive marks and localized modifications over regulatory elements [4]. It is now known that genes functionally involved in shaping the epigenome of human cells are recurrently mutated in breast cancer and various other cancers, e.g. MLL3, MLL2, ARID1A and SETD2 [5]. The discoveries of recurrent mutations in epigenetic genes provided a convincing link between disruptions in the epigenome and the development of cancer. In addition to this, earlier observations had already established repressive epigenetic marks over regulatory regions of known tumor suppressor genes in cancer cells. In breast cancer, this catalogue includes CpG promoter methylation of BRCA1, RAD51C, FOXC1, RUNX3 and L3MBTL4 [6–10]. Of these, the BRCA1 gene is well established as a cancer predisposition gene where germline mutations are found in association with greatly increased risk for breast and ovarian cancer [11]. Other high-risk breast cancer susceptibility genes such as BRCA2, PALB2, BARD1, FANCM, ATM, CHEK2 and TP53, however, are not found epigenetically silenced [12, 13].

The onset of a subset of breast cancers is strongly tied to defective repair of DNA double-stranded breaks by homologous recombination [14, 15]. In recent years, several researchers have reported loss of MGMT in breast cancer thereby implicating defective alkylation repair in breast cancer development [16–18]. The MGMT gene has an important role in removing cytotoxic adducts from O(6)-guanine in DNA [19]. Proficiency for alkylation repair is critical to protect the cell against accumulation of genetic mutations. Indeed, recent cancer genome sequencing studies have revealed a profound impact from treatment with alkylating agent temozolomide leading to a specific mutational pattern [20]. This observation provides a solid link between alkylation damage and the formation of genetic mutations.

Alkylating agents are by-products of normal cellular metabolism as well as being ubiquitous in the environment [21]. In addition to MGMT, at least two other genes are known to be involved in direct reversal of alkylation damage, i.e. the *E. coli* AlkB homologs ALKBH2 and ALKBH3 both oxidative demethylases involved in the repair of 1-methyladenine and 3-methylcytosine [21]. It is currently unknown whether ALKBH2 or ALKBH3 undergo epigenetic silencing in cancer. In this study, we demonstrate that the ALKBH3 gene, not ALKBH2, is found recurrently silenced in breast cancer by epigenetic events which, furthermore, defines a group of patients with dramatically reduced survival.

Methods

Study cohort

DNA samples were derived from primary breast tumors (n = 265) and adjacent normal breast tissue (n = 30). The normal breast tissue was obtained from non-tumorous regions of the breast. The DNA was previously isolated from freshly frozen tissue following a standard protocol based on phenol-chloroform (+proteinase K) extraction. RNA samples were available for a subset of the tumor (n = 36) and normal breast (n = 10) tissue samples isolated using Tri-Reagent (Thermo Fisher Scientific). Clinical parameters, including tumor size, nodal status, histological grade along with disease-specific follow-up times were obtained from the nationwide Icelandic Cancer Registry [22]. This work was carried out according to permits from the Icelandic Data Protection Commission (2006050307) and Bioethics Committee (VSNb2006050001/03–16). Informed consent (written) was obtained from all patients.

The cell lines used in this study were obtained from the American Type Culture Collection. The cells were cultured in DMEM (CAMA-1, MDAMB-468, MCF-7, MCF-10A, MDAMB-231 and SKBr-3) or RPMI (HCC-38, Bt-474) with added 10% serum (+penicillin/streptomycin). DNA and RNA was extracted in parallel from the cell lines using Qiagen's AllPrep DNA/RNA/miRNA Universal kit (80224; Qiagen).

DNA methylation analyses

Bisulfite conversion was carried out using the EZ-96 DNA Methylation-Gold kit from Zymo Research (D5008). We carried out the bisulfite conversion for 16 cycles of {95 °C for 30 s and 50 °C for 1 h} to then hold at 4 °C until samples were added to the DNA columns for completing the conversion following the manufacturer's guidelines (Zymo Research).

The PyroMark Assay Design 2.0 software was used to design primers for the analysis of ALKBH3 promoter methylation. The following primer sequences

were used: 5′-(Btn)-GTGGGATTATTAGGATTGAG GATT-3′ (5-biotin labelled) and 5′-CTCCAACAACTCC CAATCAC-3′. The pre-amplification PCR reaction was carried out using a hot-start PCR polymerase (Immolase DNA polymerase from Bioline; Bio-21,047). The PCR conditions were as follows: 96 °C for 10 min, 45 cycles of (96 °C for 30s, 60 °C for 30s and 72 °C for 30s) followed by 15 min hold at 72 °C and then 4 °C. The PCR products were then captured using streptavidin coated agarose beads (Streptavidin Sepharose High Performance 34 μm beads, GE Healthcare) under denaturing conditions to obtain single-stranded DNA. The pyrosequencing reaction was then carried out using the PyroMark Q24 machine (Qiagen) and PyroMark Gold-Q24 Reagents kit (Qiagen) using the following sequencing-primer: 5′-ACATCAAA CACTTCCT-3′.

CpG methylation for three CpG's were assessed (−58, −53 and −50 bp upstream of the TSS (p1) given the FANTOM5 promoterome database) [23]. The output data (obtained from PyroMark Q24 sequencing reactions), representing percent methylated cytosines over each of these three CpGs, was averaged for each sample analysed. This yielded a single measure representing a proxy for CpG methylation levels over the ALKBH3 promoter region. The statistical analysis of paired tumor and normal breast tissue samples made use of a paired Wilcoxon's test using the wilcox.test function in R.

Expression analyses in normal breast tissue, tumors and cancer cell lines

RNA was extracted from tumors and normal breast tissues using Tri-Reagent (Thermo Fisher Scientific). The RNA derived from cell lines was isolated in a simultaneous DNA/RNA isolation procedure using the Qiagen's Allprep kit (Qiagen). Reverse transcription was carried out using High-Capacity cDNA Reverse-Transcription Kit (Thermo Fisher Scientific). ALKBH3 expression was quantitatively analysed by the SYBR green method using a real-time PCR (RT-PCR) machine (Applied Biosystems 7500). HPRT1 expression was used to normalize the expression data by computing the difference in Ct as follows: $2^{-(\text{ALKBH3 Ct} - \text{HPRT1 Ct})}$. The primers used for ALKBH3 were: 5′-AGCCACGAGTGATTGACAGAG-3′ and 5′-ACAAACAGACCCTAGATACACCT-3′, and for HPRT1: 5′- CCTGGCGTCGTGATTAGTGAT-3′ and 5′-AGACG TTCAGTCCTGTCCATAA-3′. The Spearman's rank test for correlation was carried out using the cor.test function in R to assess the association between CpG methylation and expression.

Proteins were extracted from cell lines at 80% confluency using the EBC lysis buffer and measured at 490 nm using a spectrophotometer. The samples were denatured and electropherized using a 10% gel followed by transfer to PVDF membrane. The primary antibody (Millipore anti-ALKBH3 rabbit polyclonal 09–882) was used at 1:500 dilution overnight at 4 °C followed by washing with PBS-Tween. The secondary antibody (Santa Cruz donkey anti-rabbit IgG-HRP, sc-2313) was used at 1:10,000 dilution. The membrane was developed with ECL (Pierce ECL Plus Western Blotting Substrate, Thermo Scientific 32,132) and detected in a ImageQuant LAS4000. The β-actin primary antibody (MAB1501R; Millipore) was used at a dilution of 1:20,000 with secondary HRP-rabbit antibody anti-mouse IgG used at 1: 10,000 dilution (61–6020; Invitrogen).

DNA alkylation damage detection

CAMA1 and MDA-MB-468 were grown on coverslips and fixed with freshly prepared 4% para-formaldehyde solution for 15 min. After fixation, cells were treated with 1.5 M HCL for 20 min, to gain access to single stranded DNA, followed by a 2-min treatment with Sodium Borate (pH 8.5) to neutralize the acid. After permabilization (5 min, 0.2% TritonX) and 1 h of blocking (DMEM (Gibco) with 10% FBS (Gibco)) cells were stained with antibodies against 3-methylcytosine (3meC) (rabbit, Active Motif, 61111) and 5-methylcytosine (5meC) (mouse, abcam, ab10805) for 1 h at room temperature. Both antibodies were diluted 1:250 in blocking buffer. Next, samples were incubated with secondary antibodies, Alexa-Fluor 488 goat anti rabbit (A11008, Thermo Fisher Scientific) and Alexa-Fluor 555 goat anti-mouse (A21422, Thermo Fisher Scientific), diluted in blocking buffer (1:1000) for 1 h. Nuclear DNA was stained by DAPI (SIGMA, D9542). The DAPI stain was added directly to the secondary antibody solution (diluted 1:5000). After drying, the coverslips were mounted on glass slides using Fluoroshield (SIGMA, F6182) mounting medium.

