

CONSEQUENCES OF PERICENTROMERIC DNA HYPOMETHYLATION:  
LESSONS FROM AN ANIMAL MODEL OF ICF SYNDROME

A Dissertation

Presented to the Faculty of the Weill Cornell Graduate School  
of Medical Sciences

in Partial Fulfillment of the Requirements for the degree of  
Doctor of Philosophy

By

Srivarsha Rajshekar

May 2019

© 2019 Srivarsha Rajshekar

CONSEQUENCES OF PERICENTROMERIC DNA HYPOMETHYLATION:  
LESSONS FROM AN ANIMAL MODEL OF ICF SYNDROME

Srivarsha Rajshekar, Ph.D.

Cornell University 2019

The modified base, 5-methylcytosine (5mC) is enriched at repetitive DNA sequences including satellite repeats that surround chromosome centromeres. These centromeric and pericentromeric satellite repeats are important for stable chromosome structure and proper chromosome segregation. Loss of 5mC at pericentromeric repeats is common in cancer and senescence. While the general importance of 5mC is well-established, the specific functions of 5mC at pericentromeres are less clear. 5mC loss at pericentromeric repeats is a molecular hallmark of the rare genetic disease Immunodeficiency, Centromere instability and Facial abnormalities (ICF) syndrome. To date, attempts to model specific loss of 5mC at pericentromeres in mouse through mutation of ICF associated genes have been unsuccessful. Here, I develop a zebrafish model for ICF syndrome by mutating the zebrafish ortholog of *ZBTB24*, a poorly characterized gene that is disrupted in ~30% of ICF patients. *zbtb24* mutant zebrafish recapitulate key features of ICF syndrome including immunodeficiency, facial abnormalities, gastrointestinal defects, impaired growth and reduced lifespan. I also show that homozygous mutation of *zbtb24* causes a progressive loss of 5mC at pericentromeric satellite repeats in zebrafish. This progressive loss of methylation allowed for elucidation of primary vs secondary consequences of hypomethylation at these sequences. Transcriptome analysis revealed that one of the earliest consequences of

pericentromeric hypomethylation was activation of an interferon-based innate immune response. Mechanistically, I tie this response to derepression of pericentromeric satellite transcripts and I demonstrate that these aberrant transcripts are recognized through the MDA5-MAVS dsRNA-sensing machinery, which is normally associated with an innate immune response to viruses. Additional preliminary studies indicate increased incidence of DNA damage and tumor formation in *zbtb24* mutants suggesting that pericentromeric 5mC is likely important for genome stability. Taken together, this thesis describes the first viable animal model of ICF Syndrome, reveals a function for ICF-gene *zbtb24* in the long-term maintenance of pericentromeric DNA methylation and identifies roles for pericentromeric DNA methylation in preventing autoimmunity and maintaining genome integrity.

## **BIOGRAPHICAL SKETCH**

Srivarsha Rajshekar was born in Gwalior, India and grew up in Air Force stations all across the country. She graduated high school from Padma Seshadri Bala Bhavan Senior Secondary School, Chennai in 2007 and went on to pursue an Integrated B.S and M.S in Biological Sciences at the Birla Institute of Technology and Science, Pilani. She did her Master's Thesis research work at the laboratory of Dr. Carl-Philipp Heisenberg at the Max Planck Institute of Molecular Cell Biology in Dresden, Germany. In the fall of 2012, she began her PhD at the BCMB Allied program at Weill Cornell Graduate School of Medical Sciences in New York City. She joined the lab of Dr. Mary Goll who was at the Developmental Biology department of the Sloan Kettering Institute in New York and is currently at the University of Georgia in Athens, GA. For her graduate research, Varsha has investigated the effect of epigenetic changes at pericentromeric DNA through developing an animal model of a rare human disease called ICF Syndrome.

*To my grandparents, for each of their unique influences on my life.*

*To MSN Thatha, for teaching me the virtue of self-discipline,*

*To Ramu Thatha, for reminding me to live the happiest version of myself,*

*To Bala Patti, for breaking glass ceilings since 1974,*

*To Kamala Patti, for inspiring me to stay curious and to never let go of  
my inner child.*

## ACKNOWLEDGEMENTS

First and foremost, I want to extend my most sincere gratitude to my PhD mentor, Dr. Mary Goll. I am so grateful to her for giving me the opportunity to work on an exciting thesis project and for her commitment to training me as a thorough scientist as I navigated my way through this project. I am inspired by her intelligence, grit and professionalism as a scientist and her compassion as a mentor. I am grateful for the attention she has given me to hone my skills in scientific presentation and writing. I have learnt valuable life lessons from her and I am sure I will fall back on them throughout my scientific career.

Next, I would like to thank my friends and colleagues at the Goll lab both at Memorial Sloan Kettering Cancer Center and at the University of Georgia for their feedback and discussions which have been helpful in my research. I will particularly acknowledge support from postdocs in our lab, Kathrin Laue and Jonathan Schneiderman. I am grateful to Jun Yao and Paige Arnold for their contribution to my project. I especially want to thank my lab mate Cheng Li, who has been a brother to me and always supported me both scientifically and emotionally.

I would like to thank the members of my dissertation committee: Dr. Todd Evans, Dr. Iestyn Whitehouse, Dr. Joseph Sun for all their support, help and suggestions regarding my project as well as their guidance for my future. I am grateful to our collaborators, Dr. John Edwards, Dr. Teresa Bowman, Dr. Robert Schmitz, Sara Payne and Yinwen Zhang, for their contribution to my project.

Special thanks to Dr. Andy Koff, who has always given me valuable advice every time I have been at the crossroads of a crucial career decision.

I am grateful to my graduate school, the Weill Cornell Graduate School of Medical Sciences, for setting up such a wonderful environment in the great city of New York. A very special thanks to the Outdoor club of Weill Cornell, I will fondly remember our many adventures and near-death experiences. A special shout out to CrossFit Liberate in Athens, Georgia for contributing to my physical and mental health in the final year of my PhD.

I am deeply thankful to all the friends I have made and re-made these past 6.5 years of my life. They have always been there for me in difficult times, exuberant ones and everything in between.

Thank you to my family, old relatives and new. You have loved me and fed me abundantly, reminding me that I'm not so far away from home after all.

Thank you, Kris, my best friend, my confidante, my rock. Thank you for your perennial optimism in times of setbacks and your eternal excitement for my victories.

So much love and gratitude for my parents, my loudest cheerleaders.

## TABLE OF CONTENTS

<b>BIOGRAPHICAL SKETCH</b> .....	<b>iii</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>v</b>
<b>LIST OF FIGURES</b> .....	<b>x</b>
<b>LIST OF ABBREVIATIONS</b> .....	<b>xii</b>
<b>Chapter 1 : INTRODUCTION</b> .....	<b>1</b>
<b>DNA CYTOSINE METHYLATION</b> .....	<b>1</b>
The epigenetic modification 5mC.....	1
Enzymes involved in establishment and maintenance of DNA methylation	2
.....	2
Distribution of DNA methylation in mammalian genomes.....	4
DNA methylation in Zebrafish.....	7
<b>PERICENTROMERES</b> .....	<b>9</b>
Pericentromeric Satellite repeats.....	9
Pericentromeric Chromatin.....	11
Diseases associated with pericentromeric DNA hypomethylation.....	12
<b>ICF SYNDROME, A MODEL FOR PERICENTROMERIC DNA</b>	
<b>HYPOMETHYLATION</b> .....	<b>14</b>
ICF Syndrome Symptoms.....	14
Genetic Basis of ICF Syndrome.....	15
Epigenomic Assessment of ICF Syndrome.....	18
Molecular Basis of ICF Syndrome.....	19
Animal Models of ICF Syndrome.....	20
<b>DISSERTATION OVERVIEW</b> .....	<b>21</b>
<b>Chapter 2 : DEVELOPING AN ANIMAL MODEL OF ICF SYNDROME</b> .....	<b>23</b>
<b>PREFACE</b> .....	<b>23</b>
<b>EXPRESSION ANALYSIS OF ZEBRAFISH <i>zbtb24</i></b> .....	<b>26</b>
<b>MORPHOLINO-MEDIATED KNOCKDOWN OF <i>Zbtb24</i></b> .....	<b>26</b>
<b>DESCRIPTION OF <i>zbtb24</i> MUTANT ALLELES</b> .....	<b>30</b>
<i>zbtb24</i> <sup>mk22</sup> .....	31
<i>zbtb24</i> <sup>mk19</sup> .....	31
<i>zbtb24</i> <sup>mk21</sup> .....	31
<b><i>zbtb24</i> MUTANTS RECAPITULATE KEY PHENOTYPES OF ICF</b>	
<b>SYNDROME</b> .....	<b>34</b>
Growth defects.....	34
Immunodeficiency.....	34
Facial Abnormalities.....	37

Gastrointestinal defects.....	37
Infertility .....	40
<b>DISCUSSION.....</b>	<b>43</b>
<b>Chapter 3 : EPIGENETIC ANALYSIS OF AN ANIMAL MODEL OF ICF SYNDROME.....</b>	<b>45</b>
<b>PREFACE.....</b>	<b>45</b>
<b>PERICENTROMERIC DNA METHYLATION ANALYSIS.....</b>	<b>46</b>
Pericentromeric satellite repeats are hypomethylated in <i>zbtb24</i> mutants	46
Progressive hypomethylation at pericentromeric satellite repeats .....	49
<b>GENOME-WIDE METHYLOME ANALYSIS .....</b>	<b>51</b>
<b>HISTONE MODIFICATION ANALYSIS .....</b>	<b>57</b>
<b>DISCUSSION.....</b>	<b>58</b>
<b>Chapter 4 : PERICENTROMERIC HYPOMETHYLATION ELICITS AN INTERFERON RESPONSE .....</b>	<b>61</b>
<b>PREFACE.....</b>	<b>61</b>
<b>INTERFERON RESPONSE IN <i>zbtb24</i> MUTANTS .....</b>	<b>61</b>
<b>MECHANISM FOR INTERFERON RESPONSE IN <i>zbtb24</i> MUTANTS .....</b>	<b>64</b>
Introduction to interferon response .....	64
The innate immune response in <i>zbtb24</i> mutants is mediated by sensors of cytosolic RNA.....	67
Satellite RNA is upregulated in <i>zbtb24</i> mutants .....	70
Satellite RNA is sufficient to trigger an interferon response .....	72
The cytosolic RNA helicase Mda5 is required for the Interferon Response in <i>zbtb24</i> mutants.....	74
Co-mutations in Interferon activation genes do not rescue ICF Syndrome Phenotypes in <i>zbtb24</i> mutants .....	76
<b>DISCUSSION.....</b>	<b>78</b>
<b>Chapter 5 : PERICENTROMERIC DNA HYPOMETHYLATION AND GENOME INSTABILITY .....</b>	<b>81</b>
<b>PREFACE.....</b>	<b>81</b>
<b>EVIDENCE OF DNA DAMAGE AND CANCER INCIDENCE.....</b>	<b>84</b>
<b>DISCUSSION.....</b>	<b>87</b>
<b>Chapter 6 : SUMMARY AND PERSPECTIVES.....</b>	<b>89</b>
Animal model of ICF Syndrome .....	89
Function of Zbtb24 .....	90

Pericentromeric hypomethylation, genome instability and the interferon response .....	90
<b>MATERIALS AND METHODS.....</b>	<b>92</b>
Zebrafish husbandry .....	92
TALEN and CRISPR mutagenesis.....	92
Zebrafish imaging and length measurements .....	93
FACS Analysis of Whole Kidney Marrow .....	93
Morpholino oligonucleotide (MO) knockdown .....	94
Whole Mount in situ Hybridization (WISH) .....	94
DNA Methylation Analysis .....	94
Enhanced Reduced Representation Bisulfite Sequencing (ERRBS).....	95
RNA Expression Analysis .....	95
Transcriptome sequencing .....	96
RNA synthesis and injections.....	97
Histology .....	98
Sperm Count .....	98
Chromatin Immunoprecipitation (ChIP).....	98
5-aza-dC Treatment .....	99
Immunofluorescence .....	99
Microgavage.....	100
Sudan Black staining.....	100
Statistical Analysis .....	100
<b>REFERENCES.....</b>	<b>101</b>
<b>APPENDIX 1: LIST OF MUTANT ALLELES.....</b>	<b>115</b>
<b>APPENDIX 2: LIST OF PRIMERS.....</b>	<b>116</b>
<b>APPENDIX 3: LIST OF OLIGOS .....</b>	<b>121</b>
<b>APPENDIX 4: eLife PAPER .....</b>	<b>123</b>

## LIST OF FIGURES

<i>Figure 1.1: DNA Methylation</i> .....	1
<i>Figure 1.2: Schematic of the functions of DNA methyltransferase (Dnmt) enzymes</i> .....	3
<i>Figure 1.3: Distribution of DNA methylation in mammalian genomes.</i> .....	5
<i>Figure 1.4: Pericentromeric satellite repeats.</i> .....	10
<i>Figure 1.5: Genes mutated in ICF Syndrome.</i> .....	18
<i>Figure 1.6: Molecular Basis of ICF Syndrome.</i> .....	20
<i>Figure 2.1: Zebrafish Zbtb24 is conserved.</i> .....	24
<i>Figure 2.2: zbtb24 is maternally deposited and expressed at low-levels during early embryonic zebrafish development.</i> .....	27
<i>Figure 2.3: Morpholino-mediated knockdown of Zbtb24 causes developmental abnormalities in zebrafish embryos.</i> .....	28
<i>Figure 2.4: Morpholino-mediated knockdown of Zbtb24 causes modest hypomethylation of pericentromeric Sat1 repeats.</i> .....	30
<i>Figure 2.5: Generating a large deletion in the zebrafish ortholog of zbtb24 (zbtb24<sup>mk22(Δ)</sup>)</i> .....	32
<i>Figure 2.6: Additional zbtb24 mutant alleles (zbtb24<sup>mk19</sup> and zbtb24<sup>mk21</sup>)</i> .....	33
<i>Figure 2.7: Growth defects in zbtb24 mutants.</i> .....	35
<i>Figure 2.8: Immunodeficiency and Facial anomalies in zbtb24 mutants.</i> .....	38
<i>Figure 2.9: Gastrointestinal defects in zbtb24 mutants.</i> .....	39
<i>Figure 2.10: Fertility and gonad analysis in zbtb24 mutants.</i> .....	41
<i>Figure 3.1: Pericentromeric satellite repeats are hypomethylated in zbtb24 mutants.</i> .....	47
<i>Figure 3.2: Somatic tissues but not germ cells are similarly hypomethylated at pericentromeric Sat1 repeats in zbtb24 mutants.</i> .....	50
<i>Figure 3.3: zbtb24 mutation causes progressive methylation loss at pericentromeric Sat1 repeats in zbtb24 mutants.</i> .....	51
<i>Figure 3.4: zbtb24 mutants exhibit modest reductions in 5mC at non-pericentromeric sequences.</i> .....	54

<i>Figure 3.5: DNA Methylation levels at different genomic classes in fins of zbtb24<sup>+/+</sup> and zbtb24<sup>Δ/Δ</sup> zebrafish.</i>	55
<i>Figure 3.6: Methylation at interspersed repeats is unaffected in zbtb24 mutants.</i>	56
<i>Figure 3.7: Histone Modifications are unaffected in zbtb24<sup>Δ/Δ</sup> mutants.</i>	57
<i>Figure 4.1: cdca7a is downregulated in zbtb24 mutants.</i>	62
<i>Figure 4.2: Mutation of zbtb24 leads to activation of innate immune response genes.</i>	63
<i>Figure 4.3: Methylation at promoters of innate immune genes upregulated in zbtb24 mutants is unchanged.</i>	64
<i>Figure 4.4: Schematic for activation of interferon response by nucleic acids.</i>	65
<i>Figure 4.5: Mutation of zebrafish orthologs of mavs and sting.</i>	68
<i>Figure 4.6: Interferon response in zbtb24 mutants is mediated by sensors of cytosolic RNA.</i>	69
<i>Figure 4.7: Pericentromeric Sat1 transcripts are upregulated in zbtb24 mutants.</i>	71
<i>Figure 4.8: Mutation in zbtb24 upregulates Sat1 transcripts but not transposons.</i>	72
<i>Figure 4.9: Sat1 RNAs are sufficient to elicit an interferon response.</i>	73
<i>Figure 4.10: Mutation of cytosolic dsRNA receptor Mda5 mitigates the interferon response in zbtb24<sup>Δ/Δ</sup> zebrafish.</i>	75
<i>Figure 4.11: Mitigation of interferon response in zbtb24 mutants does not rescue ICF syndrome phenotypes.</i>	77
<i>Figure 4.12: Model for the activation of interferon response in zbtb24 mutants.</i>	78
<i>Figure 5.1: Gene mutation analysis of ZBTB24 in human cancers.</i>	83
<i>Figure 5.2: Increased incidence of DNA damage in zbtb24 mutants.</i>	85
<i>Figure 5.3: Increased incidence of eye tumors in zbtb24<sup>Δ/Δ</sup>; tp53<sup>-/-</sup> zebrafish.</i>	87

## LIST OF ABBREVIATIONS

5mC	5 methyl-cytosine
CDCA7	Cell division cycle associated 7
cGAS	Cyclic GMP-AMP synthase
COBRA	Combined Bisulphite Restriction Analysis
CpG	Cytosine-guanine dinucleotide
DMR	Differentially Methylated Region
DNMT	DNA Methyltransferase
dpf	Days post fertilization
dsRNA	Double stranded RNA
ERRBS	Enhanced Reduced Representation Bisulphite Sequencing
ERV	Endogenous retroviruses
H3K9me3	Histone H3 lysine 9 trimethyl
HELLS	Helicase, lymphoid-specific
ICF	Immunodeficiency, Centromeric Instability and Facial Anomalies
IRF	Interferon Regulatory Factor
ISG	Interferon stimulated genes
LINE	Long Interspersed Nuclear Elements
LTR	Long Terminal Repeats
MAVS	Mitochondrial antiviral-signaling protein
MDA5	Melanoma Differentiation-Associated protein 5;
Myd88	Myeloid differentiation primary response 88
NHEJ	Non-Homologous End Joining
PAMP	Pathogen-associated molecular patterns
PRR	Pattern-recognition receptors
RIG-I	retinoic acid-inducible gene I
RLR	RIG-I-like Receptors
Sat1	Satellite-1
SINE	Short Interspersed Nuclear Elements
STING	Stimulator of interferon genes

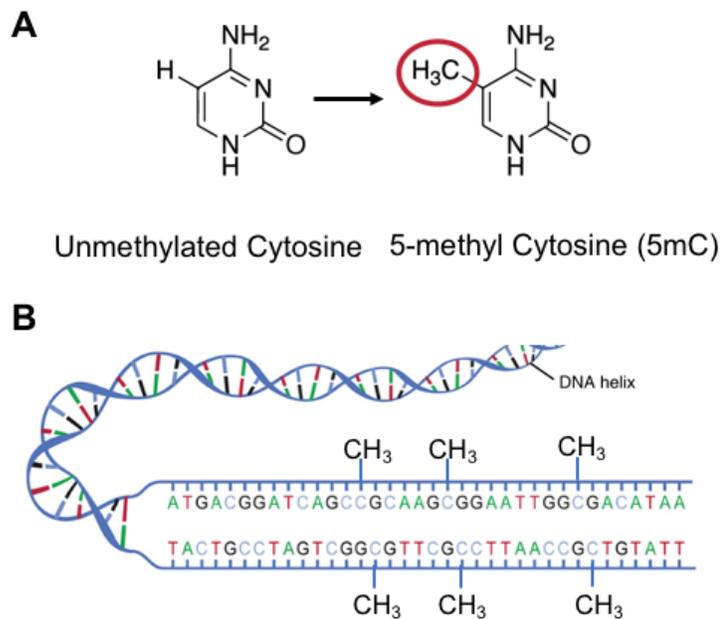
TALEN	Transcription activator-like effector nucleases
TASA-TD	TAG-aided sense/antisense transcript detection
TLR	Toll-Like Receptor
TRIF	TIR-domain-containing adapter-inducing interferon- $\beta$
UHRF1	Ubiquitin-like with PHD and ring finger domains 1
wpf	Weeks post fertilization
ZBTB24	Zinc-finger and BTB domain containing 24

## Chapter 1 : INTRODUCTION

### DNA CYTOSINE METHYLATION

#### The epigenetic modification 5mC

Methylation at the fifth position of the cytosine ring of DNA (5-methyl Cytosine, 5mC) is a well-conserved feature of many plant, animal and some fungal genomes (Goll and Bestor, 2005) (**Figure 1.1 A**). In vertebrates genomes, cytosine methylation is mostly restricted to the symmetrical cytosine-guanine (CpG) dinucleotide context (**Figure 1.1 B**). DNA methylation is mitotically



#### Figure 1.1: DNA Methylation

(A) Chemical structure of cytosine DNA base and its methylated form, 5-methyl cytosine. (B) In vertebrates genomes, cytosine methylation of DNA is mostly restricted to the symmetrical cytosine-guanine (CpG) dinucleotide context. Images modified from Wikimedia Commons.

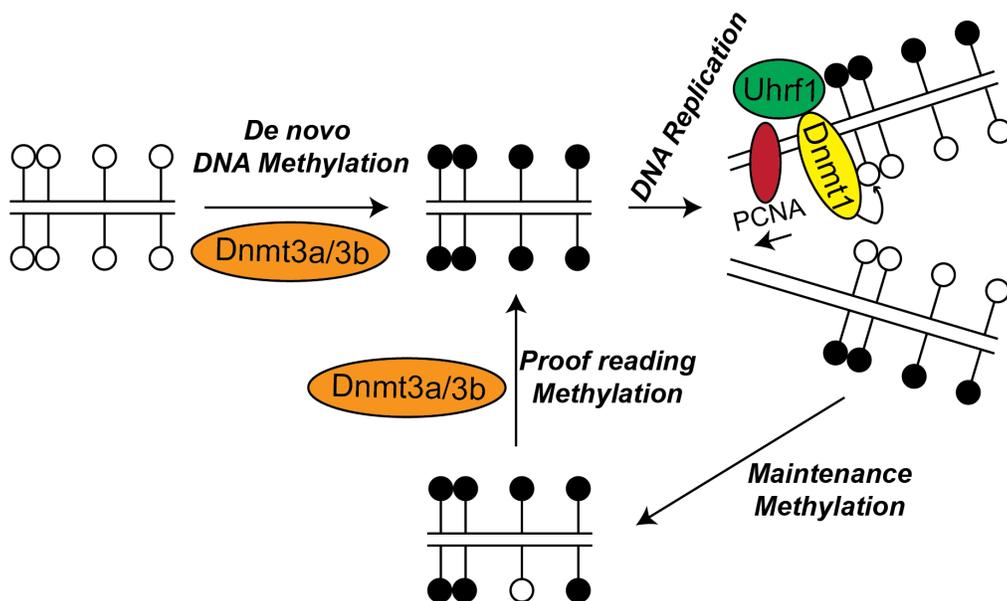
inherited and is largely associated with transcriptional repression (Suzuki and Bird, 2008).

DNA methylation plays an important role in normal development through its function in regulating gene expression, X-chromosome inactivation and genomic imprinting (Jones and Takai, 2001). Global methylation deficiencies are linked to a variety of adverse outcomes including deregulation of gene expression, developmental defects, elevated levels of DNA damage and increased genome instability during mitosis (Smith and Meissner, 2013).

### **Enzymes involved in establishment and maintenance of DNA methylation**

DNA methylation is established by the transfer of a methyl group from S-adenosyl-L-methionine to the fifth position of the cytosine ring by DNA methyltransferases (DNMT). Based on their structure and function, DNMTs largely fall into two conserved families: DNMT1, and DNMT3 (Goll and Bestor, 2005) (**Figure 1.2**). DNMT1 is a maintenance methyltransferase that propagates 5mC at CpG dinucleotides during mitosis by recognizing hemimethylated sites (Li et al., 1992). DNMT1 localizes to the replication fork in the S phase of the cell cycle and copies methylation states onto the newly synthesized DNA strand in a semi-conservative manner (Leonhardt et al., 1992). Ubiquitin-like with PHD and ring finger domains 1 (UHRF1) is a cofactor to DNMT1 that is essential for targeting DNMT1 to replication forks to maintain DNA methylation (Bostick et al., 2007). Together, DNMT1 and UHRF1 ensure the faithful maintenance of DNA methylation patterns during DNA replication. Knockout mutations of *Dnmt1* in mice are embryonic lethal and cause up to 90

% reductions in genome-wide DNA methylation (Li et al., 1992). Loss of function of MET1, the *Arabidopsis thaliana* ortholog of *Dnmt1*, leads to developmental abnormalities such as delayed flowering and reduced fertility (Xiao et al., 2006). Mutations in *dnmt1* in zebrafish also causes embryonic lethality and affects liver and pancreatic development (Anderson et al., 2009). Similar to *Dnmt1* mutations, homozygous mutations in *Uhrf1* results in genome-wide demethylation and embryonic lethality in mice and zebrafish (Feng et al., 2010; Sharif et al., 2007).



**Figure 1.2: Schematic of the functions of DNA methyltransferase (Dnmt) enzymes**

De novo establishment of 5mC (depicted by solid lollipops) is performed by Dnmt3a and Dnmt3b. After DNA replication, the newly replicated products are hemi-methylated. Uhrf1 binds to hemimethylated DNA and directs Dnmt1 to the replication fork for maintaining methylation. Dnmt3a and Dnmt3b can function as proofreaders and fill the gaps of the hemimethylated CpG sites missed by Dnmt1 especially at heavily methylated genomic sites. Adapted from (Chen et al., 2003).

The establishment of DNA methylation marks are orchestrated largely by the *de novo* DNMTs of the DNMT3 family (DNMT3A and DNMT3B) (Okano et al., 1999). During early mammalian embryogenesis, DNA methylation marks are established by Dnmt3a and Dnmt3b following a wave of global erasure of methylation. While global methylation patterns appear intact in *Dnmt3a* deficient mice and they survive to term, they are runted and die early in adulthood with a loss of germ cells in males (Okano et al., 1999). *Dnmt3b* mutant mice die around 9.5 dpc with demethylation of minor satellite repeats (Okano et al., 1999). *De novo* DNMTs are also implicated in the maintenance of methylation patterns at germline genes and repetitive elements (Chen et al., 2003; Liang et al., 2002).

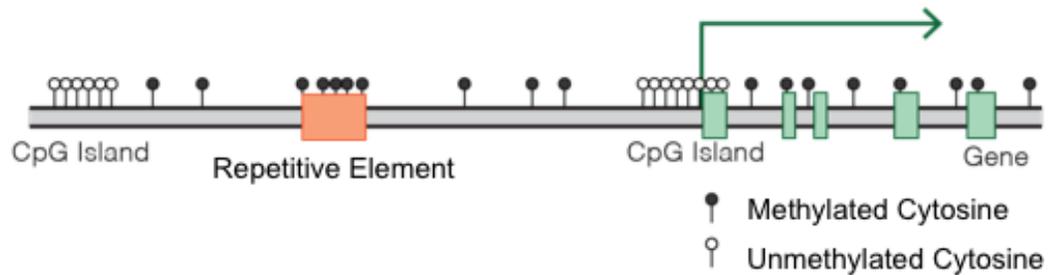
A third homolog, DNMT3L, which lacks methyltransferase activity, cooperates with DNMT3A for establishing genomic imprints in the germline (Bourc'his et al., 2001; Chedin et al., 2002). *Dnmt3l* homozygous null mice are viable however both males and females are sterile. A recently discovered enzyme, Dnmt3c, is required for the *de novo* methylation of the promoters of young transposable elements in the male germ line in mice (Barau et al., 2016).

Together, these various DNMTs play essential roles in establishment and maintenance of stable DNA methylation patterns in many cellular and developmental contexts.

### **Distribution of DNA methylation in mammalian genomes**

DNA methylation is detected using bisulfite conversion, methylation-sensitive restriction enzymes, methyl-binding proteins and anti-5mC antibodies

(Zilberman and Henikoff, 2007). High-throughput sequencing combined with 5mC detection techniques has revealed genome-wide distribution patterns of DNA methylation that have contributed to our understanding of its function (Laird, 2010).



**Figure 1.3: Distribution of DNA methylation in mammalian genomes.**

The mammalian genome is depleted of CpGs and majority of these CpGs are methylated (solid lollipops). The bulk of the methylated CpGs are found at repetitive elements. CpG islands are enriched for CpG dinucleotides, coincide with gene promoters, but are generally unmethylated (empty lollipops). Gene bodies are often CpG poor but are extensively methylated. Image from Wikimedia Commons.

The frequency of CpG dinucleotides is statistically underrepresented in organisms with 5mC, for instance, mammals contain only 20-25% of the expected CpGs. In healthy mammalian genomes, 70–80% of these CpGs, however, are methylated (Ehrlich et al., 1982). The bulk of these CpGs are found within the intergenic and intronic regions of DNA particularly within repeat sequences and transposable elements (**Figure 1.3**).

### Repetitive Elements

One of the most conserved features of DNA methylation function is the repression of repetitive elements such as pericentromeric satellite repeats and transposable elements to ensure long-term silencing. Repetitive DNA elements

constitute up to 50% of the human genome and contain approximately 52% of all CpG dinucleotides in the human genome (Lander et al., 2001). The bulk of 5mC (~90%) in vertebrate genomes is found at these repeats (Beisel and Paro, 2011). In humans, repetitive elements are arranged as interspersed repeats or tandem repeats. Interspersed repeats include transposable elements such as SINE (Short Interspersed Elements), LINE (Long Interspersed Elements), LTR (Long Terminal Repeats) and DNA transposons. Tandem repeats include Satellite repeats and Simple repeats that are typically found near the centromere or telomere (Lopez-Flores and Garrido-Ramos, 2012).

Transposable elements are heavily methylated to prevent these parasitic elements from invading and integrating into different parts of the genome. If such parasitic repeat elements are not silenced, they can pose a serious threat to the structural integrity of the genome by causing chromosomal rearrangements and disrupting genes. While the proposed role of 5mC enrichment at transposable elements is to prevent their translocation, the significance of DNA methylation abundance at pericentromeric repeats is unclear. Loss of DNA methylation from these repeat sequences is often reported in human disease (Robertson, 2005).

### Gene Body Methylation

Gene bodies are often CpG poor but are extensively methylated. Contrary to the canonical role of DNA methylation in repression, gene body methylation is often associated with transcribed genes (Jones, 2012). While the detailed role of gene body methylation is unclear, some studies suggest its role in regulating splicing (Shukla et al., 2011).

### CpG islands

Less than 10% of CpGs occur in CG-dense regions in small genomic regions of about one kilobase, known as CpG islands. Although CpG islands account for only about 1% of the genome and less than 10% of the total genome-wide CpGs, these regions contain over 50% of the unmethylated CpGs. These CpG islands are most often found in the promoters or first exons of genes and are unmethylated in healthy cells (Deaton and Bird, 2011). An open chromatin structure, enriched in acetylated forms of histones H3 and H4, further marks CpG islands for transcriptional activation (Tazi and Bird, 1990). CpG islands are hypermethylated in tumors resulting in the transcriptional silencing of tumor-suppressors (Esteller, 2002).

### **DNA methylation in Zebrafish**

While, 5mC is an important and extensively characterized epigenetic modification, it is absent or present in low-levels in many popular invertebrate model organisms such as *S. cerevisiae*, *D. melanogaster* and *C. elegans*. *Danio rerio* (Zebrafish) is an important emerging model system to study epigenetic regulation in vertebrate development and disease (Mudbhary and Sadler, 2011). ~70% of mammalian genes have at least one zebrafish ortholog and proteins involved in chromatin regulation are highly conserved (Howe et al., 2013).

The DNMT machinery is conserved in zebrafish through the presence of a maintenance methyltransferase (Dnmt1) and 6 orthologs of *de novo* DNA methyltransferases (Dnmt3) (Goll and Halpern, 2011). Mutations and

morpholino-based knock down of Dnmt1 in zebrafish resulted in global DNA hypomethylation and display several terminal differentiation defects (Rai et al., 2006) (Anderson et al., 2009). Similar to mice, *uhrf1* zebrafish mutants display global 5mC reduction (Feng et al., 2010). Systematic mutation of the six zebrafish Dnmt3 orthologs has not been performed, although morphants for Dnmt3bb.2 (an ortholog of DNMT3B) display neuronal defects (Rai et al., 2010).

Consistent with mammalian genomes, zebrafish have high levels of DNA methylation (80%) at CpG dinucleotides at all sequence contexts, and are depleted for 5mC at CpG islands in promoters of genes. DNA methylation is comparably enriched at repetitive elements and gene bodies in zebrafish as well (Feng et al., 2010).

Zebrafish are different from mammals in that they lack imprinting and X-chromosome inactivation (Corley-Smith et al., 1996). The developmental dynamics of DNA methylation during early development is distinct in zebrafish compared to mammals. In the early mammalian embryo, both parental genomes undergo genome-wide erasure and re-establishment of global methylation patterns. (Smith and Meissner, 2013). However, remodeling of DNA methylation in the early zebrafish embryo is more limited. The paternal methylome is stably inherited and the maternal methylome undergoes some remodeling to match up to the paternal genome without any global demethylation (Jiang et al., 2013; Potok et al., 2013).

Overall, the similarities in the distribution and function of DNA methylation between zebrafish and mammals set up zebrafish as a promising model to understand the role of DNA methylation in development and disease.

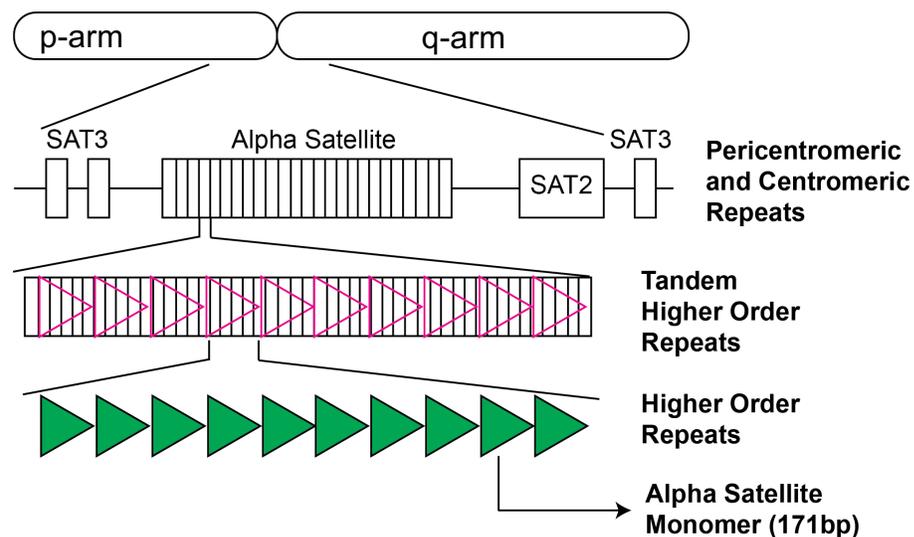
## **PERICENTROMERES**

### **Pericentromeric Satellite repeats**

The region flanking chromosome centers is abundant in several thousand copies of AT-rich, non-coding tandem repeat sequences known as satellite repeats. While satellite repeats are not conserved at a sequence level, their organization at the centromeric and pericentromeric region is a conserved feature of eukaryotic cells (Garrido-Ramos, 2017). These satellite repeats are implicated in important functions such as maintaining heterochromatin architecture (Allshire and Madhani, 2018), facilitating chromosomal segregation during mitosis and meiosis (Plohl et al., 2014) and encapsulating the genome within a single nucleus (Jagannathan et al., 2018).

In humans, the centromeric and pericentromeric region consists of several thousand copies of alpha-satellite DNA with a monomer length of ~170bp (**Figure 1.4**). The alpha satellite monomer is repeated to form a higher repeat which in turn is repeated several times to form a higher order repeat array in the peri/centromeric region. The alpha-satellite repeats are found on all chromosomes and are important for kinetochore assembly that propels cells division (Tyler-Smith and Brown, 1987). In addition to the peri/centromeric alpha-satellites, three types of classical satellite repeats found in the pericentromeric regions of chromosomes and account for ~ 4%–5% of the

human genome. Satellite type I are short AT-rich sequences found at pericentromeric regions of most chromosomes. Type II and type III satellites (SAT2 and SAT3) are made of a 5bp GGAAT repeat unit and form condensed domains in the pericentromeric regions of chromosomes 1, 9, 16 (Vourc'h and Biamonti, 2011).



**Figure 1.4: Pericentromeric satellite repeats.**

Schematic representation of human centromeric and pericentromeric satellite repeats. Adapted from (Cleveland et al., 2003)

In mice, two types of repetitive DNA sequences are associated with centromeres. Major satellite repeats (6 megabases of 234 bp units) are located pericentromerically and minor satellite repeats (~600 kb of 120 bp units) coincide with the centromeric region (Choo, 1997; Joseph et al., 1989). In zebrafish, Type I satellite-like sequence (Sat1) are 186 bp long, A+T-rich (65%), and constitute 8% of the zebrafish genome (Ekker et al., 1992). Fluorescence in situ hybridization experiments revealed that Sat1 sequences are located in the peri/centromeric regions of all chromosomes (Phillips and Reed, 2000).

Based on their location and sequence organization, Sat1 repeats in zebrafish are reminiscent of human alpha-satellite repeats.

### **Pericentromeric Chromatin**

Pericentromeric repeats are enriched with repressive epigenetic modifications. This includes the repressive histone modification Histone H3 lysine 9 trimethyl (H3K9me3) and the DNA modification, 5mC (Dejardin, 2015). The pericentromeric satellite sequences are a significant source of 5mC in vertebrate genomes. Despite their abundance the exact importance of 5mC at these sequences is unknown.

The function of de novo DNA Methyltransferases DNMT3A and DNMT3B is 5mC establishment and DNMT1 on 5mC maintenance is well known (Goll and Bestor, 2005). However, since DNMTs lack intrinsic sequence specificity beyond CpG dinucleotides, mechanisms by which DNMTs target satellite repeats are unclear. DNMT3B is predicted to be involved in establishing DNA methylation at satellite repeats. Minor satellite repeats are substantially demethylated in *Dnmt3b*<sup>-/-</sup>, but not in *Dnmt3a*<sup>-/-</sup> mouse ES cells and similar demethylation of the minor satellite repeats was also detected in E9.5 *Dnmt3b*<sup>-/-</sup> and *Dnmt3a*<sup>-/-</sup>, *Dnmt3b*<sup>-/-</sup> mouse embryos, but not in *Dnmt3a*<sup>-/-</sup> embryos (Okano et al., 1999). These results indicated that minor satellite repeats are specific targets for Dnmt3b.

Suv39h-mediated H3K9me3 is involved in directing DNA methylation to satellite repeats. The H3K9me3 methylation system interacts with Dnmt3b and

influences DNA methylation at satellite repeats (but not other repeats) in mouse ES cells (Lehnertz et al., 2003). These findings indicate a cross-talk between the two most abundant repressive epigenetic modifications found at pericentromeric satellite repeats.