Images were acquired using the FV1200 Olympus inverted confocal microscope. Dual colour confocal images were acquired with standard settings using laser lines 488 nm and 543 nm for excitation of Alexa Fluor 488 and Alexa Fluor 568 dyes, respectively. Nuclear DAPI staining was imaged using excitation by the 405 nm laser. For each condition 10 images were randomly acquired with the 20X/0.75 objective and imported into CellProfiler for downstream image analysis. For each data point, 400–600 nuclei (identified by DAPI staining) were analysed for 3meC and 5meC nuclear intensity (mean integrated intensity).

The 3meC and 5meC values presented in Fig. 2d are based on four independent staining experiments. The Wilcoxon's rank sum hypothesis test was used to assess differences in 3meC and 5meC values (in R 3.1.0).

Tissue microarrays (TMAs)

Estrogen receptor (ER), progesterone receptor (PR) and HER-2 expression were previously analysed on tumors

by immunohistochemistry (IHC) on tissue microarrays (TMAs) [24]. The TMAs were constructed as previously described (Stefansson et al. 2009). Immunohistochemistry (IHC) was then applied using 4 μm thick TMA sections using the following anti-bodies: anti-ER (1D5; DAKO), anti-PR (PgR 636; DAKO) and anti-HER2 (HercepTest Kit; DAKO). ER and PR were scored positive given any visible nuclear staining in more than 1% of tumor cell nuclei. HER-2 positivity was defined as score of 3+ according to criteria provided by the anti-body manufacturer.

Informatics and statistical analyses

Information on CpG methylation over the promoter region of ALKBH2 and ALKBH3 was obtained from pre-existing methylome analyses published by Stefansson et al. (GSE52865), Dedeurwaerder et al. (GSE20713) and Teschendorff et al. (GSE69914) available through the Omnibus repository at NCBI's website (www.ncbi.nlm.nih.gov/gds/) [7, 25, 26]. The normalized data were extracted from the SOFT formatted files using the GEOquery package in R and analysed by comparing normal breast tissue samples and breast cancers. This was carried out using the Student's t-test on M–values computed using $M_i = \log_2(B_i / (1\text{-}B_i))$ where B represents the β-value coupled with the Benjamini-Hochberg adjustment procedure to account for multiple hypothesis testing making use of the p.adjust function in R. The multiple hypothesis adjustment accounted for the total number of CpGs represented on the array platform, i.e. adjusting for the entire >450 thousand CpGs in GSE52865 and GSE69914 and >27 thousand CpGs in GSE20713.

DNA methylation (450 K Infinium) and RNAseq (V2) level 3 data were downloaded from the Cancer Genome Atlas data repository (http://cancergenome.nih.gov/) [27]. Firstly, the analysis of differential ALKBH3 mRNA expression levels between normal breast tissue and breast cancers was carried out using the Wilcoxon's rank sum hypothesis test taking into account adjustment for multiple hypothesis testing including the entire set of >20 thousand genes included in the RNAseqV2 dataset. This was carried out using the Benjamini-Hochberg (BH) procedure through the p.adjust function in R. Secondly, the relation between CpG methylation for each site represented over either ALKBH2 and ALKBH3 were studied with respect to ALKBH3 mRNA expression using Spearman's correlation analysis and, as before, with genome-wide adjustment of the *P*-values using the BH procedure to account for multiple hypothesis testing.

Information on epigenetic marks for the ALKBH3 promoter region in variant human mammary epithelial cells (vHMEC) was obtained from the Roadmap Epigenomics browser (egg2.wustl.edu/roadmap/web_portal/) [28]. This includes information on ALKBH3 expression based on RNA sequencing and chromatin marks based on ChIP-seq along with data on chromatin accessibility based on DNA sequencing. Data on CpG methylation for the vHMEC cells was derived from methylCRF computational analysis using MeDIP-seq and MRE-seq data to infer whole-genome 5-methylcytosine states as carried out and provided by the Roadmap Epigenomics project [28].

Information on nucleotide positions for ALKBH3 gene structure (introns and exons) was downloaded from Ensembl (GRCh37 browser; HG19). Data on transcriptional start site (TSS) and CpG islands were obtained from the FANTOM5 promoterome [23]. Using the UCSC genome browser, the chromStart/chromEnd fields in the hg19.cpgIslandExt table provided the CpG island positional information. The R statistical software (R 3.1.0) was then used to graphically represent the ALKBH3 promoter with respect to the TSS, 1st Exon and CpG island.

The association between ALKBH3 promoter methylation and subtype-specific markers was assessed using wilcoxon's rank sum hypothesis testing (wilcox.test in R). The chi-squared test was used to assess the association between tumor subtype classification, histological grade, tumor size and nodal status (chisq.test in R). Differences in breast cancer-specific patient survival with respect to ALKBH3 methylation in tumor tissue was assessed using the log-rank hypothesis test (survdiff function in R). Cox's proportional hazards regression model was use for multivariate analysis of survival (coxph function in R). The cox.zph function in R was applied to assess the assumptions of the regression model with respect to proportionality over time.

Results

Methylome analyses identify ALKBH3 as a target of CpG promoter methylation in breast cancer

By making use of methylome data for breast cancers and normal breast tissues, we specifically asked whether aberrant CpG methylation events are found over the promoter region of either ALKBH2 or ALKBH3. To achieve this, we used datasets available through Omnibus including those published by Stefansson et al. (GSE52865), Dedeurwaerder et al. (GSE20713) and Teschendorff et al. (GSE69914) [7, 25, 26]. The analysis of these datasets consistently identify aberrant CpG methylation over the ALKBH3 gene promoter in breast cancers. Figure 1a illustrates this finding where statistically significant CpG methylation events are seen over the ALKBH3 promoter region ($P_{adj} < 0.001$). In contrast, differential methylation between breast cancer and normal breast tissue was not identified over the promoter region of ALKBH2 (Fig. 1a).

We used the Cancer Genome Atlas dataset to assess the impact of ALKBH3 promoter methylation on mRNA