While the classical roles of DNMT1 as a maintenance methyltransferase and DNMT3s as *de novo* enzymes are well established, these roles are not mutually exclusive, especially in 5mC dense repetitive sequences. Mouse ES cells deficient in Dnmt3a and Dnmt3b but containing functional Dnmt1 progressively lost 5mC at repetitive elements after prolonged culture (Chen et al., 2003). It is proposed that due to the risk of errors in maintenance methylation at highly repetitive 5mC rich sequences, Dnmt3a/Dnmt3b ensure methylation fidelity, though the molecular details of this role are unclear (Liang et al., 2002).

Together, these studies indicate an essential role for Dnmt3b in establishment and maintenance of 5mC at peri/centromeric satellite repeats.

### **Diseases associated with pericentromeric DNA hypomethylation**

#### Cancer

Pericentromeric satellite sequences appear particularly susceptible to methylation loss in several cancers. (Enukashvily et al., 2007; Fanelli et al., 2008; Nakagawa et al., 2005; Narayan et al., 1998; Qu et al., 1999b; Suzuki et al., 2002; Tsuda et al., 2002). It is believed that hypomethylation of condensed pericentromeric repeat sequences may predispose cells to cancer-promoting chromosomal rearrangements. However, pericentromeric hypomethylation is

neither necessary nor sufficient to cause chromosomal rearrangements in tumors (Ehrlich et al., 2003; Tsien et al., 2002). Further studies are necessary to understand the consequences of this hypomethylation and its impact on cancer progression. Understanding the consequences of pericentromeric hypomethylation has been challenging since most studies are in the context of global hypomethylation, masking the specific effects of satellite DNA hypomethylation.

### Senescence

Pericentromeric repeats are also found hypomethylated in senescent cells (Erukashvily et al., 2007, Suzuki et al., 2002). Methylome analysis of senescent cells compared to proliferating cells revealed global hypomethylation including hypomethylation of Sat2 sequences in senescent cells, reminiscent of methylation changes reported in cancer (Cruickshanks et al., 2013). This study suggested that such a similarity in methylome between senescent cells and cancer cells might facilitate late-life onset of many cancers.

### ICF Syndrome

Individuals with the rare, autosomal recessive disorder, Immunodeficiency, Centromere and Facial anomalies (ICF) syndrome, also show extensive hypomethylation of pericentromeric Sat2 and Sat 3 and alpha-satellite repeats, while methylation across the rest of the genome is relatively intact (Velasco et al., 2018). Chromosome anomalies including whole-arm deletions and multiradial chromosomes have also been reported in mitogen-stimulated lymphocytes from ICF-patients. However, similar chromosome anomalies are not observed in primary tissues from affected individuals (Ehrlich, 2003). ICF

syndrome presented a unique opportunity to study consequences of pericentromeric hypomethylation in the context of a relatively intact global methylome.

## **ICF SYNDROME, A MODEL FOR PERICENTROMERIC DNA HYPOMETHYLATION**

### **ICF Syndrome Symptoms**

#### Centromeric Instability

Molecular hallmark of this disease is hypomethylation of classical satellite 2 and 3 DNA, the main DNA components of 1qh, 16qh, and 9qh. In addition, a subset of patients displays hypomethylation at centromeric alpha-satellite repeats. This hypomethylation is associated with cytogenetic aberrations in lymphocyte cultures of blood cells from patients. Chromosome anomalies include whole-arm deletions and multiradial chromosomes. Though the aberrations seen in these cultured cells increase with time, they are rarely seen in other tissues. (Ehrlich et al., 2006)

In addition to the molecular hallmark of pericentromeric DNA hypomethylation, ICF patients have other characteristic symptoms.

#### Immunodeficiency

ICF patients are characterized by severe immunodeficiency. Decreased immunoglobulins with often normal level of lymphoid cells is the most common feature of ICF syndrome. Infections are recurrent in patients and often the cause of death at a young age. The cause of immunodeficiency in ICF patients is

unknown. Regular intravenous infusions of immunoglobulin are the most common treatment offered to ICF patients currently.

### Facial Anomalies

ICF Syndrome patients exhibit mild but characteristic facial anomalies which include a flat nasal bridge, hypertelorism (widely spaced eyes), epicanthic folds, macroglossia (enlarged tongue), micrognathia (small jaw), and low set ears.

### Additional symptoms

ICF patients display growth retardation, gastrointestinal defects and an overall failure to thrive. Variable levels of intellectual impairment and neurological defects have also been reported in some patients.

## **Genetic Basis of ICF Syndrome**

Homozygosity mapping and whole-exome sequencing have separately implicated four genes in ICF syndrome: DNA Methyltransferase 3B (*DNMT3B*, ICF type-1), Zinc-finger and BTB domain containing 24 (*ZBTB24*, ICF type-2), Cell division cycle associated 7 (*CDCA7*, ICF type-3) and Helicase, lymphoid-specific (*HELLS*, ICF type-4) (de Greef et al., 2011; Thijssen et al., 2015; Xu et al., 1999).

### DNMT3B (ICF-1)

Approximately half of the reported ICF syndrome patients carry mutations in the de novo DNA methyltransferase *DNMT3B* (OMIM 602900) (Xu et al., 1999). Most of the described mutations in *DNMT3B* cause amino acid substitutions

within the C-terminal catalytic domain, suggesting they may be hypomorphic (**Figure 1.5 A**). Hypomethylation is restricted to pericentromeric classical satellite repeats, Sat2 and Sat3. The alpha-satellite repeats are not hypomethylated in ICF-1. Mutations in *DNMT3B* in ICF Syndrome are consistent with the role of DNMT3B in satellite DNA methylation.

#### ZBTB24 (ICF-2)

Zinc-finger and BTB domain-containing 24 (ZBTB24, OMIM 614064), belongs to a large ZBTB family of transcription factors that form homo- or hetero dimers through their protein interacting BTB domain and bind to target DNA via their zinc finger domains (Siggs and Beutler, 2012). ~30% of ICF patients carry null mutations in *ZBTB24* (**Figure 1.5 B**) (de Greef et al., 2011). In addition to the classical satellite repeats, alpha-satellite repeats are also hypomethylated in ICF Syndrome patients with mutations in *ZBTB24* (ICF-2).

The detailed molecular functions of ZBTB24 are unclear. So far, it has been characterized as a transcription factor that controls the expression of another ICF gene, *CDCA7* (Wu et al., 2016). Three ZBTB family members, ZBTB38, ZBTB33, and ZBTB4 bind both unmethylated and methylated DNA, and function as readers of DNA methylation (Siggs and Beutler, 2012). Immunofluorescence revealed that mouse *Zbtb24* binds to the pericentromere and can do so in the absence of DNA methylation (Nitta et al., 2013). More recently, ZBTB24 was shown to coordinate with DNMT3B to control DNA methylation at intergenic sites (Thompson et al., 2018). Knockdown of ZBTB24 in B-cells and knockout in Human Embryonic Kidney cell-lines significantly reduced proliferation (Liang

et al., 2016; Unoki et al., 2018). The molecular details of the role of ZBTB24 in methylation of peri/centromeric satellite repeats remain to be clarified.

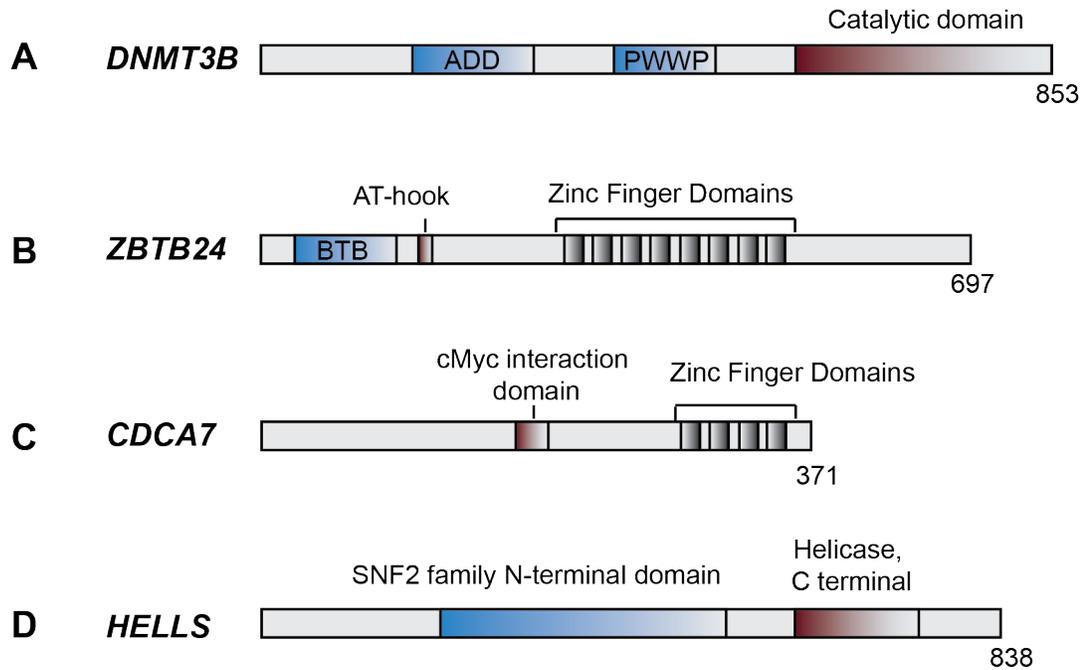
#### CDCA7 (ICF-3)

Cell division cycle associated 7 (CDCA7, OMIM 609937) contains a highly conserved 4CXXC zinc finger domain and has been suspected to be a transcription factor under the control of c-Myc (Gill et al., 2013). Loss of function mutations in *CDCA7* have been identified in 4 ICF patients (**Figure 1.5 C**) (Thijssen et al., 2015). Current understanding of CDCA7 function is that it forms a complex another factor mutated in ICF Syndrome, HELLS (Jenness et al., 2018). CDCA7 also forms a complex with factors involved in non-homologous end joining (NHEJ) and mutations in *CDCA7* in cell culture promote DNA damage (Unoki et al., 2018).

#### HELLS (ICF-4)

Helicase, lymphoid-specific (HELLS, OMIM 603946) is a relatively well characterized chromatin remodeler that is involved in de novo DNA methylation, through its interaction with DNMT3B, dependent on its ATPase domain (**Figure 1.5 D**) (Myant and Stancheva, 2008). Five ICF patients with homozygous null mutations in *HELLS* have been identified (Thijssen et al., 2015). *Hells* mutant mice are perinatal lethal and show signs of genome-wide hypomethylation in addition to methylation loss at repeat sequences including satellite repeats (Tao et al., 2011).

In addition to these 4 ICF genes, more genes involved in ICF Syndrome remain to be identified (ICF-X).



**Figure 1.5: Genes mutated in ICF Syndrome**

Four genes are mutated in ICF Syndrome. **(A)** *De novo* DNA methyltransferase *DNMT3B* **(B)** Zinc-finger and BTB domain-containing 24 (*ZBTB24*), **(C)** Cell division cycle associated 7 (*CDCA7*) and **(D)** Helicase, lymphoid-specific (*HELLS*).

### Epigenomic Assessment of ICF Syndrome

A key molecular hallmark of ICF syndrome is extensive hypomethylation of classical satellite 2 and 3 DNA, the main DNA components of 1qh, 16qh, and 9qh. In addition, ICF-2, ICF-3 and ICF-4 patients display hypomethylation at centromeric alpha-satellite repeats. ICF syndrome is the only human disease that shows Mendelian inheritance of aberrant DNA methylation. Outside of pericentromeres methylation changes in ICF patients appear limited. Comparative methylome analysis of whole blood from ICF syndrome patients of all four genotypes revealed that only 2-3% of the total probes in Illumina Infinium HumanMethylation450 BeadChip were hypomethylated (Velasco et al., 2018).

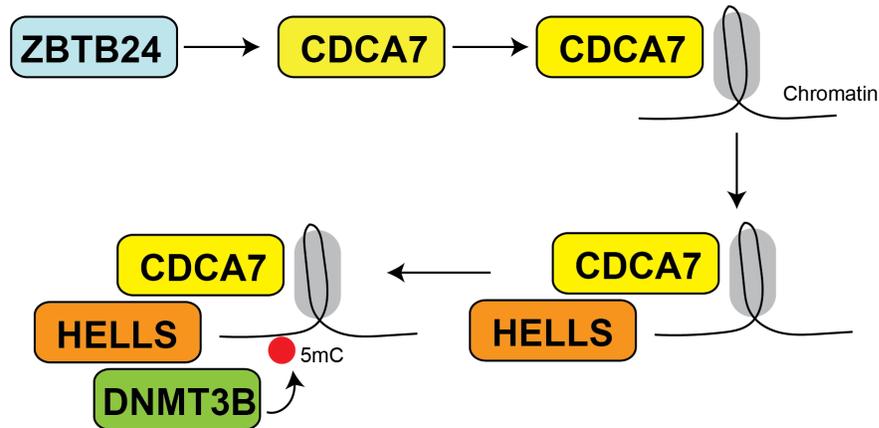
These probes assayed more than 485,000 CpGs throughout the human genome, covering 99% of Refseq genes and approximately 25% of the probes were designed in intergenic regions. This study revealed a strong correlation between methylation landscapes in ICF-2, ICF-3 and ICF-4 patients, distinct from ICF-1. More than 90% of the hypomethylated CpGs in ICF-2, 3 and 4 samples were located in CpG-poor, late-replicating heterochromatic regions compared with less than 70% in ICF-1. In contrast, 25% of hypomethylated CpGs in ICF-1 were located in functional promoters and enhancers but were absent from hypomethylated probes common to ICF-2, 3 and 4 cells (Velasco et al., 2018). The difference in methylome of ICF-1 compared to ICF-2, 3 and 4 indicates that DNMT3B access at several sites outside of the peri/centromeric region is independent of the other ICF factors, ZBTB24, CDCA7 and HELLS.

### **Molecular Basis of ICF Syndrome**

Mechanistically, ZBTB24, CDCA7 and HELLS are thought to converge in a singular pathway that facilitates DNMT3B access to pericentromeric DNA. ZBTB24 promotes the transcription of *CDCA7* (Wu et al., 2016) (**Figure 1.6**). CDCA7 and HELLS form a nucleosomal remodelling complex that is proposed to facilitate DNMT3B access to DNA (Jenness et al., 2018). The similarity in the methylome of ICF-2, 3 and 4 primary cells gives credence to this model suggesting that the three genes are likely acting in the same pathway in the methylation of satellite repeats (Velasco et al., 2018).

However, this model does not clarify how DNMT3B is targeted to satellite repeats. Further, these studies fail to explain why alpha-satellite repeats are

hypomethylated in ICF-2, 3, and 4 but not ICF-1. Understanding the functions of the genetic players associated with ICF Syndrome will provide mechanistic insights regarding the establishment and maintenance of DNA methylation at pericentromeric DNA.



**Figure 1.6: Molecular Basis of ICF Syndrome.**

Predicted model for the role of ICF genes in regulating DNA methylation. ZBTB24 stimulates transcription of CDCA7. CDCA7 associates with chromatin, recruits HELLS, and remodels nucleosomes to allow DNMT3B-mediated DNA methylation. Adapted from (Jenness et al., 2018).

### Animal Models of ICF Syndrome

Previous attempts to generate viable mouse models of pericentromeric hypomethylation through mutation of ICF genes have had limited success. Mice harboring ICF-like mutations in *Dnmt3b* exhibit some characteristics of ICF syndrome including small size and facial anomalies. However, most mice die within 24 hours of birth (Ueda et al., 2006). Deletion of the mouse *Zbtb24* gene was reported to cause embryonic lethality; but methylation changes in these mutants have not been investigated (Wu et al., 2016). Animal models with mutations in *CDCA7* have not yet been reported. Perinatal lethality was observed following deletion of the mouse *HELLS* orthologue. In this case,

mutations were accompanied by roughly 50% reductions in 5mC, and methylation loss was noted at pericentromeres, retroviruses and some single copy sequences (Tao et al., 2011). ICF specific *HELLS* mutations have not been introduced in any organism.

## **DISSERTATION OVERVIEW**

While the general importance of 5mC is well characterized, its role at pericentromeric satellite repeats is less clear. Progress has been hampered by the lack of model systems with large-magnitude hypomethylation limited to pericentromeric repeats. To circumvent this problem, I have developed the first viable animal model of ICF syndrome, a disease characterized by the loss of methylation from pericentromeric repeats. The generation and phenotypic characterization of a zebrafish model of ICF Syndrome are described in **Chapter 2**.

In **Chapter 3**, I present the characterization of the epigenetic features of a zebrafish model of ICF syndrome with mutations in *zbtb24*. This analysis revealed progressive loss of methylation at satellite repeats indicating a role for Zbtb24 in maintaining methylation at these repeat sequences.

The progressive loss of methylation observed in this ICF disease model further allowed for elucidation of primary vs secondary consequences of hypomethylation at these sequences. In **Chapter 4**, I present my findings that identify an interferon response as one of the earliest consequences of pericentromeric hypomethylation. Further, I identify aberrantly overexpressed

satellite RNAs as a trigger for this immune response elicited by pericentromeric hypomethylation.

Hypomethylation of pericentromeric repeats has been reported in several cancers. In **Chapter 5**, I present evidence for a link between pericentromeric hypomethylation and genomic instability in an animal model of ICF Syndrome.

The discussion in **Chapter 6** places these findings in context and identifies potential future directions.

## Chapter 2 : DEVELOPING AN ANIMAL MODEL OF ICF SYNDROME

### PREFACE

The zebrafish genome encodes for 6 orthologs of DNMT3 and one conserved ortholog of ZBTB24. At the onset of this project, CDCA7 and HELLS were not yet identified as genes mutated in ICF syndrome. With the primary aim of creating an animal model for ICF Syndrome, I decided to introduce mutations in the single copy of the zebrafish ortholog of ZBTB24. Additionally, since the function of ZBTB24 was largely unclear, mutational analysis of *zbtb24* held the promise of providing insights into the role of ZBTB24 in establishment and/or maintenance of DNA methylation at pericentromeric DNA.

The zebrafish genome encodes a well-conserved orthologue of ZBTB24 (**Figure 2.1 A**). The BTB and Zinc finger domains are well-conserved (**Figure 2.1 B**). Phylogenetic analysis reveals that *Zbtb24* clusters with its orthologs in other species when compared to its closest members of the *Zbtb* family (*Zbtb14* and *16*). However, *Zbtb24* is also more divergent among its orthologs in various species compared to *zbtb14* and *zbtb16* (**Figure 2.1 C**). This divergence is typical of proteins associated with the centromere and pericentromere, consistent with the rapid divergence of satellite DNAs observed between different species (Henikoff et al., 2001).

**Figure 2.1: Zebrafish Zbtb24 is conserved**

(A) Schematic of human and zebrafish Zbtb24 proteins. (B) Sequence alignment of mouse (Mm), human (Hs) and zebrafish (Dr) Zbtb24 showing conservation of BTB (blue) and Zinc finger (orange) domains. (C) Phylogenetic tree of Zbtb24 with closest members of Zbtb family of proteins, Zbtb14 and Zbtb16 mouse, rat, human, chick, and Tetradon. Alignments were performed on full-length amino acid sequences via ClustalW and DRAWTREE (<http://mobyli.pasteur.fr/cgi-bin/portal.py>).



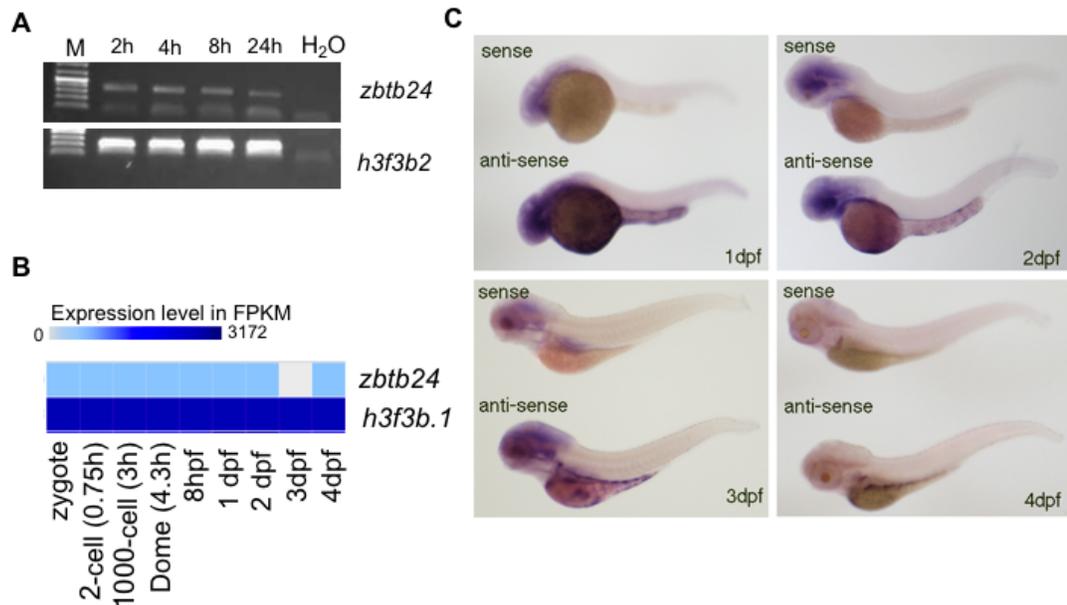
## **EXPRESSION ANALYSIS OF ZEBRAFISH *zbtb24***

Expression of *zbtb24* was first investigated using Reverse Transcriptase-PCR (RT-PCR) by extracting total RNA from various stages of early development in zebrafish. Low-level expression of *zbtb24* was detected in the first 24 hours of development (**Figure 2.2 A**). *zbtb24* mRNA transcripts were detected at 2 hours post fertilization indicating that *zbtb24* is maternally deposited albeit at low levels. This observation was consistent with the expression of *zbtb24* reported in a study that assessed mRNAs expressed during zebrafish development across several time points from 1-cell to 5 days post-fertilization (White et al., 2017). This study indicates that *zbtb24* is expressed at comparable levels across early development (**Figure 2.2 B**).

To determine the spatial-temporal expression of *zbtb24*, I performed whole mount *in situ* hybridization (WISH) using probes specific for *zbtb24* on zebrafish embryos from 1-4 dpf. WISH using an antisense RNA probe for *zbtb24* gene showed low level expression across most tissues in the animal (**Figure 2.2 C**). Sense RNA probe for *zbtb24* is used as a negative control. *zbtb24* appears to be somewhat prominently expressed in the Yolk Syncytial Layer compared to other tissues.

## **MORPHOLINO-MEDIATED KNOCKDOWN OF Zbtb24**

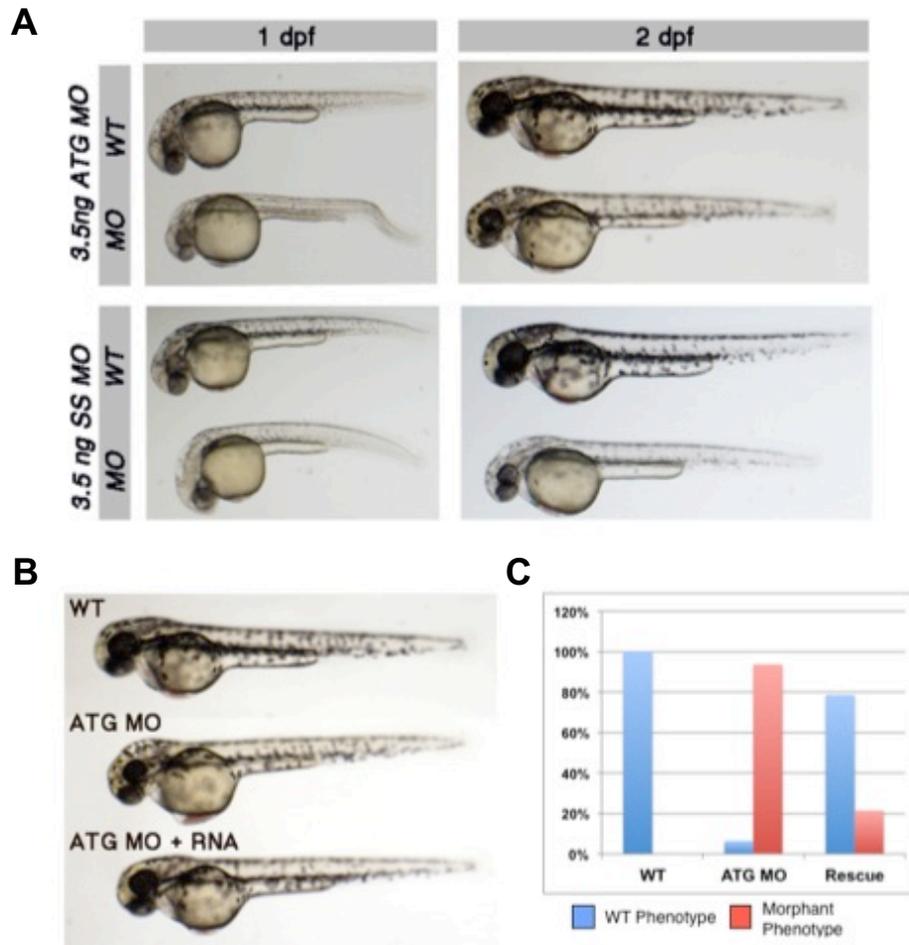
To determine whether the zebrafish Zbtb24 is essential for embryonic development, I injected separately two morpholino-oligonucleotides (MO), an



**Figure 2.2: *zbtb24* is maternally deposited and expressed at low-levels during early embryonic zebrafish development.**

(A) RT-PCR analysis of *zbtb24* in whole zebrafish embryos at indicated stages of development. *zbtb24* expression is compared to Histone H3, family 3B (*h3f3b2*). (h: hours post fertilization) (B) Expression level of *zbtb24* and *h3f3b.1* (Histone H3, family 3B.1) measured in FPKM. Values were obtained in an mRNA expression study of zebrafish development performed by RJ White and Colleagues, 2017. (C) *In situ* hybridization of *zbtb24* from 1 day post fertilization (dpf) to 4 dpf. Sense probe for *zbtb24* is used as a negative control.

ATG blocking (ATG MO) and a splice site blocking (SS MO) morpholino targeting *zbtb24* into 1-cell stage wildtype embryos. Phenotypes were analyzed at both 1 dpf and 2 dpf. Most embryos displayed developmental defects such as smaller eyes, enlarged yolk sac and thinner yolk extension upon injecting 3.5ng of ATG MO (Figure 2.3 A). Similar phenotypes were observed upon injecting 3.5ng of ssMO (Figure 2.3 A). These defects could be rescued by co-injecting

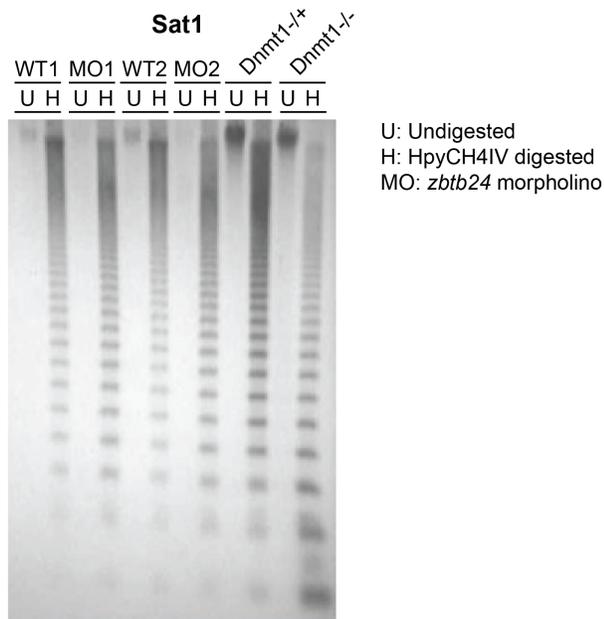


**Figure 2.3: Morpholino-mediated knockdown of Zbtb24 causes developmental abnormalities in zebrafish embryos.**

(A) Morphology of embryos injected with ATG blocking (ATG MO) and splice site blocking (SS MO) morpholinos imaged at indicated time-points alongside uninjected sibling controls. (B) Developmental phenotypes in zebrafish injected with *zbtb24* ATG MO are partially rescued by co-injecting full length *zbtb24* mRNA. Images obtained at 2 dpf. (C) Quantification of embryos demonstrating a rescue of *zbtb24* ATG MO-induced phenotypes upon co-injecting with full length *zbtb24* RNA at 2 dpf.

a full-length *zbtb24* mRNA (**Figure 2.3 B**). The morpholino-mediated knockdown and their rescue was quantified (**Figure 2.3 C**). Most morpholino-injected embryos died by 5dpf. These *zbtb24* knockdown phenotypes are reminiscent of morphant embryos with defects in proliferation.

I was next interested to test whether knockdown of *zbtb24* was sufficient to cause the molecular hallmark of ICF syndrome, hypomethylation of pericentromeric satellite repeats. Pericentromeric satellite type-1 (Sat1) repeats are found on all zebrafish chromosomes and comprise 5-8% of the zebrafish genome (Phillips and Reed, 2000). Based on their location and sequence organization, Sat1 repeats in zebrafish are reminiscent of human alpha-satellite repeats. To test whether Sat1 repeats are hypomethylated in *zbtb24* morpholino injected embryos, I performed a methylation sensitive southern blot. Zebrafish Sat1 has a site that is recognized by the methylation sensitive restriction enzyme HpyCH4IV which is resistant to digestion when methylated. I observed that morpholino-injected embryos showed mild hypomethylation at satellite repeats as evidenced by increased digestion of Sat1 DNA (**Figure 2.4**). *dnmt1* mutant zebrafish embryos at 6dpf were used as a positive control (**Figure 2.4**). While these results were encouraging, morpholino-based loss-of-function approaches are limiting due to the incomplete knockdown of the gene of interest. Moreover, since most embryos died by 5dpf, I was unable to assess ICF syndrome-specific developmental phenotypes in *zbtb24* morphants. At the time of these preliminary analyses, the genome editing revolution was dawning upon the zebrafish model system and so I set out to generate loss-of-function mutations in the zebrafish ortholog of *zbtb24*.



**Figure 2.4: Morpholino-mediated knockdown of *Zbtb24* causes modest hypomethylation of pericentromeric *Sat1* repeats.**

Southern blot of genomic DNA digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish *Sat1* sequence. Genomic DNA was isolated at 2 dpf zebrafish larvae injected with morpholino against *zbtb24*. Uninjected siblings are used as a negative control. DNA from *dnmt1*<sup>-/-</sup> zebrafish larvae at 7 days post fertilization and their phenotypically wild-type siblings (WT) provides a positive control.

## DESCRIPTION OF *zbtb24* MUTANT ALLELES

With the goal to generate an animal model of ICF Syndrome, I introduced null mutations in the zebrafish ortholog of *ZBTB24* using Transcription activator-like effector nucleases (TALENs) for mutagenesis. To introduce mutations in zebrafish *zbtb24*, I designed two (TALENs) targeting the BTB domain and the Zinc Finger domain respectively (**Figure 2.5 A**). I was successful in recovering three different mutant alleles.

### ***zbtb24<sup>mk22</sup>***

Upon co-injection of two TALENs, each targeting the BTB and Zinc Finger Domain respectively, I recovered a 7.9 kb deletion allele (*zbtb24<sup>mk22</sup>*; here after referred to as *zbtb24<sup>Δ</sup>*) that eliminates coding sequence between exons 2 and 5 (**Figure 2.5 A-B**). Generation of this large deletion allele was confirmed through sequencing (**Figure 2.5 C**). Genotyping for the mutant allele is described in **Figure 2.5 D**. Animals that were homozygous for this deletion lacked detectable *zbtb24* transcripts, suggesting *zbtb24<sup>Δ</sup>* is a null allele (**Figure 2.5 E**). Primary analysis in this thesis uses the *zbtb24<sup>Δ</sup>* allele.

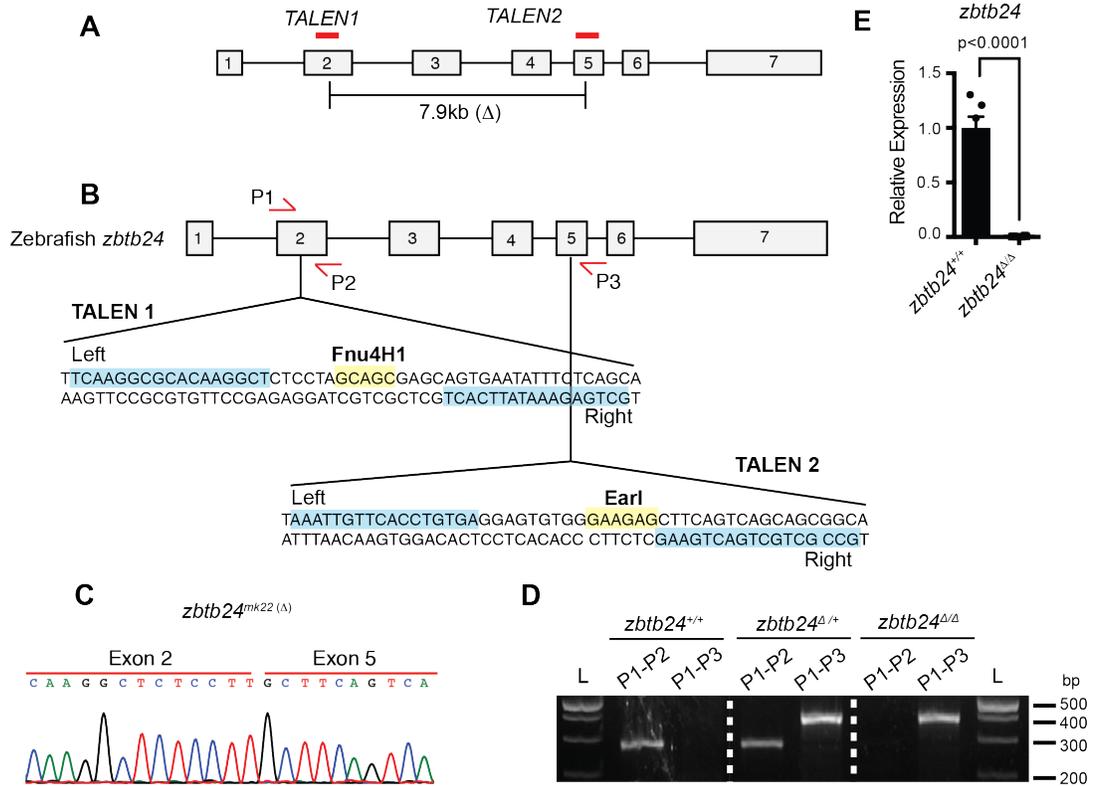
### ***zbtb24<sup>mk19</sup>***

When I injected the TALEN targeting the BTB domain, I recovered a mutant allele that resulted in an 8bp deletion in the second exon. This deletion, causes a frame shift resulting in introduction of a STOP codon at amino acid 82 (*zbtb24<sup>mk19</sup>*) (**Figure 2.6 A**). The mutant allele is resistant to digestion by the restriction enzyme Fnu4H1 which was used to genotyping the allele (**Figure 2.6 B**). Generation of the *zbtb24<sup>mk19</sup>* allele was confirmed through sequencing (**Figure 2.6 C**).

### ***zbtb24<sup>mk21</sup>***

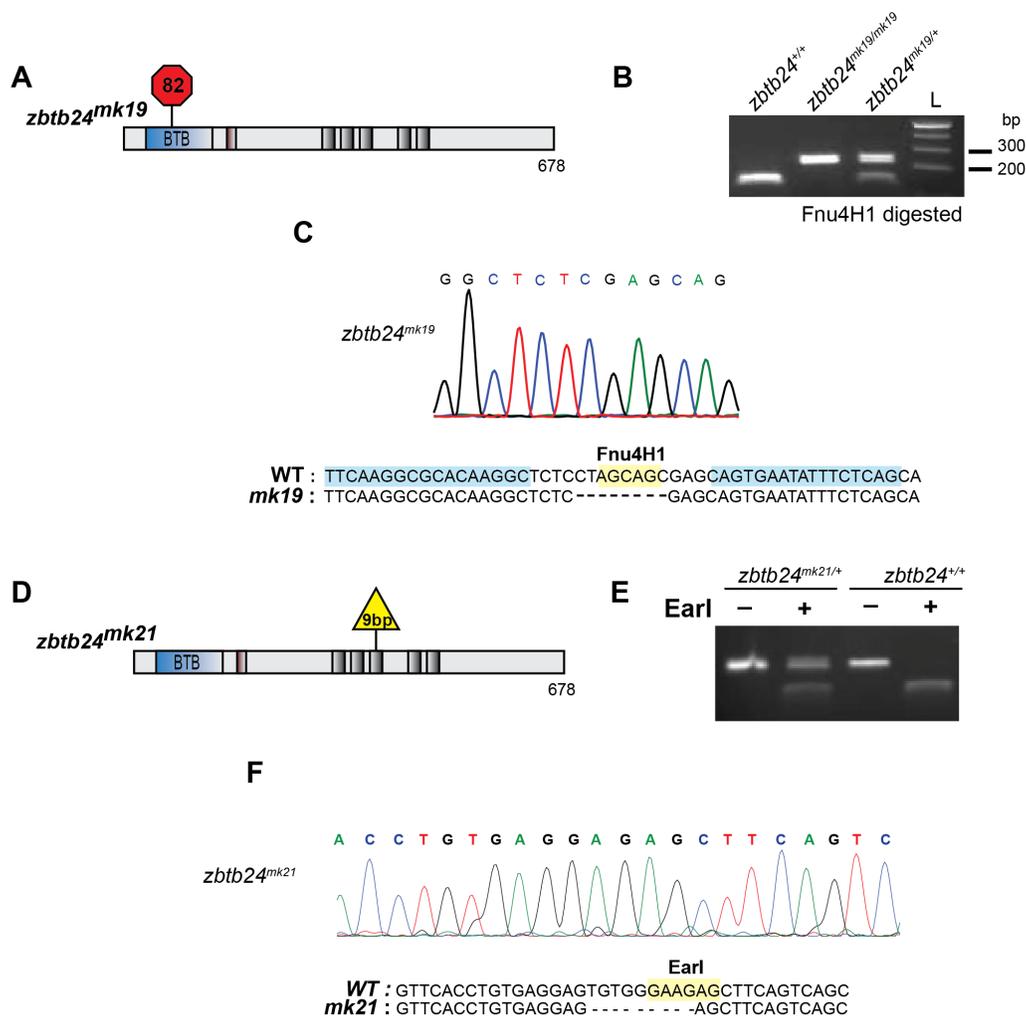
When I injected the TALEN targeting the Zinc Finger domain, I recovered a mutant allele with a 9bp (or 3 amino acid) in-frame deletion in region of the *zbtb24* gene encoding the second zinc finger of Zbtb24 (*zbtb24<sup>mk21</sup>*) (**Figure 2.6 D**). The mutant allele is resistant to digestion by the restriction enzyme Earl which was used to genotyping the allele (**Figure 2.6 E**). Generation of a

*zbtb24*<sup>mk21</sup> allele was confirmed through sequencing (Figure 2.6 F). This allele may be hypomorphic.