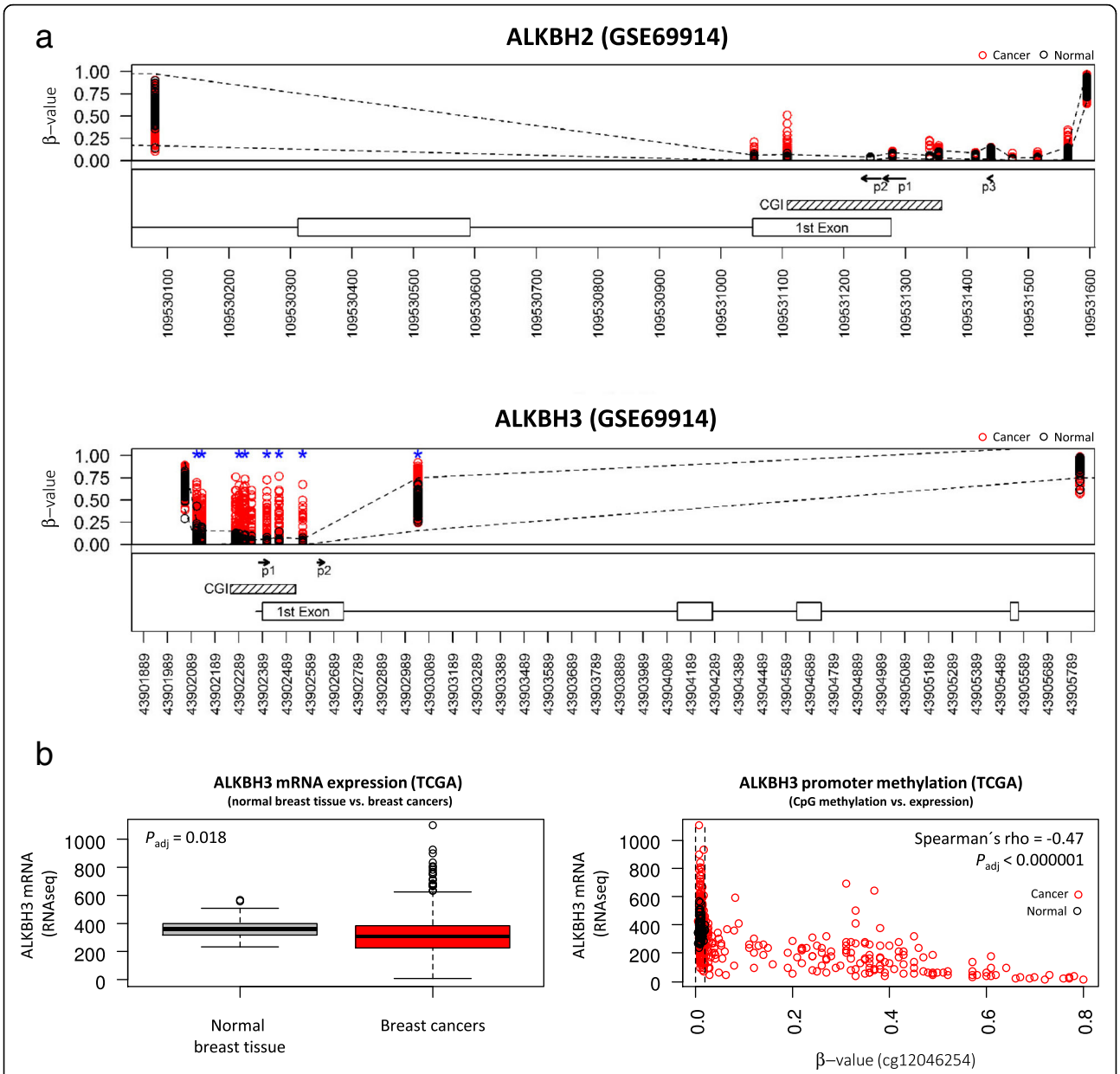


Fig. 1 CpG methylation over the promoter region ALKBH2 and ALKBH3 in breast cancers and normal breast tissue samples. **a** CpG methylation states obtained from GSE69914 over ALKBH2 (upper panel) and ALKBH3 (lower panel) in normal breast tissue samples and breast cancers in *black* and *red*, respectively. The y-axis represents β (beta)-values (reflecting the degree of 5-cytosine methylation) for CpG's included on the Infinium arrays located within or proximal to the promoter region for ALKBH2 and ALKBH3 arranged on the x-axis according to nucleotide position. The dashed lines (in *black*) represent the upper and lower 99% confidence intervals for the distribution of β-values in normal breast tissue samples – thus displaying the "normal range" for each of the CpGs analysed. Differentially methylated CpGs between breast cancers and normal breast tissues are indicated by *blue* asterisk marks taking into account correction for multiple hypothesis testing including all CpGs represented on the 450 K array ($P_{adj} < 0.001$). FANTOM5 regions for ALKBH2 and ALKBH3 are shown as arrows indicating transcription start sites (TSS) where p1 represents the major TSS (while p2 and p3 are less prominently used as TSS). Additionally, the location of UCSC defined CpG islands (CGI; strikethrough patterned boxes) and the 1st Exon for each of the two genes are labelled. UCSC defined CpG islands from the UCSC genome table browser. **b** Left panel; ALKBH3 mRNA expression levels by RNA sequencing (RNAseq) obtained from the TCGA dataset analysed with respect to normal breast tissue samples compared with breast cancers. These differences reflect generally lowered expression levels in breast cancers compared with normal breast tissue samples. The *P*-value indicated in the upper-left corner was derived from Wilcoxon's hypothesis testing after adjusting for multiple hypotheses accounting for >20 thousand protein-coding genes represented in the RNAseq dataset ($P_{adj} = 0.018$). The right panel displays the topmost significant CpG (ranked according to the adjusted *P*-value), i.e. cg12046254, illustrating the relation between ALKBH3 mRNA expression (y-axis) and CpG promoter methylation (x-axis). Again, the dashed lines (in *black*) represent the lower and upper 99% confidence limits for the normal breast tissue samples – reflecting the "normal range" of 5-methylcytosine levels for this particular CpG (cg12046254). The *P*-value indicated in the top-right corner, based on Spearman's rho correlation analysis, was highly significant even after adjustment for multiple hypothesis testing

expression levels. Firstly, these data provide additional support for differential methylation over the ALKBH3 promoter region between normal breast tissue and tumours (data not shown). Secondly, significantly lower expression levels of ALKBH3 mRNA were seen in breast cancers compared with normal breast tissue samples. This finding holds statistically significant after adjusting for multiple hypothesis testing taking into account the entire >20 thousand protein-coding genes in the RNA-seqV2 TCGA dataset (Wilcoxon's rank sum test; $P_{adjusted}$ = 0.018) as shown in Fig. 1b (left panel). Thirdly, the majority of the CpGs identified as differentially methylated are also found significantly associated with down-regulation of ALKBH3 mRNA levels in breast cancers (Fig. 1b, right panel; Additional file 1). The top-most significant CpG derived from this analysis (Spearman's rho = −0.47; $P_{adj} < 0.00001$) is represented by cg12046254 shown in Fig. 1b (right panel).

Epigenetics and expression of ALKBH3 in normal breast epithelial cells

Information on epigenetic regulation and expression of the ALKBH3 gene in human mammary epithelial cells (vHMEC) is displayed in Fig. 2a demonstrating transcriptionally active chromatin configuration over the promoter region. This can be seen in DNase-seq signal peaks found upstream of the promoter and extending into the first exon together with active histone markings, i.e. H3 lysine 4 tri-methylation (H3K4Me3) localized over the first exon and H3 lysine 36 tri-methylation (H3K36Me3) over the gene body region collectively indicating open chromatin and active transcription (Fig. 2a). Notably, the H3K4Me3 activation marks are found in the absence of repressive H3 lysine 4 mono-methylation (H3K4me1), H3 lysine 27 tri-methylation (H3K27Me3) and H3 lysine 9 tri-methylation (H3K9Me3) as shown in Fig. 2a. Indeed, the expression data (RNAseq track) shows clear signals from all ten exons of the ALKBH3 gene – the first three exons are shown in Fig. 2a.

In agreement with transcriptionally active chromatin configuration, the ALKBH3 gene promoter region and first exon are lacking of repressive CpG methylation marks (DNA methyl track in Fig. 2a). The ALKBH3 promoter is associated with a CpG island (spanning a region from 43,902,254 bp to 43,902,528 bp) extending into the first exon (Fig. 2b) and, indeed, the entire CpG island (CGI) is found unmethylated in the vHMEC breast epithelial cells. The CGI was identified from the UCSC genome table browser defined as regions of at least 200 bp where the GC content is at least 50%. The ALKBH3 promoter associated CGI was found to be 274 base pairs in length.

We designed a DNA methylation pyrosequencing assay to carry out CpG methylation analysis for the ALKBH3 gene promoter region (Fig. 2b). The DNA region assayed, labelled R in Fig. 2b, includes three CpG

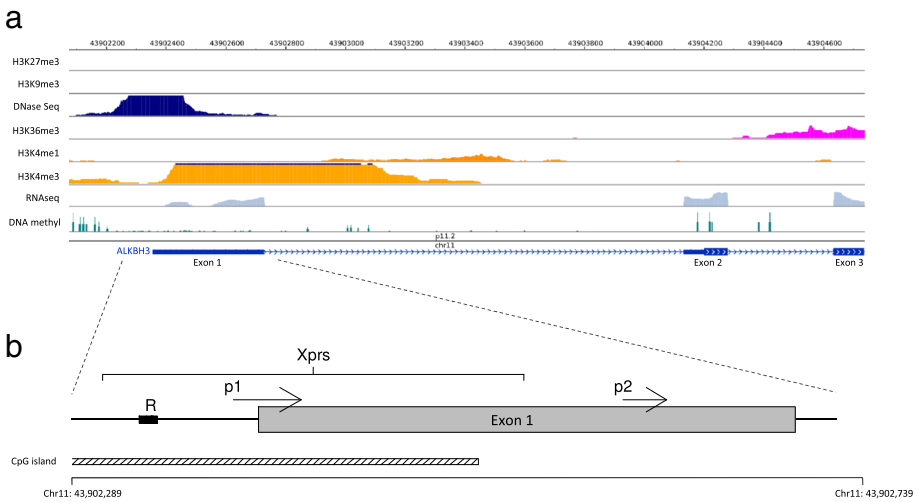


Fig. 2 The ALKBH3 gene promoter region. **a** The ALKBH3 gene promoter is transcriptionally active in variant human mammary epithelial cells (vHMEC) based on available data from the Roadmap Epigenetic Consortium (Roadmap Epigenomics Consortium, Nature 2015). Data for vHMEC are shown here for markers associated with active transcription (H3K4me3 and H3K36me3) together with repressive markers (H3K27me3, H3K9me3, H3K4me1 and 5-methylcytosine (DNA methyl)). Additionally, DNAse and RNA sequencing results from vHMEC cells are shown – reflecting chromatin accessibility and mRNA expression, respectively. **b** The ALKBH3 gene promoter region is shown with respect to the FANTOM5 transcription start site (TSS) as arrows p1 and p2 along with the 1st exon and the promoter-associated CpG island (UCSC defined). The CpG methylation assay for ALKBH3 was designed to include CpG sites proximal to the TSS and the regions selected is indicated by a *black* box (labelled R) covering three closely spaced CpG dinucleotides found -50, −53 and -58 bp upstream of the major TSS (p1 region in FANTOM5). Additionally, the region where statistically significant associations were revealed between CpG methylation and loss of expression for the ALKBH3 gene in tumors is marked out and labelled for expression as "Xprs" (see further information in Additional file 1)

sites found −50, −53 and -58 bp upstream of the major TSS (FANTOM5 element p1). In this way, the assay was designed to reflect ALKBH3 promoter methylation status which we then applied across 303 DNA samples, i.e. 8 breast-derived cell lines, 30 normal breast tissue samples and 265 primary breast tumors.

ALKBH3 epigenetic repression found in two breast cancer cell lines

Out of the eight breast-derived cell lines analysed with respect to ALKBH3 promoter methylation, seven were derived from breast tumors (CAMA-1, Bt-474, HCC-38, SKBr-3, MDA-MB-231, MCF-7, MDA-MB-468) and one was derived from a fibrocystic breast lesion (MCF-10A). We identified ALKBH3 promoter methylation in two of the seven breast cancer cell lines, i.e. in CAMA-1 and Bt-474 (Fig. 3a). The MCF10A cell line, often used to reflect normal breast epithelial cells, was clearly unmethylated over the ALKBH3 promoter.

ALKBH3 mRNA expression was not detected in either CAMA-1 or Bt-474 whereas all other cell lines showed ALKBH3 mRNA expression (Fig. 3b). The association between ALKBH3 promoter methylation and mRNA expression was found to be statistically significant (Spearman's rho = −0.73; P = 0.039). Further, complete loss of ALKBH3 protein expression was seen only in the two cell lines showing ALKBH3 promoter methylation, i.e. CAMA-1 and Bt-474 (Fig. 3c).

ALKBH3 has been reported to catalyse the removal of 3-methylcytosine (3meC) on single stranded DNA [29]. In line with that, RNAi mediated knockdown of ALKBH3 has previously been shown to cause an increase in 3meC levels on single-stranded DNA [30]. To determine the functional impact of ALKHB3 promoter methylation, we developed a novel imaging-based assay for the quantification of 3meC on single stranded DNA. Using this method we confirmed previous findings [30] demonstrating increased formation of 3meC damage following siRNA knock-down of ALKBH3 (Additional file 2). Further, we show statistically significant differences in 3meC damages between ALKBH3 deficient cell line CAMA-1 compared with ALKBH3 expressing cell line MDA-MB-468 (P = 0.029), see Fig. 3d and e. For reference, we show that 5meC intensities are not significantly different between CAMA-1 and MDA-MB-468 (P = 0.69). This indicates that, comparable to knockdown of ALKBH3 [30], epigenetic inactivation of ALKBH3 results in a higher burden of 3meC, most likely because of less efficient repair of alkylation damage.

CpG promoter methylation and expression of the ALKBH3 gene in normal and primary breast tumor samples

Out of the 265 primary tumors analysed with respect to ALKBH3 promoter methylation, a subset of 30 tumor samples were matched with adjacent normal breast tissue samples from the same patients – thereby enabling assessment of differential methylation in paired samples. This analysis reveals clear differences in CpG methylation over the ALKBH3 promoter region between primary tumors and normal breast tissue from the same 30 individuals (Fig. 4a).

This analysis further shows that, in normal breast tissue, the mean methylation levels over the ALKBH3 promoter region is approximately 1.0% (99%CI: 0.12–1.9%). Deviations from the 99% confidence interval for ALKBH3 promoter methylation levels in normal breast tissue samples can be considered "aberrant". On the basis of the upper 99%CI for normal breast tissue samples (99%CI upper limit ~2% methylation), the incidence of ALKBH3 promoter methylation in primary breast tumors is approximately 27%, i.e. 72 out of 265 primary breast tumors show aberrant ALKBH3 promoter methylation.

RNA samples available from normal breast tissue and tumor samples were used to assess ALKBH3 mRNA expression. Firstly, this analysis demonstrates statistically significant differences in ALKBH3 mRNA expression between normal breast tissue and tumor samples wherein breast tumors generally show reduced ALKBH3 expression (Fig. 4b). This provides an independent confirmation for the previous observations based on the use of available data from the TCGA project shown in Fig. 1b. Secondly, by using DNA samples available from a subset of the same tumor samples as analysed with respect to ALKBH3 expression, we were able to assess the association between mRNA expression levels and promoter methylation status. This analysis, shown in Fig. 4c, further validates the impact of ALKBH3 promoter methylation on mRNA expression levels.

ALKBH3 promoter methylation with respect to clinical relevance

Table 1 displays clinical and pathological characteristics of the patient cohort (n = 265). ALKBH3 promoter methylation was not found to be significantly associated with the expression of subtype-specific markers, i.e. estrogen-receptor (ER), progesterone-receptor (PR), erb-b2 receptor tyrosine kinase 2 (known as HER2) and MKI67 (known as Ki-67) as shown in Fig. 5a. Additionally, no associations were found for ALKBH3 promoter methylation in relation to discrete subtype classification based on these four subtype-specific markers, histological grade, tumor size or nodal status (Additional file 3).

Nonetheless, we found significantly reduced survival in patients with tumors showing ALKBH3 promoter methylation (Fig. 5b). This was seen in breast tumors with high cytosine methylation levels over the ALKBH3 promoter region, i.e. those showing at least 20% cytosine methylation (Fig. 5b). This level of methylation is indeed