**Figure 2.5: Generating a large deletion in the zebrafish ortholog of *zbtb24* (*zbtb24*<sup>mk22(Δ)</sup>)**

(A) Schematic of zebrafish *zbtb24* gene. Location of TALEN target sequences are indicated in red (not to scale). Brackets indicate the region deleted by the *zbtb24*<sup>Δ</sup> allele. (B) Schematic of TALEN sequences showing target sites for introducing mutations in *zbtb24*. P1, P2 and P3 indicate locations for genotyping primers. Sequence in blue indicate target site. Sequence in yellow indicates site of restriction enzyme digestion. (C) Sequence trace confirming generation of large deletion, *zbtb24*<sup>Δ</sup>. (D) Representative genotyping of *zbtb24*<sup>Δ</sup> allele. L: Ladder. P1, P2, P3 represent primers from panel A used for amplifying product in specified lane. (E) qRT-PCR analysis of *zbtb24* mRNA in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 2 wpf (n=6 for each group).



**Figure 2.6: Additional *zbtb24* mutant alleles (*zbtb24<sup>mk19</sup>* and *zbtb24<sup>mk21</sup>*)**  
**(A)** Schematic of *zbtb24<sup>mk19</sup>* with premature stop codon at aa 82. **(B)** Representative genotyping of *zbtb24<sup>mk19</sup>* allele. **(C)** Sequence trace confirming generation of small deletion allele, *zbtb24<sup>mk19</sup>*. **(D)** Schematic of *zbtb24<sup>mk21</sup>* with a 9bp in-frame deletion. **(E)** Representative genotyping of *zbtb24<sup>mk21</sup>* allele. **(F)** Sequence trace confirming generation of small in-frame deletion allele, *zbtb24<sup>mk21</sup>*.

## ***zbtb24* MUTANTS RECAPITULATE KEY PHENOTYPES OF ICF SYNDROME**

### **Growth defects**

*Zbtb24* homozygous mutant embryos were born to heterozygous parents at the expected Mendelian ratios and had no obvious morphological abnormalities during the first two weeks of development (**Figure 2.7 A-B**). Phenotypes that were reminiscent of ICF syndrome emerged as animals matured. Consistent with the small stature observed in ICF syndrome, by 3-4 weeks post fertilization (wpf), *zbtb24*<sup>Δ/Δ</sup> mutant zebrafish were smaller than wild-type siblings raised under identical conditions (**Figure 2.7 C**), and this size reduction persisted into adulthood (**Figure 2.7 D-E**). Similar growth phenotypes were also observed in *zbtb24*<sup>mk19/mk19</sup> (**Figure 2.7 F-G**) and *zbtb24*<sup>mk21/mk21</sup> mutant fish (**Figure 2.7 I-J**). Significant death was noted among *zbtb24*<sup>Δ/Δ</sup> homozygous mutants at 4 months of age and fewer than 10% of *zbtb24*<sup>Δ/Δ</sup> animals survived beyond 8 months (**Figure 2.7 K**). Although not rigorously quantified, reduced lifespans were also observed in *zbtb24*<sup>mk19/mk19</sup> and *zbtb24*<sup>mk21/mk21</sup> mutants. Of the three alleles, *zbtb24*<sup>Δ/Δ</sup> animals displayed the strongest growth defects followed by *zbtb24*<sup>mk19/mk19</sup>, which were more severe than *zbtb24*<sup>mk21/mk21</sup>.

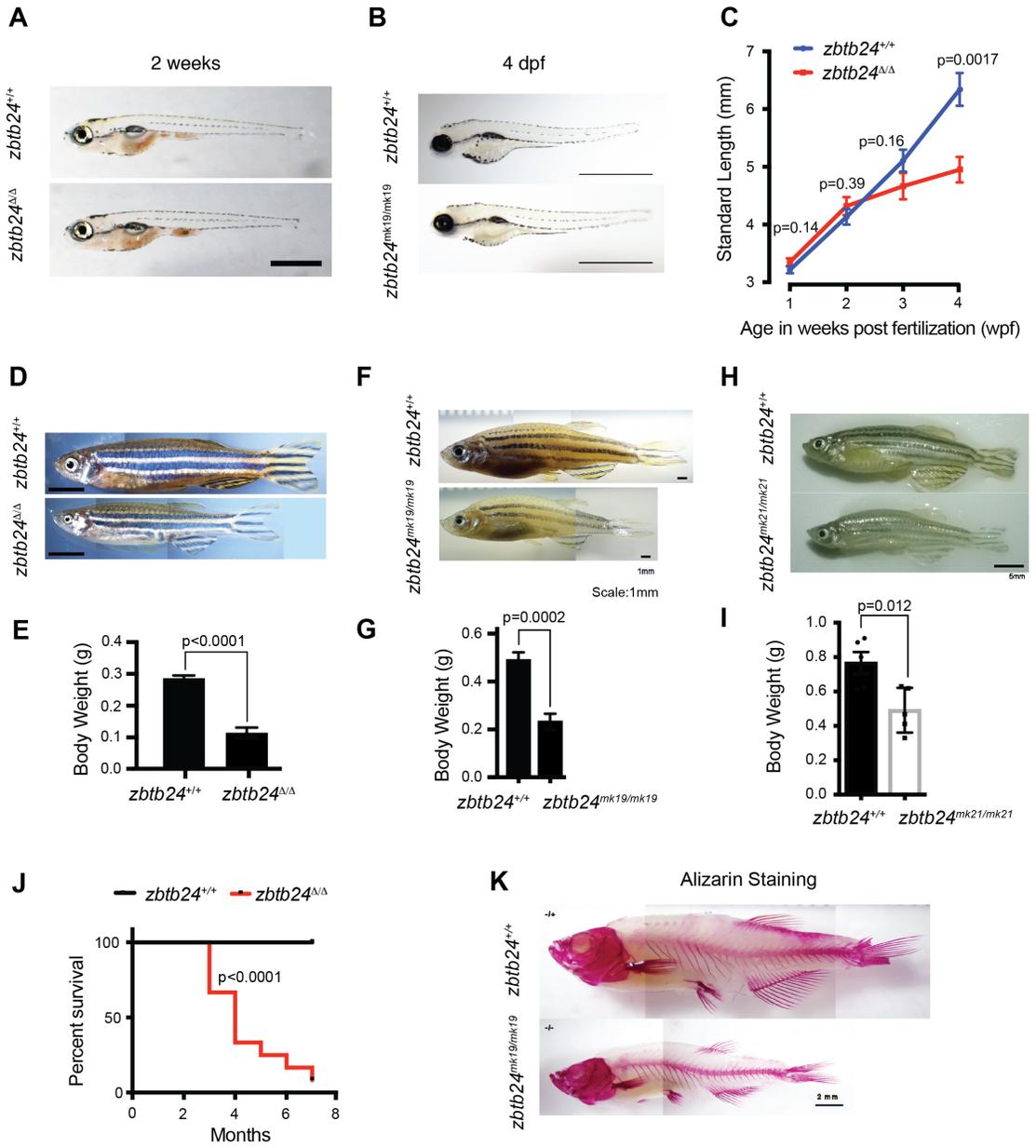
### **Immunodeficiency**

A characteristic feature of ICF syndrome is immunodeficiency characterized by hypogammaglobulinemia or the decrease in immunoglobulins, in the presence of normal lymphoid cell numbers. To test for immunoglobulin deficiency in *zbtb24* mutant fish, I performed expression analysis for zebrafish immunoglobulins *IgM*, *IgD* and *IgZ*. B cell development in zebrafish initiates at

### Figure 2.7: Growth defects in *zbtb24* mutants

(A) Representative images of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 2 wpf. Scale bar: 1 mm. (B) Representative images of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> zebrafish at 4 dpf. Scale bar: 1 mm. (C) Standard length measurements for *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 1, 2, 3 and 4 wpf (n≥6 for each group). (D) Representative images of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 5 months. Scale bar: 5 mm. (E) Average weight of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 5 months (n=5 for each group). (F) Representative images of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> zebrafish at 5 months. Scale bar: 1 mm. (G) Average weight of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> zebrafish at 5 months (n=5 for each group). (H) Representative images of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk21/mk21</sup> zebrafish at 5 months. Scale bar: 1 mm. (I) Average weight of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk21/mk21</sup> zebrafish at 5 months (n=5 for each group). (J) Kaplan-Meier curve indicating survival among groups of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish (n=12 for each group). (K) Alizarin Red bone staining of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> zebrafish adults at 5 months. Scale bar: 2 mm.

**Figure 2.7 (Continued)**



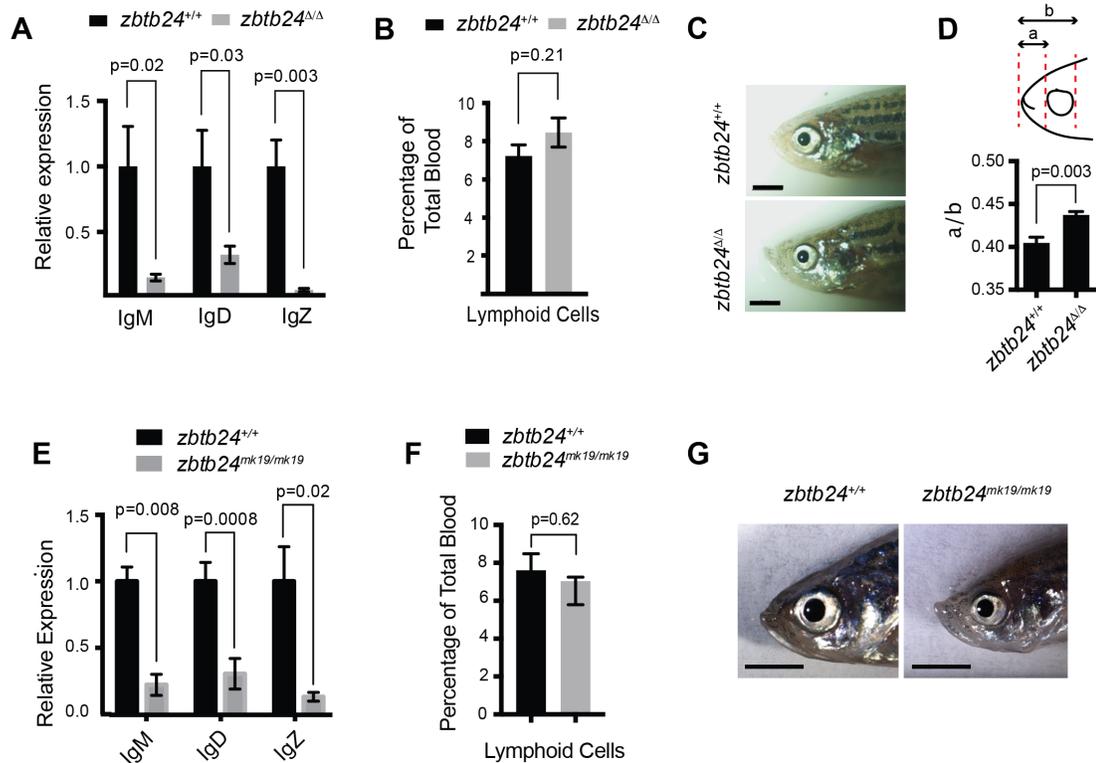
6 wpf (Trede et al., 2004). I observed significantly decreased expression of these immunoglobulins in *zbtb24* mutants (**Figure 2.8 A and E**) at 6 weeks post fertilization. I tested for the number of lymphoid cells using FACS sorting of whole kidney marrow in adult zebrafish. I observed normal lymphoid cell numbers (**Figure 2.8 B and F**). Collectively these data suggest the presence of ICF-like immunological defects in *zbtb24* mutant zebrafish.

### **Facial Abnormalities**

As adults, *zbtb24*<sup>Δ/Δ</sup> mutants exhibited unique facial anomalies that were characterized by a quantifiable elongation of the snout (**Figure 2.8 C and G**). To quantify these abnormalities, I took the ratio of the distance between the tip of the fish's snout and the anterior end of the eye to the posterior end of the eye (**Figure 2.8 D**). This defect is reminiscent of a flattened nasal bridge often reported in ICF Syndrome patients.

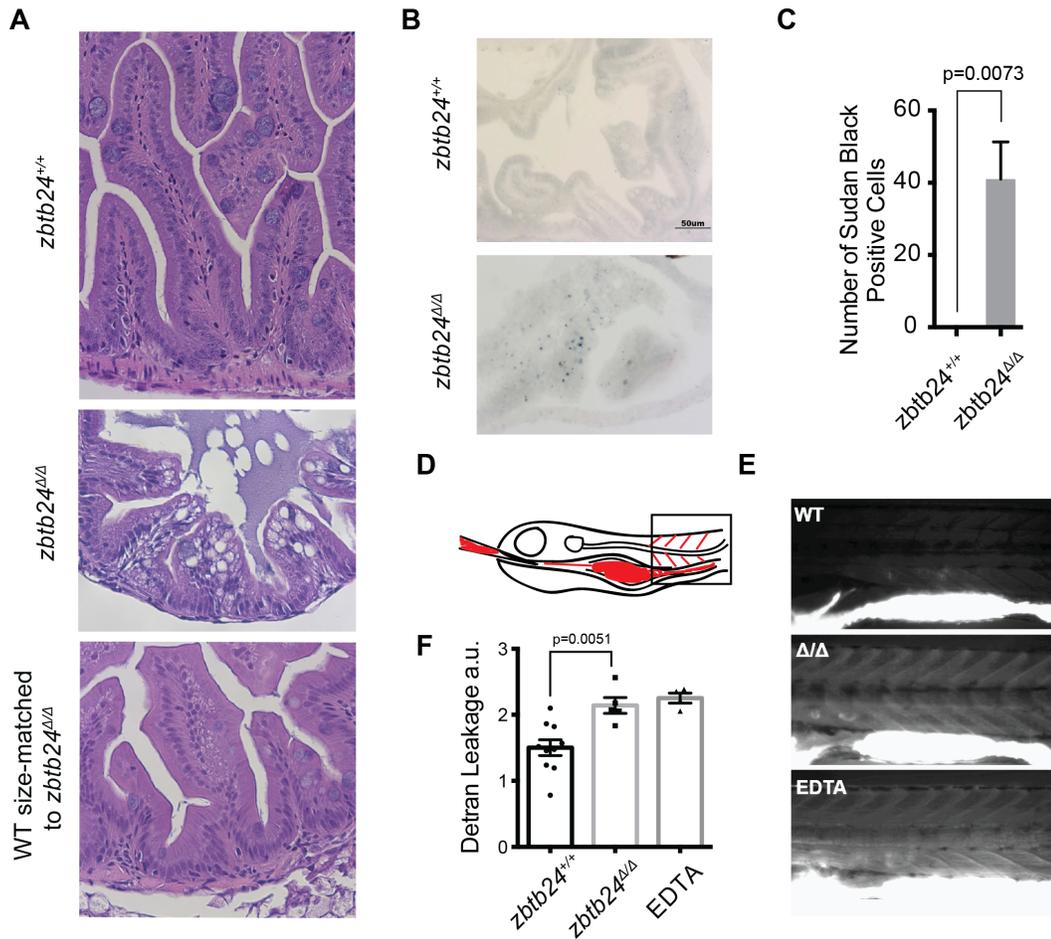
### **Gastrointestinal defects**

ICF patients have often reported gastrointestinal defects such as diarrhea and recurrent gastrointestinal tract infections (Ehrlich, 2003). To test for any pathological defects in the intestinal tract of *zbtb24* mutant fish, I performed H&E staining on sagittal sections of one-month old *zbtb24* mutant fish and their wild-type siblings as a control. Given the differences in size between *zbtb24* mutant fish and their wild-type siblings at this age, I also included sections of wild-type fish size-matched with *zbtb24* mutant fish as an additional control. I observed signs of intestinal inflammation as observed by villous blunting and enlarged goblet cells in the intestinal villi (**Figure 2.9 A**). I further confirmed increased



**Figure 2.8: Immunodeficiency and Facial anomalies in *zbtb24* mutants**

(A) Abundance of IgM, IgD and IgZ transcripts in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 6 weeks post fertilization (n=5 for each group). (B) Quantification of lymphoid cell populations in total blood isolated from *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> kidney marrow from 8-month-old adults, measured by Forward/Side scatter flow cytometry (n=11 for each group). (C) Representative images of facial abnormalities in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> adults at 6 months. Scale bar: 2 mm. (D) Schematic and quantification of facial abnormalities in *zbtb24*<sup>Δ/Δ</sup> zebrafish (n=5 for each group). (E) qRT-PCR analysis of IgM, IgD and IgZ zebrafish immunoglobulins in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> zebrafish measured at 6 weeks post fertilization (n=5 biological replicates). (F) Quantification of lymphoid cell populations in total blood isolated from *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> 11-month-old adult kidney marrow, measured by Forward/Side scatter flow cytometry (n=16 biological replicates). All error bars indicate standard error of the mean (SEM). (G) Facial abnormalities in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> zebrafish. Scale bar: 3 mm.



**Figure 2.9: Gastrointestinal defects in *zbtb24* mutants.**

(A) Representative image of H&E staining of the gut of *zbtb24<sup>+/+</sup>* and *zbtb24<sup>Δ/Δ</sup>* at 8 wpf. (B) Sudan black staining in the gut of *zbtb24<sup>+/+</sup>* and *zbtb24<sup>Δ/Δ</sup>* at 3 wpf. (C) Quantification of Sudan black positive cells in the gut of *zbtb24<sup>+/+</sup>* and *zbtb24<sup>Δ/Δ</sup>* at 3 weeks. Results are representative of 4 independent experiments. Error bars indicate SEM. (D) Schematic of dextran micro-gavage in zebrafish larvae at 3 wpf. (E) Dextran gavage in *zbtb24<sup>+/+</sup>* and *zbtb24<sup>Δ/Δ</sup>* at 3 wpf. EDTA is co-gavaged as a positive control. (F) Quantification of dextran leakage into the vasculature as measured by fluorescence intensity (measured in arbitrary units, a.u.) in *zbtb24<sup>+/+</sup>* and *zbtb24<sup>Δ/Δ</sup>* at 3 wpf.

incidence of inflammation in the gut as assayed by staining for sudan black, a stain for neutrophils, which are first responders to an inflammation site. Sudan Black-positive cells are detected in the villi of *zbtb24*<sup>Δ/Δ</sup> (**Figure 2.9 B-C**) and *zbtb24*<sup>mk19/mk19</sup> larvae (**Figure 2.9 D**). This intestinal inflammation also led to epithelial barrier disruption in the guts of *zbtb24*<sup>Δ/Δ</sup> larvae as assayed by the leakage of microgavaged fluorescent dextran (**Figure 2.9 E-F**). Together these data demonstrate inflammatory bowel-disease like symptoms in *zbtb24* mutants. While gastrointestinal tract infections and diarrhea have been reported in ICF patients, inflammatory phenotypes of the gut have not been categorically described in ICF patients before. These results raise the possibility that intestinal inflammation could contribute to the gastrointestinal defects observed in ICF patients and is an underappreciated aspect of ICF pathology.

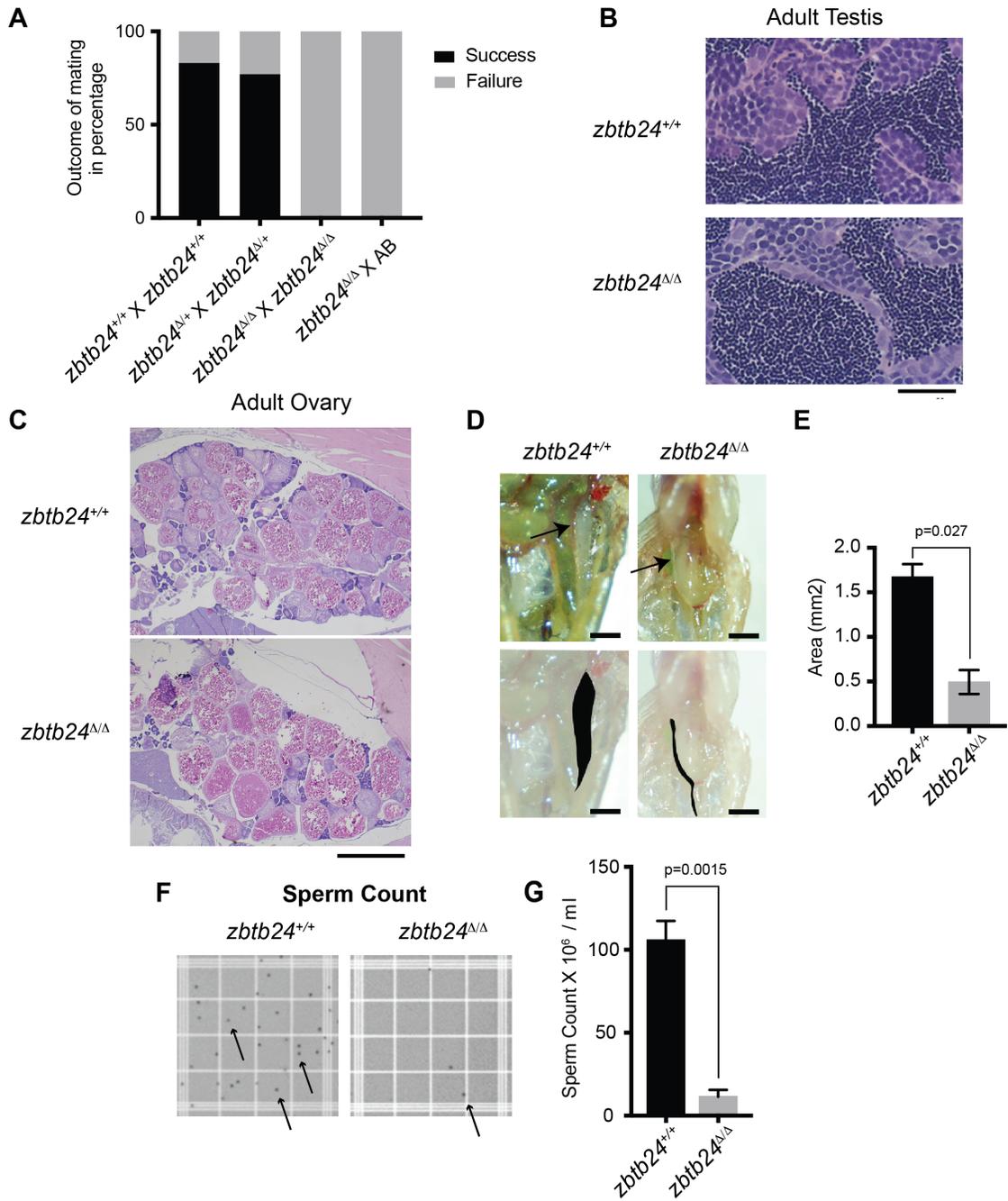
### **Infertility**

All attempts to set up mutant fish either as intercross or outcrossed to ABs have been unsuccessful (**Figure 2.10 A**). To better understand the sterility phenotype, we performed H&E staining on testes and ovary of *zbtb24* mutant adults. Histological sections did not reveal any obvious differences in morphology of male and female gonads (**Figure 2.10 B-C**). We also observe formation of mature sperm. However, the testes in *zbtb24*<sup>Δ/Δ</sup> mutants are thinner compared to wildtype siblings (**Figure 2.10 D**) and sperm count in *zbtb24*<sup>Δ/Δ</sup> mutants is also significantly reduced (**Figure 2.10 E-F**). This defect likely contributes to the infertility observed in the mutants.

**Figure 2.10: Fertility and gonad analysis in *zbtb24* mutants.**

(A) Outcome of mating for the indicated crosses. Success and failure are defined respectively as generation and lack of fertilized embryos from a male x female cross. (B) H&E staining of cross-section of testes dissected from 8-month-old adult *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish. Scale: 20um. (C) H&E staining of cross-section of 8-month-old adult *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish ovaries. Scale: 1mm. (D) Brightfield images of adult testes in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish. The lower panel highlights the trace of the testis shaded in black. Scale: 1mm. (E) Area of the black shaded region in panel D was used to measure the size of the testis. Error bars indicate SD from 2 biological replicates. (F) Brightfield images of counting chamber grid used to count sperm. Black specks (arrows) indicate individual sperm. (G) Sperm count from crushed testes dissected from adult *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish. Error bars indicate SEM from 3 biological replicates.

**Figure 2.10 (Continued)**



## DISCUSSION

Taken together, these findings identify *zbtb24* homozygous mutant zebrafish as a faithful animal model of ICF syndrome phenotypes. In this study, I describe a viable animal model of ICF syndrome which recapitulates key phenotypic hallmarks of the disease including slow growth, facial anomalies, immunoglobulin deficiencies and reduced lifespan. Given that previous attempts to model ICF syndrome in mice have resulted in perinatal or embryonic lethality (Geiman et al., 2001; Ueda et al., 2006; Wu et al., 2016), this zebrafish model provides an important new resource for understanding ICF disease etiology during juvenile and adult life stages. Most research on ICF Syndrome has been performed on immortalized lymphoblastoid cell lines from ICF patients which do not provide any information on the tissue specific phenotypes observed in patients. *zbtb24* mutant zebrafish will be useful for understanding phenotypes, such as immunoglobulin deficiency, which have not been reported in mouse models and are difficult to study in cell culture systems.

While the developmental progression of ICF syndrome phenotypes has been anecdotally reported, it has not been well-characterized. The lack of model systems for this disease has been a major impediment for understanding the developmental aspects of ICF syndrome. Analysis of developmental phenotypes in *zbtb24* mutants reveal that the *zbtb24* mutant fish are born normal but demonstrate ICF-like phenotypes with age. Developmental defects emerge in the juvenile stage between 3-4 weeks post fertilization. Combined immunodeficiency has also been reported to develop with age, at least in one patient (von Bernuth et al., 2014). This study indicates that developmental

progression of ICF phenotypes is an essential but underappreciated feature of the disease.

Gastrointestinal infections and diarrhea are commonly reported clinical manifestations of ICF syndrome. More recently, a 17-month-old patient with Very-Early-Onset IBD who was reported to have mutations in *ZBTB24* (Conrad et al., 2017) In this study, I provide evidence for intestinal bowel disease (IBD)-like inflammatory symptoms in *zbtb24* mutant zebrafish. As early as 3 weeks post fertilization, I observe signs of intestinal inflammation which develop into more chronic IBD-like symptoms with age. This is unlikely a byproduct of immunodeficiencies because at this stage the adaptive immune system has not yet developed (Trede et al., 2004). Together, this study identifies an underappreciated link between ICF syndrome and intestinal inflammation. Verification of this pathological feature in ICF patients could pave the way for appropriate therapy to alleviate the gastrointestinal symptoms observed in ICF patients.

## **Chapter 3 : EPIGENETIC ANALYSIS OF AN ANIMAL MODEL OF ICF SYNDROME**

### **PREFACE**

A zebrafish model of ICF syndrome with mutations in *zbtb24* described in Chapter 2, recapitulates the principle clinical and pathological manifestations of the human disease. The molecular hallmark of this disease is hypomethylation of pericentromeric satellite DNA repeats. ICF syndrome is the only human disease that involves Mendelian inheritance of aberrant DNA methylation. A majority of the methylome analysis of ICF syndrome has employed immortalized cell lines from patients. However, DNA methylation pattern changes have been reported in such long-term cultured lymphoblastoid cell lines (Grafodatskaya et al., 2010). Thus, representative methylome changes in ICF syndrome patients has not been well characterized. Recently, the first comparative methylation profiling in primary blood samples from ICF patients with all four genotypes was reported which revealed that, despite unifying hypomethylation of pericentromeric repeats and a few common loci, methylation profiling clearly distinguished ICF1 from ICF2, 3 and 4 patients (Velasco et al., 2018). The study showed that in addition to satellite repeats, ZBTB24, CDCA7 and HELLS mutations affect CpG-poor regions with heterochromatin features. DNMT3B mutations also affected Sat2 DNA methylation, though, preferential hypomethylation of CpG islands was an exclusive feature of ICF-1 methylome. While this study provided insight into methylome differences in the four types of ICF Syndrome, analysis on the developmental dynamics of methylation changes at satellite repeats (the unifying feature of ICF syndrome) has been

completely lacking due to the rarity of patients and lack of model systems of ICF syndrome. The viable zebrafish model of ICF syndrome with mutations in *zbtb24* provided a unique opportunity to assess the methylation status at pericentromeric repeats *in vivo*, over time. Given that ICF 2, 3 and 4 methylome show similarity based on analysis by Velasco and colleagues, findings from such an analysis on *zbtb24* mutants could likely be generalized to ICF 2, 3 and 4 (Velasco et al., 2018).

## **PERICENTROMERIC DNA METHYLATION ANALYSIS**

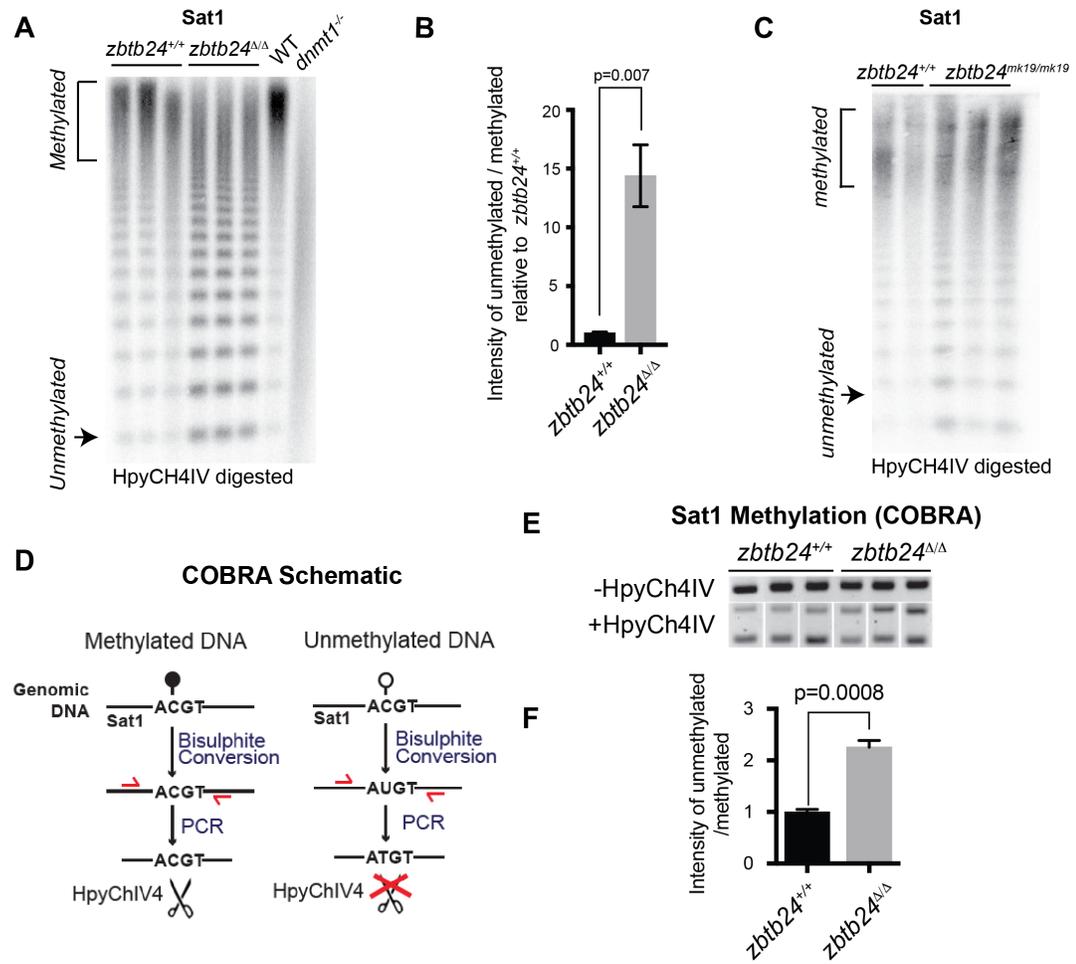
### **Pericentromeric satellite repeats are hypomethylated in *zbtb24* mutants**

To test for satellite DNA hypomethylation, I performed methylation-sensitive Southern blot. Zebrafish Sat1 has a site that is recognized by the methylation sensitive restriction enzyme HpyCH4IV which is resistant to digestion when methylated. As expected, I found that Sat1 sequences from wild-type one-month fish were resistant to digestion by HpyCH4IV, indicating that these pericentromeric repeats were heavily methylated. In contrast to wildtype, Sat1 sequences from *zbtb24*<sup>Δ/Δ</sup> and *zbtb24*<sup>mk19/mk19</sup> mutant siblings were readily digested with HpyCH4IV, indicating extensive loss of methylation at these repeats (**Figure 3.1 A-C**). The hypomethylation of zebrafish Sat1 was independently verified using a secondary approach called Combined Bisulphite Restriction Analysis (COBRA) which employed bisulphite conversion (Xiong and Laird, 1997) (**Figure 3.1 D-E**). In this approach, total genomic DNA is bisulphite converted, Sat1 DNA is PCR amplified and the amplicon is digested with the restriction enzyme, HpyCH4IV. If Sat1 is methylated, the sequence recognized by HpyCH4IV is intact and digested. If Sat1 is unmethylated,

**Figure 3.1: Pericentromeric satellite repeats are hypomethylated in *zbtb24* mutants.**

(A) Southern blot of genomic DNA digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish Sat1 sequence. Genomic DNA was isolated from one-month-old *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> animals. Each lane represents DNA isolated from one adult individual of the indicated genotype. DNA from *dnmt1*<sup>-/-</sup> zebrafish larvae at 7 days post fertilization and their phenotypically wild-type siblings (WT) provides a positive control. (B) Quantification of methylation changes in panel A. Error bars indicate SEM from the 3 biological replicates. (C) Southern blot of genomic DNA digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish Sat1 probe. Genomic DNA was isolated from one-month-old *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> animals. Each lane represents a biological replicate for the indicated genotype. (D) Schematic of Combined Bisulphite Restriction Analysis (COBRA). (E) DNA methylation changes in Sat1 assayed by COBRA in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> adult fins. Bisulfite-treated DNA was amplified with Sat1 specific primers and digested with HpyCH4IV restriction enzyme. The products from COBRA were run on the same agarose gel, then cropped and presented. (F) COBRA quantification of 5mC levels at Sat1 repeats in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish. For each sample, data is presented as the ratio of undigested (unmethylated)/digested (methylated) DNA fragment normalized to *zbtb24*<sup>+/+</sup>. Error bars indicate SEM from 3 biological replicates.

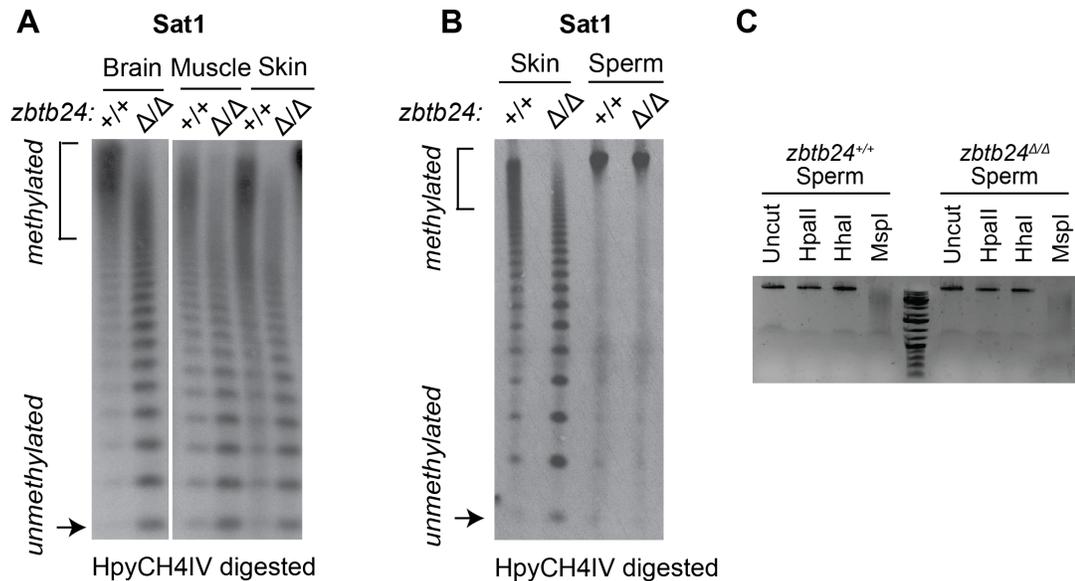
**Figure 3.1 (Continued)**



cytosine at the site of CpG methylation is converted to thymine, is no longer recognized by HpyCH4IV and the amplicon is resistant to digestion. Comparable Sat1 methylation deficiencies were observed when DNA was isolated from dissected adult brain, skin, muscle and fin, suggesting that these sequences were similarly hypomethylated in most adult somatic tissues (**Figure 3.2 A**). However, methylation levels at Sat1 repeats appeared normal in sperm extracted from *zbtb24*<sup>Δ/Δ</sup> mutant adults, suggesting methylation loss may be restricted to somatic tissues (**Figure 3.2 B-C**).

### **Progressive hypomethylation at pericentromeric satellite repeats**

ICF syndrome developmental phenotypes are progressive in *zbtb24* mutants. Methylation analysis in ICF patients has never been performed over time. I thus tested for methylation differences at Sat1 over embryonic and juvenile stages of development. I found that pericentromeric methylation loss in *zbtb24*<sup>Δ/Δ</sup> mutants was progressive. While extensive hypomethylation of Sat1 sequences was detected in adults lacking *zbtb24*, similar hypomethylation was not observed in mutants at 1 wpf. At 2 wpf, *zbtb24* mutants exhibited roughly 3-fold increases in HpyCH4IV digestion, and sensitivity to digestion became increasingly pronounced in older animals. By 32 weeks, Sat1 sequences from *zbtb24* mutants exhibited a 23-fold increase in HpyCH4IV digestion compared to wildtype, suggesting a greater than 95% reduction in methylation of these repetitive sequence blocks (**Figure 3.3 A-B**). Progressive hypomethylation at Sat1 repeats in *zbtb24*<sup>Δ/Δ</sup> was further verified through COBRA (**Figure 3.3 C-D**). Collectively, these findings are suggestive a role for *zbtb24* in the long-term maintenance of 5mC at pericentromeric satellite repeats.