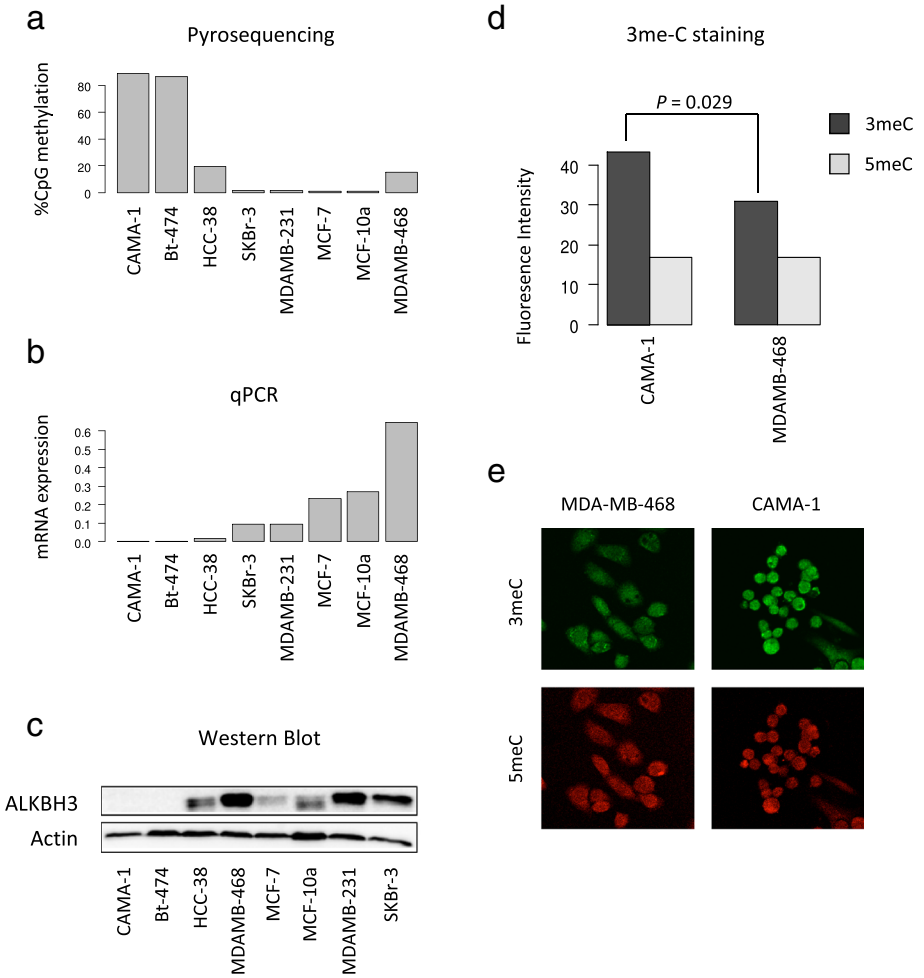


Fig. 3 Epigenetic silencing of ALKBH3 and accumulation of 3-methylcytosine damage. **a** CpG methylation analysis for the ALKBH3 gene promoter by pyrosequencing and (**b**) ALKBH3 mRNA expression analysis by qPCR in breast cancer cell lines (CAMA1, Bt-474, HCC-38, SKBr-3, MDA-MB-231, MCF-7 and MDA-MB-468) and a mammary epithelial cell line derived from a fibrocystic lesion of the breast (MCF10A). The association between ALKBH3 promoter methylation and mRNA expression was found to be statistically significant (Spearman's rho = −0.73; $P = 0.039$). **c** ALKBH3 protein expression analysed by western blotting using the same panel of breast-derived cell lines revealing lack of expression in two cell lines (CAMA-1 and Bt-474). Actin protein expression is shown for comparison. **d** Immunofluorescent staining for 3-methylcytosine and 5-methylcytosine in CAMA-1 (ALKBH3 methylated) and MDA-MB-468 (ALKBH3 unmethylated). The *P*-value shown was derived from Wilcoxon's rank sum testing for differences between CAMA-1 and MDA-MB-468 with respect to intensity for 3-methylcytosine. As expected, no statistically significant differences were found with respect to overall 5-methylcytosine intensity levels between CAMA-1 and MDA-MB-468. **e** Representative images showing 3-methylcytosine (*green*) and 5-methylcytosine (*red*) fluorescence staining in CAMA-1 and MDA-MB-468

substantially higher than the 2% threshold level defining aberrant CpG methylation as was established by looking at the distribution seen in normal breast tissue samples (see section 3.4).

The incidence of clinically relevant ALKBH3 methylation (thresholding at the 20% cytosine methylation level) is approximately 5% in our cohort (13 of 265; 4.9%) and, on the basis of Cox's proportional hazards regression analysis, we found that these patients were at approximately 2.3-fold increased risk of death resulting from breast cancer in a multivariate model including adjustment for age and year at diagnosis (HR = 2.3; $P = 0.012$).

Discussion

ALKBH3 repression by epigenetic mechanisms in breast cancer

In this study, we show that the ALKBH3 gene promoter region undergoes aberrant epigenetic repression in a significant proportion of primary breast tumours. The ALKBH3 gene is therefore a novel addition to the catalogue of DNA repair genes found inactivated in breast cancer. Previous studies have shown that treatment with the DNA alkylating agent MMS induces the expression of ALKBH3, highlighting the important role of ALKBH3 in repair of alkylation DNA damage [21]. Indeed, our results indicate that inactivation of ALKBH3 is associated

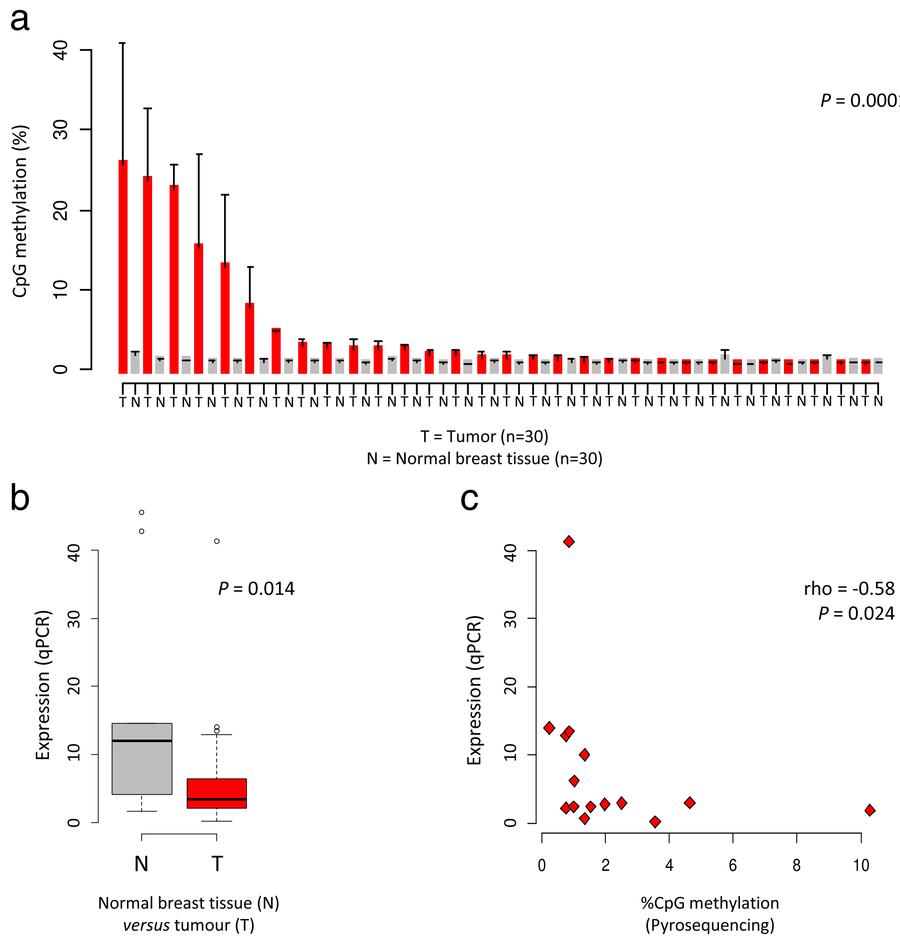


Fig. 4 ALKBH3 promoter methylation in primary breast tumors and adjacent normal breast tissue. **a** Differential CpG methylation over the ALKBH3 promoter is observed in a subset of primary breast tumors. The tumor (T; red coloured bars) and normal breast tissue (N; grey coloured bars) samples are matched, i.e. derived from the same individual and arranged side-by-side on the x-axis with ALKBH3 promoter methylation for each sample represented as a bar. The standard deviation is shown (line extensions from the bars). The *P*-value shown was derived from a paired Wilcoxon's hypothesis test ($P = 0.00012$). **b** Box and whisker plot of ALKBH3 mRNA expression in normal breast tissue samples and primary breast tumors. The *P*-value shown was derived from a Wilcoxon's hypothesis test ($P = 0.014$). **c** ALKBH3 promoter methylation data plotted on x-axis and mRNA expression data on y-axis for primary breast tumors. The *P*-value shown was derived from Spearman's correlation testing (Spearman's rho = −0.577; $P = 0.024$)

with an increased burden of unrepaired alkylation DNA damage involving 3-methylcytosine (3meC) suggesting that our findings has biological consequences. If left unrepaired these modifications can lead to alkylation-induced cell death or can be converted into mutations through the use of error-prone translesion DNA polymerases such as Pol η (eta), Pol ι (iota), and Pol κ (kappa) [31, 32]. Collectively, our observations support the hypothesis that defective repair of alkylation damage occurs in the development of a substantial fraction of breast cancers.

Our results furthermore emphasize the importance of quantification-based methods for DNA methylation analyses in clinical applications. This becomes clear by looking at varied levels of CpG methylation over the ALKBH3 promoter with respect to survival outcomes. Here, increasingly higher levels of promoter methylation were associated with shortened patient survival. Declaring tumours as either ALKBH3 promoter methylated or unmethylated is therefore not a straightforward task. In this study, two different methods are applied for this purpose, i.e. 1) making use of normal tissue samples as reference to then define "abnormally" high levels of ALKBH3 promoter methylation in tumours and 2) by identifying clinically relevant levels of ALKBH3 promoter methylation. Making use of normal breast samples as a reference, does not necessarily provide a means to identify tumours showing ALKBH3 inactivation events. This is because only slightly elevated, but still "abnormal", promoter methylation might simply reflect passenger events of no relevance to the course of disease progression. The second

Table 1 Clinical and pathological characteristics of the patient cohort

	Positive	*Negative*		
Estrogen-Receptor	175 (72%)	68 (28%)		
Progestrone-Receptor	136 (56%)	108 (44%)		
HER2-positive (over-expressed)	20 (14%)	121 (86%)		
Ki67-positive (>14% positivity)	84 (60%)	55 (40%)		
	1986–1991	*1991–1996*	*1996–2001*	*2001–2004*
Year of Diagnosis	81 (30%)	121 (46%)	60 (23%)	3 (1%)
	26–42	*42–58*	*58–74*	*74–91*
Age at Diagnosis	36 (14%)	104 (39%)	80 (30%)	45 (17%)
	5–30	*30–55*	*55–80*	*80–100*
Tumour size (mm)	94 (77%)	24 (20%)	1 (1%)	3 (2%)
Nodal status	112 (64%)	63 (36%)		
	+	++	+++	
Histological Grade	19 (13%)	62 (41%)	69 (46%)	

method refers to making use of clinical parameters involving tumour phenotype or disease survival to then declare a "clinically relevant" threshold for defining ALKBH3 promoter methylated tumours. Given the observed impact on disease progression, this second method is more likely to hold relevant at least as a proxy for identifying tumours affected by ALKBH3 promoter methylation. Using this definition, we report the incidence of ALKBH3 promoter methylation at approximately 5% in our cohort (13 out of 265 primary breast cancers).

DNA repair deficiency in breast cancer

Defective DNA repair capacity is frequently observed in various human cancers [14]. In breast cancer, this mostly

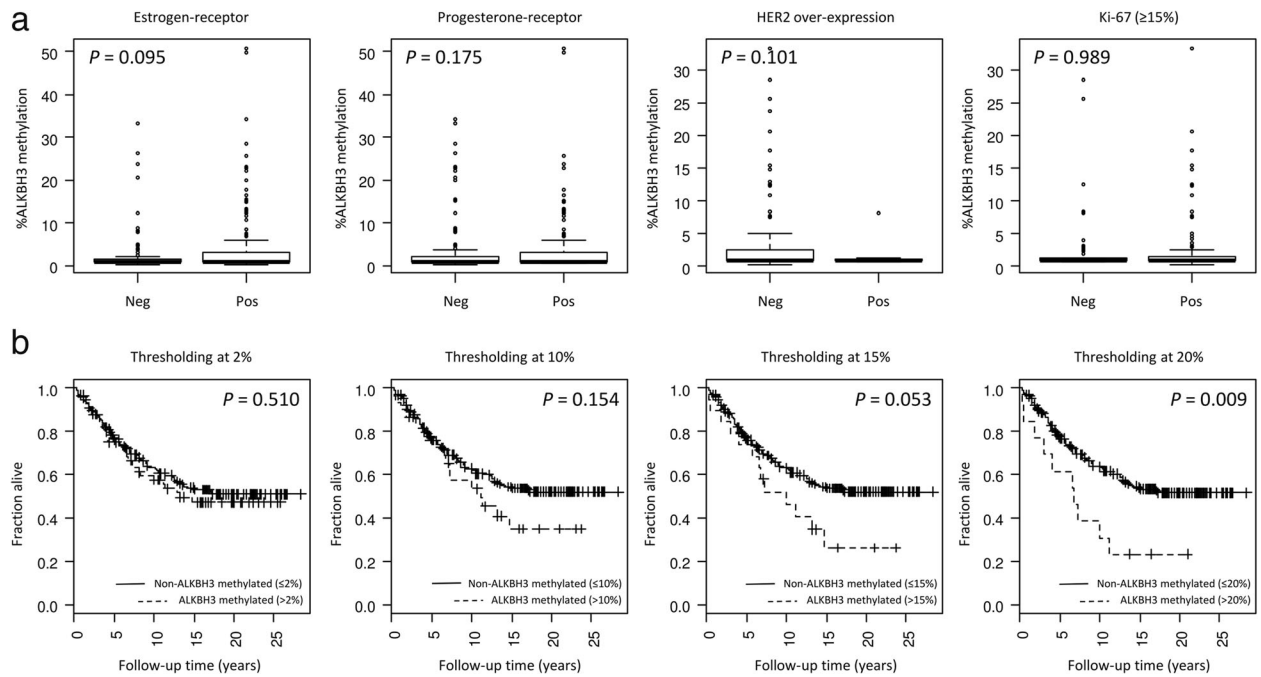


Fig. 5 ALKBH3 promoter methylation with respect to clinical parameters and breast cancer-specific survival. **a** ALKBH3 promoter methylation analysed with respect to clinical and subtype-specific markers: estrogen-receptor, progesterone-receptor, HER-2 over-expression and Ki-67. The *P*-values shown were derived from Wilcoxon's hypothesis testing. **b** ALKBH3 promoter methylation analysed with respect to patient survival (breast cancer-specific survival). The analyses shown were carried out using increasingly higher threshold levels for ALKBH3 promoter methylation as indicated. The *P*-values were derived from log-rank hypothesis testing for differences in survival outcomes

involves defective repair of DNA double-stranded breaks by homologous recombination (HR), e.g. as is known to occur in tumor cells arising in carriers of either BRCA1 or BRCA2 germline mutations [33]. Other potential sources of DNA repair deficiency in breast cancer, also involving DSB repair processes, include germline mutations in DNA repair genes PALB2, FANCM, ATM, CHEK2 (and possibly also RAD51C) characterized as loss-of-function variants [27, 34, 35]. Additionally, somatic mutations occur in both BRCA1 and BRCA2 as well as the ATM gene accounting for a small fraction of patients [27].

In addition to genetic mutations, our previous studies along with others have shown that the BRCA1 gene undergoes CpG promoter methylation in at least 5–10% of all sporadically arising breast tumors [6]. More recently, RAD51C (a RAD51 paralog involved in double-strand break repair) and the RAD51 recombinase have been identified as targets of epigenetic silencing in breast cancer [8, 36, 37]. Other contributors include deregulated expression of miRNAs including miR-182 and miR-146a/b targeting BRCA1 mRNA transcripts for degradation [38, 39].

In summary, the development of breast cancer is clearly linked to inactivation of genes involved in the repair of DNA double-stranded breaks by HR. The involvement of other DNA repair pathways in breast cancer development is, however, currently unclear. Of these, DNA repair of alkylation damage by direct reversal has been suggested following the identification of CpG promoter methylation of the MGMT gene in breast cancer [17]. Our results support this notion by identifying the ALKBH3 gene as a novel addition to the catalogue of DNA repair genes found inactivated in breast cancer.

The functional consequences of ALKBH3 epigenetic silencing

According to our results, ALKBH3 represents a candidate tumor suppressor gene. A high burden of mutations, caused by ineffective DNA repair, is generally accepted as an important factor in cancer development. Incomplete removal of alkyl groups on DNA has been shown to cause DNA damage, cell cycle arrest and apoptosis. Indeed, cancer genome sequencing has already established an important role for alkylation damage leading to genetic mutations [20]. Consequently, it seems likely that ALKBH3 mediates its tumor suppressive function via its role in DNA alkylation repair. This is supported by the increased level of alkylation damage in ALKBH3 inactivated cells, reported here and in previous studies [30]. It can, however, not be ruled out that other described functions of ALKBH3 also contribute to its anti-tumor activities – including alkylation repair of RNA and the recently described link to regulation of 1-methyladenine mark in mRNA [29, 40–42].

While our data suggest that ALKBH3 has tumor suppressor properties in breast cancer, other research has shown ALKBH3 overexpression in various cancer types. ALKBH3 overexpression is found in prostate cancer [43] and other cancer types [44–48]. In this context, ALKBH3 overexpression likely relates to adaptation of cancer cells to tolerate endogenous alkylation damage to DNA or RNA. This interpretation is in fact relevant in the context of a recent pre-clinical study showing resistance to alkylating drug temozolomide following experimental overexpression of MGMT [49]. Similar results have also been described with respect to ALKBH2 [50].