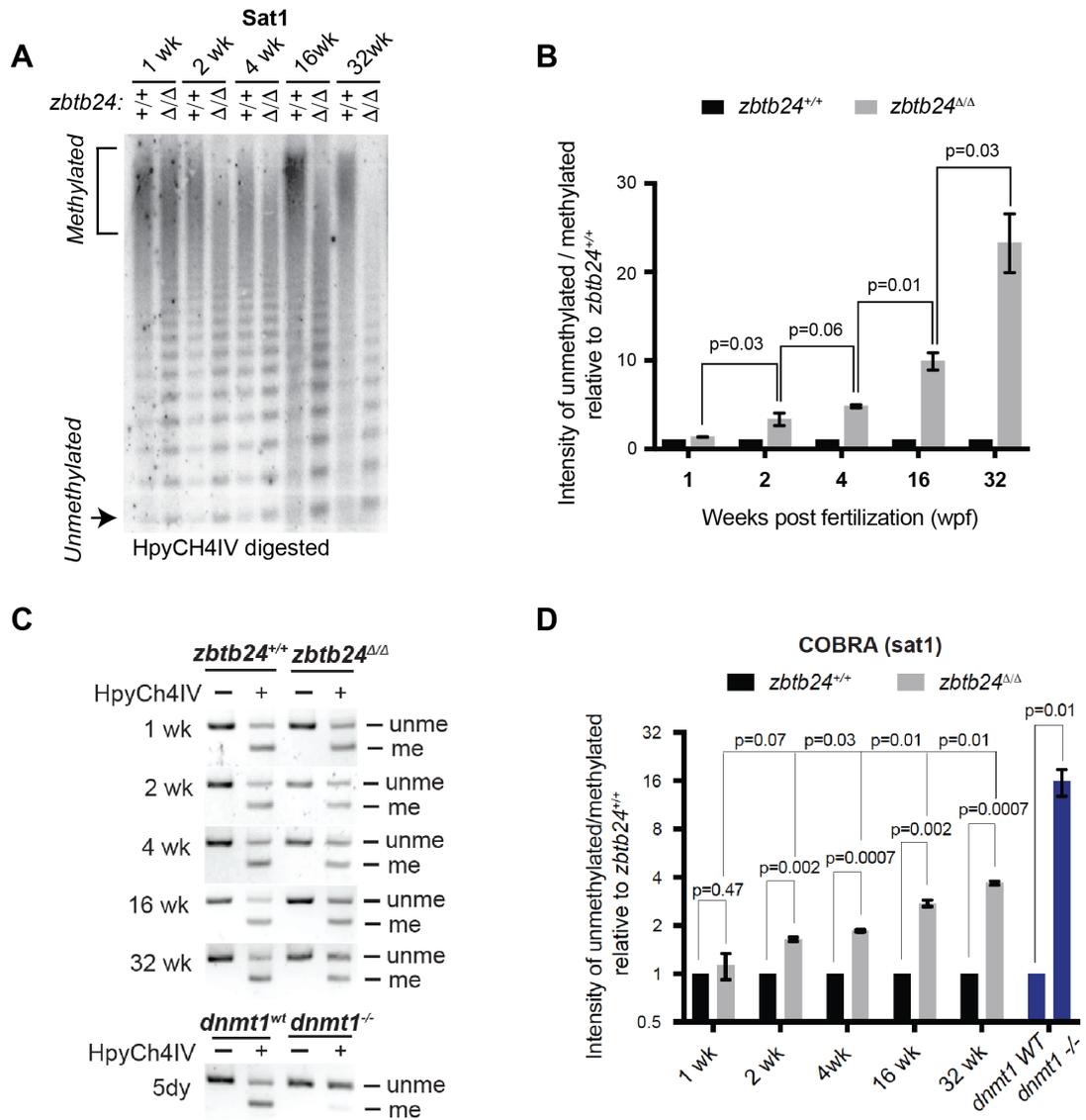
**Figure 3.2: Somatic tissues but not germ cells are similarly hypomethylated at pericentromeric Sat1 repeats in *zbtb24* mutants.**

(A) Southern blot of genomic DNA from different tissues in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish sat1 probe. Each lane represents pooled DNA samples of the indicated tissue from 3 zebrafish 8-month-old adults. (B) Southern blot of genomic DNA from Skin and sperm from 8-month-old adult *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish Sat1 sequence. (C) Ethidium bromide gel with genomic DNA collected from sperm digested with methylation sensitive enzymes, HpaII and HhaI and the methylation-insensitive isoschizomer of HpaII, MspI.

**Figure 3.3: *zbtb24* mutation causes progressive methylation loss at pericentromeric Sat1 repeats in *zbtb24* mutants.**

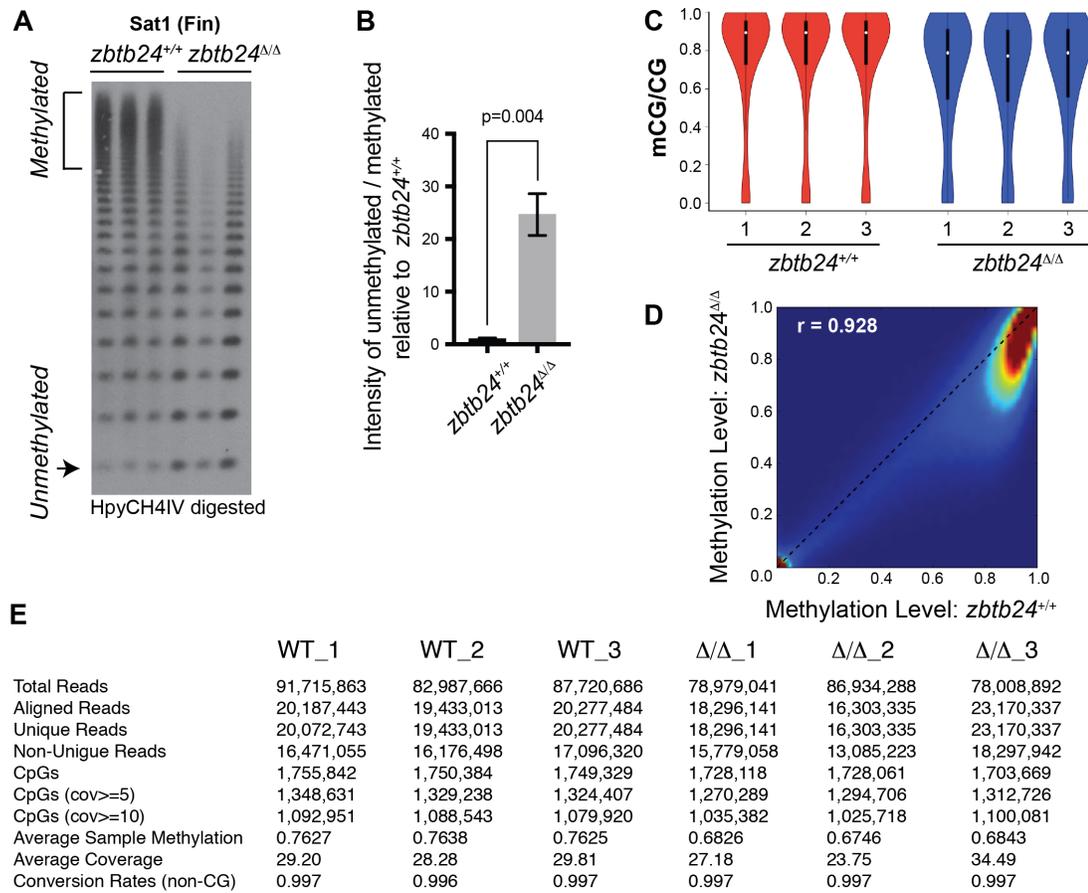
A) Southern blot of genomic DNA digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish Sat1 sequence. Genomic DNA was isolated from *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> animals at 1, 2, 4, 16 and 32 wpf as indicated. (B) Quantification of methylation changes at Sat1 sequences in panel C. Data represent averages from two independent experiments. Error bars represent the standard deviation (SD). (C) DNA methylation changes in sat1 assayed by COBRA in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> animals at 1, 2, 4, 16 and 32 wpf. Bisulfite-treated DNA was amplified with sat1 specific primers and digested with HpyCH4IV restriction enzyme. The products from COBRA were run on the same agarose gel, then cropped and presented. Degree of hypomethylation was expressed as a ratio of the intensity of undigested (unmethylated: unme) fragment divided by the digested (methylated: me) fragment (D) COBRA quantification of 5mC levels at sat1 repeats in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 1, 2, 4, 16 and 32 wpf. For each sample, data is presented as the ratio of undigested (unmethylated) / digested (methylated) DNA fragment normalized to *zbtb24*<sup>+/+</sup> Error bars indicated standard deviation (SD) from 2 independent experiments.

**Figure 3.3 (Continued)**



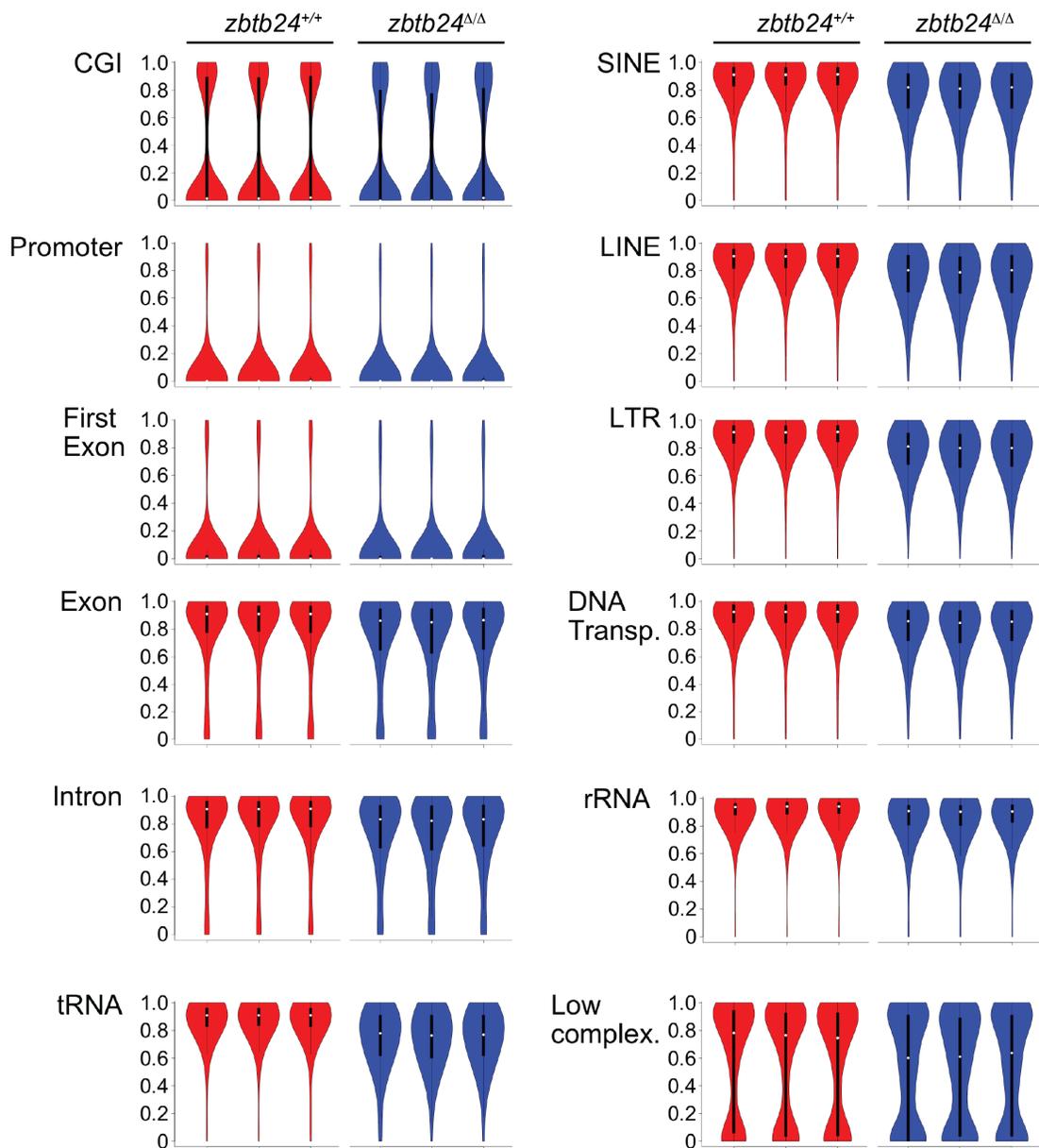
## GENOME-WIDE METHYLOME ANALYSIS

To clarify whether additional sequences were hypomethylated in *zbtb24* mutants, we performed Enhanced Reduced Representation Bisulfite Sequencing (ERRBS) using genomic DNA isolated from the fins of three *zbtb24*<sup>Δ/Δ</sup> mutant 6-month-old adults and three wild-type siblings (**Garrett-Bakelman et al., 2015**). This approach is used to assess the methylation status at CG rich genomic loci. The enrichment for such loci is achieved by digesting with restriction enzyme MspI (recognizing CCGG) followed by bisulphite conversion and next generation sequencing. At 6 months, Sat1 sequences from isolated fins were 20-fold more sensitive to HypCH4IV in *zbtb24* mutants compared to controls, indicating extensive loss of DNA methylation at pericentromeric repeats (**Figure 3.4 A-B**). In collaboration with John Edwards (Washington University at St. Louis, USA), we used ERRBS data to interrogate the methylation status of 979,971 non-pericentromeric CpG sites across the genome in the same tissue samples. Our analysis revealed a strong correlation between genome wide 5mC levels in wild-type and *zbtb24*<sup>Δ/Δ</sup> mutant adults (Pearson's correlation value of 0.928), although overall methylation levels were reduced by ~10% in mutants (**Figure 3.4 C**). This 10% methylation reduction in mutants consisted primarily of small-magnitude changes in 5mC across all methylated sequence features in the zebrafish methylome (**Figure 3.5**). Only 1.3% (13,205) of examined CpG dinucleotides exhibiting methylation differences of greater than 20%. Consistent with this finding, at a threshold of 20% change (p-value<0.01), only 55 differentially methylated regions (DMRs) were identified between wild-type and *zbtb24*<sup>Δ/Δ</sup> adults.



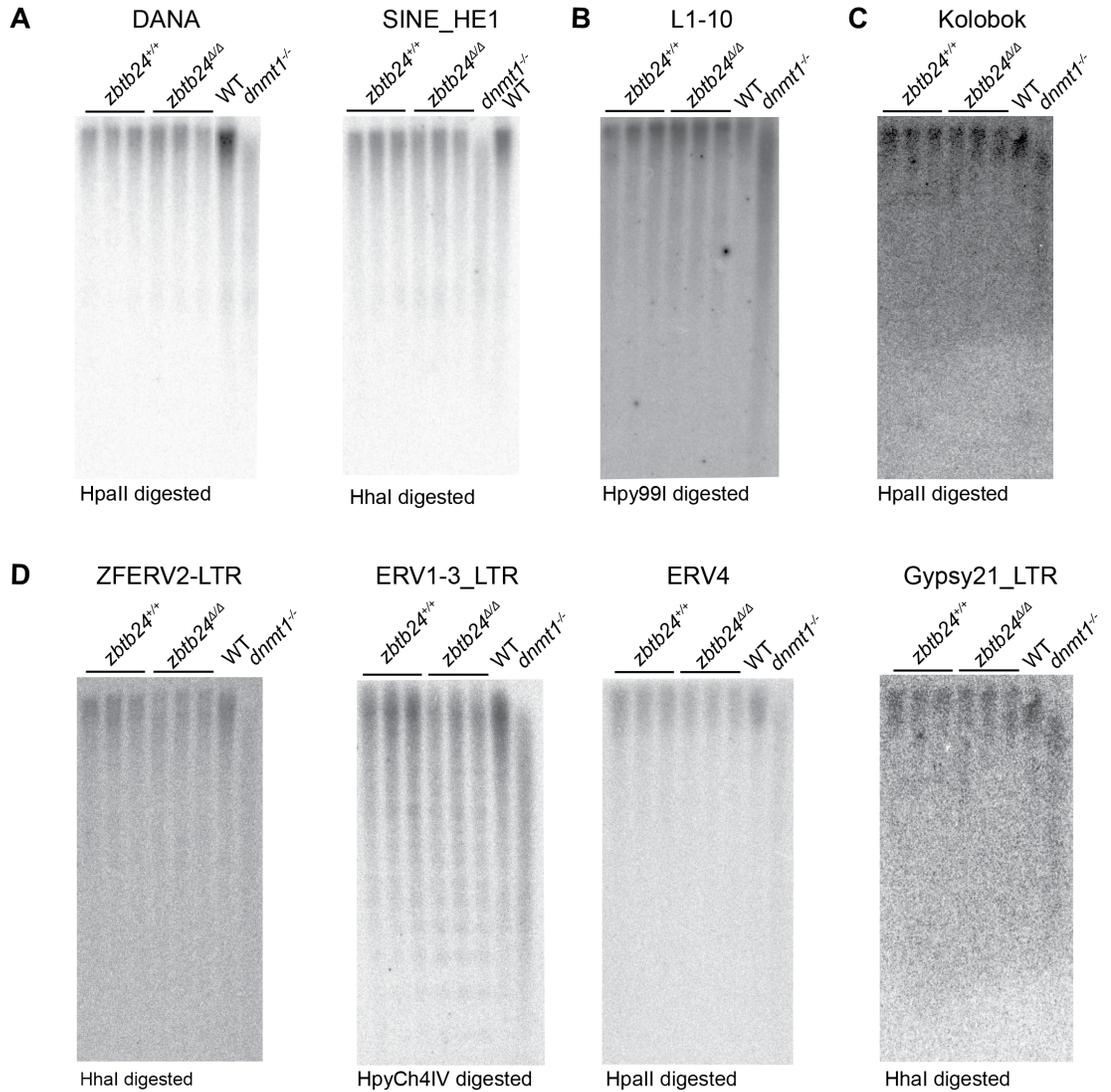
**Figure 3.4: *zbtb24* mutants exhibit modest reductions in 5mC at non-pericentromeric sequences.**

(A-B) Southern Blot and quantification of HpyCH4IV digestion at Sat1 sequences in fin tissues from 6-month-old adult zebrafish used for ERRBS. (C) Violin Plots indicating overall CpG methylation levels in fins from 6-month-old adult *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish. (D) Correlation heatmap of CpG methylation levels in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> as assessed by ERRBS (Data reflects 3 biological replicates of each genotype). The density of CpGs increases from blue to dark red. (E) Basic statistics of ERRBS analysis in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish.



**Figure 3.5: DNA Methylation levels at different genomic classes in fins of  $zbtb24^{+/+}$  and  $zbtb24^{\Delta/\Delta}$  zebrafish.**

Violin plots indicating CpG methylation in fins from 6-month-old adult  $zbtb24^{+/+}$  and  $zbtb24^{\Delta/\Delta}$  zebrafish at various genomic classes. The annotations are based on CGI, Repeatmasker, and RefSeq (only “cpl” entries) tracks from the UCSC genome browser (DanRer7). Promoters are defined as the region +/- 500 bp around the TSS. All CpGs had coverage  $\geq 10$ . DNA Transp.: DNA Transposons, Low complex.: Low Complexity.



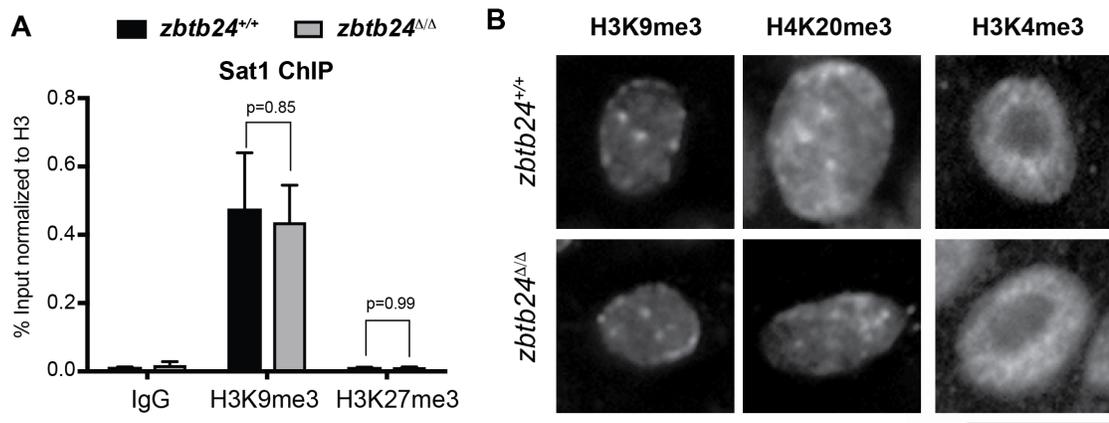
**Figure 3.6: Methylation at interspersed repeats is unaffected in *zbtb24* mutants.**

Southern blot analysis of DNA methylation at interspersed repeats (A) Short Interspersed Nuclear Element (SINE), DANA and SINE\_HE1 (B) Long Interspersed Nuclear Element (LINE), L1-10 (C) DNA Transposon, Kolobok (D) Endogenous Retroviral (ERV) elements, ZFERV2, ERV1-3, ERV4 and Gypsy21 in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> 1-month old zebrafish. Each lane represents a biological replicate of the indicated genotype. Genomic DNA from *dnmt1*<sup>-/-</sup> embryos is used as a positive control. The methylation-sensitive enzyme used to digest genomic DNA is indicated below the respective blot.