Defective repair of alkylation damage and precision medicine

Precision medicine is aimed at optimizing treatment benefits by looking at each patient in terms of genomic or epigenomic variants and, on the basis of this information, to then select the most appropriate combination of cytotoxic or targeted drugs. This is highly relevant with respect to CpG promoter methylation of the MGMT gene now widely recognized as a predictor for patient response to alkylating agents such as temozolomide [19, 51]. Indeed, temozolomide induces methylation damage in DNA involving O6-methylguanine along with N7-methylguanine. In this way, tumor cells lacking the MGMT repair gene as a result of CpG promoter methylation events are highly sensitive to temozolomide [19]. Similarly, in breast cancer, loss of MGMT was recently linked to temozolomide sensitivity in a pre-clinical study [49]. Whether the same principle can be applied with respect to the ALKBH3 gene, i.e. by inducing 3meC and 1meA in patients having developed breast tumors with ALKBH3 promoter methylation, remains to be determined.

Conclusions

We propose here that the ALKBH3 gene is a novel addition to the catalogue of DNA repair genes found inactivated in breast cancer. The value of ALKBH3 promoter methylation as a prognostic marker was revealed through quantification of CpG methylation levels by pyrosequencing. In this way, our results emphasize the use of quantification-based methods for the assessment of CpG methylation marks in clinical applications. In our cohort, clinically relevant ALKBH3 promoter methylation occurs in at least 5% of all breast cancers and, although independent cohorts will be needed for confirmation, this event appears to be associated with highly aggressive disease behaviour. These observations underscore defective repair of alkylation damage occurring in the development of breast cancer.

Additional files

Additional file 1: The relation between CpG methylation and expression for the ALKBH3 gene shown here based on available data from the Cancer Genome Atlas project. This catalogue includes only CpG's found differentially methylated between breast cancers and normal breast tissue samples. The table lists out statistically significant CpG's with rho < −0.30 and at least two-fold change in expression between unmethylated and methylated tumours. (XLSX 10 kb)

Additional file 2: RNAi for ALKBH3 analysed with respect to 3-methyl-cytosine immunostaining. A) U2OS cells were transfected with a control (scrambled siRNA) and ALKBH3 siRNA for 72 h. After fixation cells were denatured in 1.5 M HCL for 30 min (to gain access to single stranded DNA) and immunostained for 3-me-C and 5-me-C. As expected, decreased ALKBH3 expression resulted in increased accumulation of 3-me-C, indicating less efficient repair, without any detectable changes in 5-me-C levels. B) Quantification of at least 100 cells reveals approximately 1, 6-fold differences in 3-me-C damages based on nuclear intensity measured using CellProfiler. The fold differences are computed as the average signal derived form ALKBH3 siRNA treated cells divided by the average signal derived from scrambled siRNA treated cells (by default set to one in the figure). 3-me-C = 3-methyl-cytosine, 5-me-C = 5-methyl-cytosine. (TIFF 3866 kb)

Additional file 3: ALKBH3 methylation analysed with respect to prognostic parameters (breast cancer subtype, histological grade, tumour size and nodal status). (XLSX 12 kb)

Abbreviations

1meA: 1-methyladenine; 3meC: 3-methylcytosine; 5meC: 5-methylcytosine; ALKBH3: alkB homolog 3, alpha-ketoglutarate-dependent dioxygenase; MMS: Methyl methanesulfonate; qPCR: Quantitative Polymerase Chain Reaction; TCGA: The Cancer Genome Atlas; vHMEC: Variant human mammary epithelial cells

Acknowledgements

We thank Vilmundur Gudnason and his team at the Icelandic Heart Association (Hjartavernd), Kopavogur, Iceland, for generously providing use of their PyroMark Q24 machine.

Funding

The authors would like to thank The Icelandic Centre for Research (www.rannis.is) grant IDs #14193–051 and #152077–051, Gongum Saman (www.gongumsaman.is) and Minningarsjodur Eggerts Briem for financial support.

Availability of data and materials

Methylome datasets used in this study: Stefansson et al. (GSE52865), Dedeurwaerder et al. (GSE20713) and Teschendorff et al. (GSE69914).

Authors' contributions

OAS was responsible for the pyrosequencing and qPCR analyses (assay design, execution and data analysis) and performed statistical and informatics analyses. SH carried out pyrosequencing for the ALKBH3 gene in breast cancer cell lines. ZS and JH performed the qPCR expression analyses for ALKBH3 in the breast cancer cell lines. TG and SH performed the western blotting assay for ALKBH3 in cell lines. TG developed and performed the 3meC damage assay. HH and SH carried out pyrosequencing in primary breast tumors and normal breast tissue samples. JH carried out qPCR analyses in RNA samples derived from primary breast tumors and normal breast tissue samples. JGJ and LT contributed information on clinical parameters. OAS and JGJ were responsible for the TMA expression analyses. OAS and SS conceived of the study and were in charge of its design, coordination and writing of the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This work was carried out according to permits from the Icelandic Data Protection Commission (2006050307) and Bioethics Committee (VSNb2006050001/03–16). Informed consent (written) was obtained from all patients.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Cancer Research Laboratory, Biomedical Center, Vatnsmyrarvegur 16 (4th floor), 101 Reykjavik, Iceland. [2]Faculty of Medicine, University of Iceland, Vatnsmyrarvegur 16 (4th floor), 101 Reykjavik, Iceland. [3]Department of Biochemistry and Molecular Biology, Biomedical Center, Vatnsmyrarvegur 16 (5th floor), 101 Reykjavik, Iceland. [4]Icelandic Cancer Registry, Skogarhlid 8, Reykjavik, Iceland. [5]Department of Pathology, Landspitali University Hospital, Reykjavik, Iceland.

Received: 13 September 2016 Accepted: 26 June 2017
Published online: 05 July 2017

References

1. Cutter AR, Hayes JJ. A brief Review of nucleosome structure. FEBS Lett. 2015;589(20 Pt A):2914–22.
2. Hesson LB, Sloane MA, Wong JW, Nunez AC, Srivastava S, Ng B, et al. Altered promoter nucleosome positioning is an early event in gene silencing. Epigenetics. 2014;9(10):1422–30.
3. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484–92.
4. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011;11(10):726–34.
5. Simó-Riudalbas L, Esteller M. Cancer genomics identifies disrupted epigenetic genes. Hum Genet. 2014;133(6):713–25.
6. Stefansson OA, Jonasson JG, Olafsdottir K, Hilmarsdottir H, Olafsdottir G, Esteller M, et al. CpG island hypermethylation of BRCA1 and loss of pRb as co-occurring events in basal/triple-negative breast cancer. Epigenetics. 2011; 6(5):638–49.
7. Stefansson OA, Moran S, Gomez A, Sayols S, Arribas-Jorba C, Sandoval J, et al. A DNA methylation-based definition of biologically distinct breast cancer subtypes. Mol Oncol. 2015;9(3):555–68.
8. Cunningham JM, Cicek MS, Larson NB, Davila J, Wang C, Larson MC, et al. Clinical characteristics of ovarian cancer classified by BRCA1, BRCA2, and RAD51C status. Sci Rep. 2014;4:4026.
9. Muggerud AA, Rønneberg JA, Wärnberg F, Botling J, Busato F, Jovanovic J, et al. Frequent aberrant DNA methylation of ABCB1, FOXC1, PPP2R2B and PTEN in ductal carcinoma in situ and early invasive breast cancer. Breast Cancer Res. 2010;12(1):R3.
10. Subramaniam MM, Chan JY, Soong R, Ito K, Ito Y, Yeoh KG, et al. RUNX3 inactivation by frequent promoter hypermethylation and protein mislocalization constitute an early event in breast cancer progression. Breast Cancer Res Treat. 2009;113(1):113–21.
11. Lalloo F, Evans DG. Familial breast cancer. Clin Genet. 2012;82(2):105–14.
12. Collins N, Wooster R, Stratton MR. Absence of methylation of CpG dinucleotides within the promoter of the breast cancer susceptibility gene BRCA2 in normal tissues and in breast and ovarian cancers. Br J Cancer. 1997;76(9):1150–6.
13. Brandes JC, Carraway H, Herman JG. Optimal primer design using the novel primer design program: MSPprimer provides accurate methylation analysis of the ATM promoter. Oncogene. 2007;26(42):6229–37.
14. Lord CJ, Ashworth A. BRCAness revisited. Nat Rev Cancer. 2016;16(2):110–20.
15. Saal LH, Gruvberger-Saal SK, Persson C, Lövgren K, Jumppanen M, Staaf J, et al. Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. Nat Genet. 2008;40(1):102–7.
16. Spitzwieser M, Holzweber E, Pfeiler G, Hacker S, Cichna-Markl M. Applicability of HIN-1, MGMT and RASSF1A promoter methylation as biomarkers for detecting field cancerization in breast cancer. Breast Cancer Res. 2015;17(1):125.
17. Fumagalli C, Pruneri G, Possanzini P, Manzotti M, Barile M, Feroce I, et al. Methylation of O6-methylguanine-DNA methyltransferase (MGMT) promoter gene in triple-negative breast cancer patients. Breast Cancer Res Treat. 2012; 134(1):131–7.

18. Drabløs F, Feyzi E, Aas PA, Vaagbø CB, Kavli B, Bratlie MS, et al. Alkylation damage in DNA and RNA–repair mechanisms and medical significance. DNA Repair (Amst). 2004;3(11):1389–407.
19. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, OF H, Vanaclocha V, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med. 2000;343(19):1350–4.
20. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500(7463):415–21.
21. Sedgwick B, Bates PA, Paik J, Jacobs SC, Lindahl T. Repair of alkylated DNA: recent advances. DNA Repair (Amst). 2007;6(4):429–42.
22. Sigurdardottir LG, Jonasson JG, Stefansdottir S, Jonsdottir A, Olafsdottir GH, Olafsdottir EJ, Tryggvadottir L. Data quality at the Icelandic cancer Registry: comparability, validity, timeliness and completeness. Acta Oncol. 2012;51(7): 880–9.
23. Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol. 2015;16:22.
24. Stefansson OA, Jonasson JG, Johannsson OT, Olafsdottir K, Steinarsdottir M, Valgeirsdottir S, Eyfjord JE. Genomic profiling of breast tumors in relation to BRCA abnormalities and phenotypes. Breast Cancer Res. 2009;11(4):R47.
25. Dedeurwaerder S, Desmedt C, Calonne E, Singhal SK, Haibe-Kains B, Defrance M, et al. DNA methylation profiling reveals a predominant immune component in breast cancers. EMBO Mol Med. 2011;3(12):726–41.
26. Teschendorff AE, Gao Y, Jones A, Ruebner M, Beckmann MW, Wachter DL, et al. DNA methylation outliers in normal breast tissue identify field defects that are enriched in cancer. Nat Commun. 2016;7:10478.
27. Network CGA. Comprehensive molecular portraits of human breast tumors. Nature. 2012;490(7418):61–70.
28. Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317–30.
29. Ougland R, Rognes T, Klungland A, Larsen E. Non-homologous functions of the AlkB homologs. J Mol Cell Biol. 2015;7(6):494–504.
30. Dango S, Mosammaparast N, Sowa ME, Xiong LJ, Wu F, Park K, et al. DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. Mol Cell. 2011;44(3):373–84.
31. Goodman MF, Woodgate R. Translesion DNA polymerases. Cold Spring Harb Perspect Biol. 2013;5(10):a010363.
32. Waters LS, Minesinger BK, Wiltrout ME, D'Souza S, Woodruff RV, Walker GC. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. Microbiol Mol Biol Rev. 2009;73(1):134–54.
33. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC, Van Loo P, Ju YS, Smid M, Brinkman AB, Morganella S, Aure MR, Lingjærde OC, Langerød A, Ringnér M, Ahn SM, Boyault S, Brock JE, Broeks A, Butler A, Desmedt C, Dirix L, Dronov S, Fatima A, Foekens JA, Gerstung M, Hooijer GK, Jang SJ, Jones DR, Kim HY, King TA, Krishnamurthy S, Lee HJ, Lee JY, Li Y, McLaren S, Menzies A, Mustonen V, O'Meara S, Pauporte I, Pivot X, Purdie CA, Raine K, Ramakrishnan K, Rodríguez-González FG, Romieu G, Sieuwerts AM, Simpson PT, Shepherd R, Stebbings L, Stefansson OA, Teague J, Tommasi S, Treilleux I, Van den Eynden GG, Vermeulen P, Vincent-Salomon A, Yates L, Caldas C, Veer LV, Tutt A, Knappskog S, Tan BK, Jonkers J, Borg Å, Ueno NT, Sotiriou C, Viari A, Futreal PA, Campbell PJ, Span PN, Van Laere S, Lakhani SR, Eyfjord JE, Thompson AM, Birney E, Stunnenberg HG, van de Vijver MJ, Martens JW, Børresen-Dale AL, Richardson AL, Kong G, Thomas G, Stratton MR. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature. 2016; doi:10.1038/nature17676.
34. Ghoussaini M, Pharoah PD, Easton DF. Inherited genetic susceptibility to breast cancer: the beginning of the end or the end of the beginning? Am J Pathol. 2013;183(4):1038–51.
35. Peterlongo P, Catucci I, Colombo M, Caleca L, Mucaki E, Bogliolo M, et al. FANCM c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity and is a familial breast cancer risk factor. Hum Mol Genet. 2015;24(18):5345–55.
36. Chang CJ, Yang JY, Xia W, Chen CT, Xie X, Chao CH, et al. EZH2 promotes expansion of breast tumor initiating cells through activation of RAF1-beta-catenin signaling. Cancer Cell. 2011;19:86–100.
37. Stefansson OA, Esteller M. EZH2-mediated epigenetic repression of DNA repair in promoting breast tumor initiating cells. Breast Cancer Res. 2011; 13(3):309.
38. Moskwa P, Buffa FM, Pan Y, Panchakshari R, Gottipati P, Muschel RJ, et al. miR-182-mediated downregulation of BRCA1 impacts DNA repair and sensitivity to PARP inhibitors. Mol Cell. 2011;41(2):210–20.
39. Garcia AI, Buisson M, Bertrand P, Rimokh R, Rouleau E, Lopez BS, et al. Down-regulation of BRCA1 expression by miR-146a and miR-146b-5p in triple negative sporadic breast cancers. EMBO Mol Med. 2011;3(5):279–90.
40. Aas PA, Otterlei M, Falnes PO, Vågbø CB, Skorpen F, Akbari M, et al. Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature. 2003;421(6925):859–63.
41. Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, Yi C. Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome. Nat Chem Biol. 2016;
42. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, et al. The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature. 2016;530(7591):441–6.
43. Konishi N, Nakamura M, Ishida E, Shimada K, Mitsui E, Yoshikawa R, et al. High expression of a new marker PCA-1 in human prostate carcinoma. Clin Cancer Res. 2005;11(14):5090–7.
44. Tasaki M, Shimada K, Kimura H, Tsujikawa K, Konishi N. ALKBH3, a human AlkB homologue, contributes to cell survival in human non-small-cell lung cancer. Br J Cancer. 2011;104(4):700–6.
45. Yamato I, Sho M, Shimada K, Hotta K, Ueda Y, Yasuda S, et al. PCA-1/ALKBH3 contributes to pancreatic cancer by supporting apoptotic resistance and angiogenesis. Cancer Res. 2012;72(18):4829–39.
46. Cetica V, Genitori L, Giunti L, Sanzo M, Bernini G, Massimino M, Sardi I. Pediatric brain tumors: mutations of two dioxygenases (hABH2 and hABH3) that directly repair alkylation damage. J Neuro-Oncol. 2009;94(2):195–201.
47. Choi SY, Jang JH, Kim KR. Analysis of differentially expressed genes in human rectal carcinoma using suppression subtractive hybridization. Clin Exp Med. 2011;11(4):219–26.
48. Shimada K, Fujii T, Tsujikawa K, Anai S, Fujimoto K, Konishi N. ALKBH3 contributes to survival and angiogenesis of human urothelial carcinoma cells through NADPH oxidase and tweak/Fn14/VEGF signals. Clin Cancer Res. 2012;18(19):5247–55.
49. Palmieri D, Duchnowska R, Woditschka S, Hua E, Qian Y, Biernat W, et al. Profound prevention of experimental brain metastases of breast cancer by temozolomide in an MGMT-dependent manner. Clin Cancer Res. 2014; 20(10):2727–39.
50. Johannessen TC, Prestegarden L, Grudic A, Hegi ME, Tysnes BB, Bjerkvig R. The DNA repair protein ALKBH2 mediates temozolomide resistance in human glioblastoma cells. Neuro-Oncology. 2013;15(3):269–78.
51. Weller M, Tabatabai G, Kästner B, Felsberg J, Steinbach JP, Wick A, et al. MGMT promoter Methylation is a strong prognostic biomarker for benefit from dose-intensified Temozolomide Rechallenge in progressive Glioblastoma: the DIRECTOR trial. Clin Cancer Res. 2015;21(9):2057–64.