Methylation levels at endogenous retroviruses and other transposable elements were also examined by methylation sensitive restriction digest. All tested elements revealed similar resistance to digestion in *zbtb24*<sup>Δ/Δ</sup> mutant adults and wild-type siblings, indicating that these specific sequences are heavily methylated in both genotypes (**Figure 3.6**). Collectively, these data reveal limited methylation changes at genic regions with modest decreases at repeat classes outside the pericentromeric region.

## HISTONE MODIFICATION ANALYSIS

Despite significant changes in DNA methylation at satellite repeats, I did not observe any changes in histone modifications H3K9me3 or H3K27me3 at pericentromeric sequences in *zbtb24*<sup>Δ/Δ</sup> zebrafish when assayed at 1 month using chromatin immunoprecipitation (ChIP) (**Figure 3.7 A**). At this stage Sat1 repeats are significantly depleted for 5mC in *zbtb24*<sup>Δ/Δ</sup> zebrafish. I further



**Figure 3.7: Histone Modifications are unaffected in *zbtb24*<sup>Δ/Δ</sup> mutants**

(A) ChIP-qPCR analysis of H3K9me3 and H3K27me3 at Sat1 repeats in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup>. Chromatin was collected at 1 month post fertilization. Error bars indicate SEM with 3 biological replicates. (B) Immunofluorescence analysis of H3K9me3, H4K20me3 and H3K4me3 in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup>.

observed no change in the repressive histone modifications H3K9me3 and H4K20me3 or the activating histone modification H3K4me3 when assayed through immunofluorescence (**Figure 3.7 B**). These data suggest that DNA methylation status does not influence enrichment of these modifications at pericentromeres. This is consistent with a study that shows that H3K9me3 is not lost in *Dnmt1* deficient or *Dnmt3a/Dnmt3b* double-deficient mouse ES cells that show impaired DNA methylation at satellite repeats (Lehnertz et al., 2003).

## DISCUSSION

The progressive loss of 5mC we observe in somatic tissues between larval and adult stages implicates *Zbtb24* in regulating the long-term maintenance of methylation at pericentromeric repeats. It is also possible that *Zbtb24* functions in both maintenance and establishment of pericentromeric methylation, but that requirements for establishment are masked by maternally deposited RNA in *zbtb24* mutant zebrafish lines. Unfortunately, *zbtb24* homozygous mutant zebrafish are sterile, preventing the generation of the maternal-zygotic mutants required to address this question.

We note that the onset of ICF-like growth defects in *zbtb24* mutant zebrafish emerged in the weeks following *Sat1* methylation loss. In at least one case of ICF syndrome type 2, growth reductions and immunodeficiency were also reported to develop with age, raising the possibility that similar progressive methylation loss may impact ICF etiology in humans (von Bernuth et al., 2014). Knockdown of ZBTB24 using shRNAs inhibits replication of a B-cell line (Liang et al., 2016). It is likely that the developmental growth defects observed in

*zbtb24* mutants is a consequence of decreased replication from pericentromeric hypomethylation. While methylation at satellite repeats has not been tested in the same ICF individual over time, I predict progressive methylation loss to be observed in humans too, contributing to development of growth defects with age. Whether and how pericentromeric hypomethylation causes immunodeficiency in ICF syndrome is still an open question.

Methylation levels at pericentromeric Sat1 sequences could not be quantified by ERRBS, as this technique relies on MspI restriction digest to enrich for CpG containing sequences, and zebrafish Sat1 repeats are lacking in this restriction site. For highly repetitive sequences, methylation sensitive restriction digest followed by Southern blot remains the most effective way to assess methylation levels. By this approach, we observed increases in HpyCH4IV digestion that are consistent with up to 95% reductions in methylation at Sat1 pericentromeric repeats in *zbtb24* mutants. ERRBS analysis suggested that the general methylation landscape in human ICF syndrome and in *zbtb24* mutant zebrafish is similar. Methylation analysis of primary blood from ICF patients identified methylation changes of greater than 20% at roughly 3% of examined CpG dinucleotides. Significant changes in methylation of retroviruses and other dispersed repeats were not observed in these patients (Velasco et al., 2018). Consistent with these findings, our ERRBS analysis revealed methylation changes of greater than 20% at roughly 1.3% of assayed CpG dinucleotides and found methylation of dispersed repeats to be similar between wildtype and in *zbtb24* mutant zebrafish. The low-level methylation changes outside of the pericentromeres observed in ICF syndrome and our mutants raise the possibility that *zbtb24* may have additional modest roles in maintaining methylation at non

pericentromeric sequences. One important caveat of ERRBS analysis is that CpG poor sequences can be under represented, leaving open the possibility that additional DMRs in CpG poor regions of the genome were overlooked by our approach.

While similar hypomethylation was observed in all adult somatic tissues that we examined, we unexpectedly observed that methylation levels in sperm from *zbtb24* mutants and wildtype animals appeared comparable. This finding raises the possibility that different pathways act to control pericentromeric methylation in germ and somatic cells and *zbtb24* is only required for regulating DNA methylation in somatic cells. Alternatively, such a difference between somatic and germ cell methylation in *zbtb24* mutants could be zebrafish-specific due to the differences in sperm chromatin in zebrafish compared to humans or mice (Ward and Coffey, 1991; Wu et al., 2011). Assessment of methylation state at satellite repeats in human sperm from ICF patients can help reveal whether such differences are species-specific.

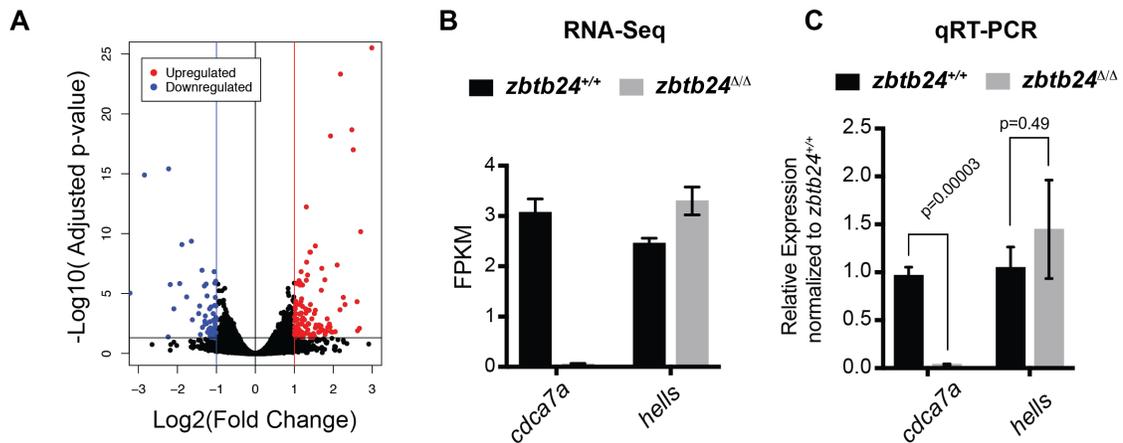
## Chapter 4 : PERICENTROMERIC HYPOMETHYLATION ELICITS AN INTERFERON RESPONSE

### PREFACE

The zebrafish *zbtb24* mutants described in this thesis provide a previously unavailable animal system to investigate the functions of DNA methylation at pericentromeric sequences. These highly repetitive, pericentromeric sequences are enriched in DNA methylation in vertebrate genomes, and are particularly susceptible to methylation loss in disease. However, the significance of this methylation loss is not well understood. This study describes the first animal model that exhibits pericentromeric hypomethylation without extensive genome-wide hypomethylation. The progressive loss of methylation observed in this model further allows for elucidation of primary vs secondary consequences of hypomethylation at these sequences.

### INTERFERON RESPONSE IN *zbtb24* MUTANTS

To gain insights into the early consequences of methylation loss in *zbtb24* mutants, I performed transcriptome analysis on total RNA isolated from wild-type and *zbtb24<sup>Δ/Δ</sup>* zebrafish at 2 wpf. At this stage, *zbtb24<sup>Δ/Δ</sup>* mutants remain morphologically indistinguishable from wildtype, but show clear hypomethylation of pericentromeric sequences. RNA-seq identified 58 genes that were downregulated by more than 2-fold in *zbtb24<sup>Δ/Δ</sup>* larvae at 2 wpf, while 119 were upregulated by 2-fold or more (**Figure 4.1 A**). No gene enrichment signature was observed among downregulated genes. However, I did observe near complete loss of *cdca7* expression in *zbtb24* mutants in our RNA-seq data

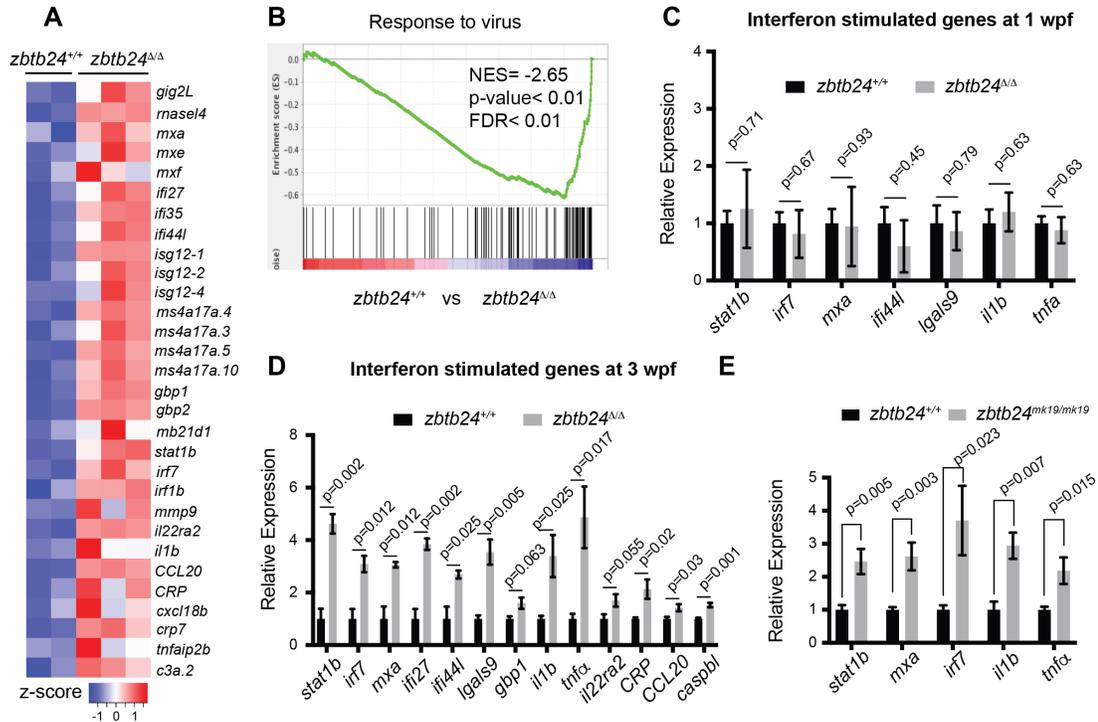


**Figure 4.1: *cdca7a* is downregulated in *zbtb24* mutants**

(A) Volcano plot representation of differential gene expression in *zbtb24*<sup>+/+</sup> vs *zbtb24*<sup>Δ/Δ</sup> zebrafish at 2 wpf. Blue and red points mark genes with >2-fold downregulation or upregulation respectively. (B) Expression levels of *cdca7* and *hells* through RNA Seq (n=3 biological replicates for each group). Error bars indicate SEM. (C) qRT-PCR analysis of *cdca7* and *hells* mRNA in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 3 wpf (n=4 biological replicates for each group). Error bars indicate SEM.

set and by qRT-PCR (Figure 4.1 B-C). This is consistent with previous studies that have identifies *Zbtb24* as a transcription factor required for the expression of *Cdca7* (Thompson et al., 2018; Wu et al., 2016).

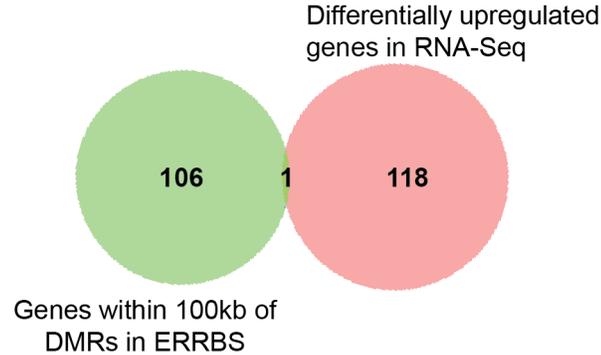
However, roughly 30% of upregulated genes were associated with activation of the innate immune system. In particular, upregulated transcripts included those associated with interferon stimulated genes (ISGs) and inflammatory cytokines (Figure 4.2 A). Consistent with these observations, Gene Set Enrichment Analysis (GSEA) identified significant enrichment of genes involved in viral response, a key function of innate immune pathways (Figure 4.2 B). Upregulation of ISGs was also observed in *zbtb24*<sup>Δ/Δ</sup> and *zbtb24*<sup>mk19/mk19</sup> mutants by qRT-PCR at 3 wpf (Figure 4.2 D-E), whereas the same genes were expressed at wild-type levels at 1 wpf (Figure 4.2 C).



**Figure 4.2: Mutation of *zbtb24* leads to activation of innate immune response genes.**

(A) RNA-seq heatmap showing innate immune genes upregulated in *zbtb24*<sup>Δ/Δ</sup> mutant compared to *zbtb24*<sup>+/+</sup> siblings at 2 wpf. Shown are Z-score normalized gene expression values. (B) GSEA of a set of genes involved in Response to Virus in zebrafish comparing *zbtb24*<sup>+/+</sup> vs *zbtb24*<sup>Δ/Δ</sup>. NES, normalized enrichment score; FDR, false discovery rate. (C) qRT-PCR analysis reveals similar expression of interferon genes in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> larvae at 1 wpf. Error bars represent SEM from at least 5 biological replicates. (D) qRT-PCR demonstrating upregulated interferon and inflammatory response genes in *zbtb24*<sup>Δ/Δ</sup> mutants at 3 wpf. Expression levels are reported relative to b-actin. Error bars indicate SEM from at least 3 independent biological replicates with n=8 total animals for each replicate. (E) qRT-PCR validation of genes that are part of the innate immune response pathway that are also upregulated in *zbtb24*<sup>mk19/mk19</sup> zebrafish at 3 wpf. Error bars indicate SEM from 5 biological replicates.

No immune-related genes (and only 1 gene differentially upregulated in the RNA-Seq) were found within 100kb of identified DMRs, suggesting that direct loss of methylation at these sequences was unlikely to cause the response (Figure 4.3).



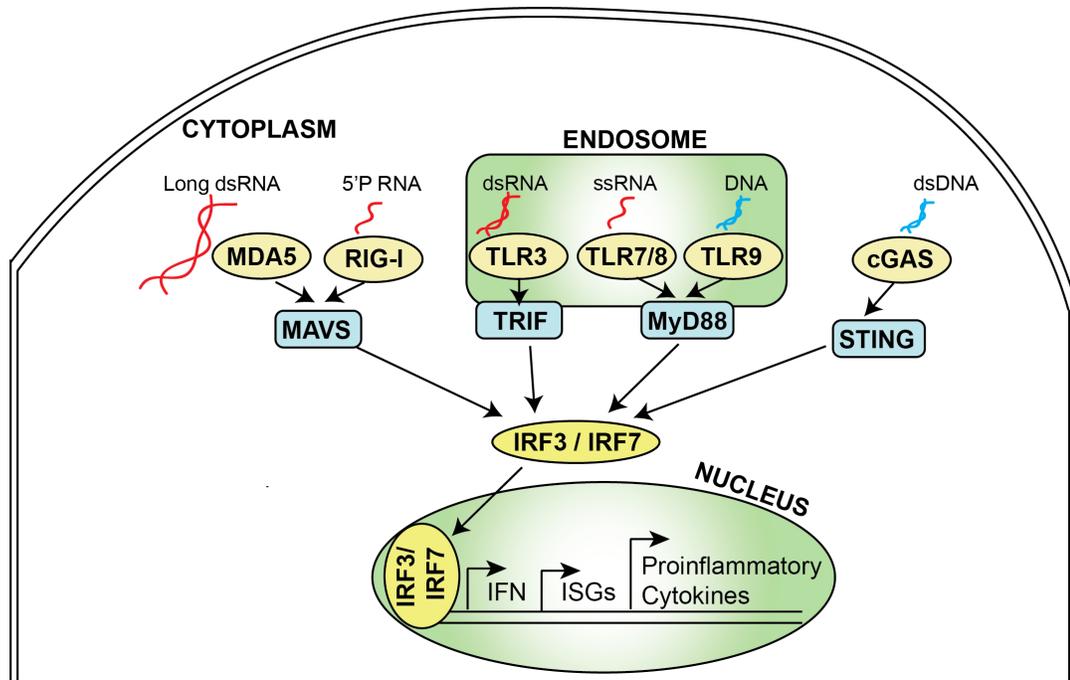
**Figure 4.3: Methylation at promoters of innate immune genes upregulated in *zbtb24* mutants is unchanged.**

Venn diagram summarizing the overlap between differentially upregulated genes from RNA-Seq (right circle) and genes within 100kb of all DMRs in ERRBS tested for differential expression (left circle).

## **MECHANISM FOR INTERFERON RESPONSE IN *zbtb24* MUTANTS**

### **Introduction to interferon response**

The interferon response is an innate immune response that plays an essential role in defense against pathogen infection via induction of interferon stimulated genes (ISGs) through Janus kinase (JAK)–signal transducer and activator of transcription (STAT) signaling pathway. The interferon response is triggered by the cellular detection of pathogen-associated molecular patterns (PAMPs) in viruses, bacteria, fungi and protozoa. These PAMPs are detected by molecules called pattern-recognition receptors (PRRs). Following detection of a PAMP by its associated PRR, signaling cascades are induced that drive the production of interferons and expression of ISGs and other inflammatory cytokines with antiviral and immune modulating functions (Schneider et al., 2014).



**Figure 4.4: Schematic for activation of interferon response by nucleic acids.**

The type I interferon response is activated by pattern recognition receptors (PRRs), which recognize pathogen-associated molecular patterns (PAMPs) such as viral nucleic acids. PRR signals are transduced to transcription factor activity via adaptor molecules. Activation of such adaptor molecules leads to activation of IRF3/7 which trigger the expression of interferons (IFN) and interferon stimulated genes (ISGs) with anti-viral roles. dsRNA: double stranded RNA; 5'P RNA: RNA with 5'-triphosphate caps; ssRNA: single stranded RNA; dsDNA: double stranded DNA; MDA5: Melanoma Differentiation-Associated protein 5; RIG-I: retinoic acid-inducible gene I; MAVS: Mitochondrial antiviral-signaling protein; TLR: Toll-Like Receptor; TRIF: TIR-domain-containing adapter-inducing interferon- $\beta$ ; Myd88: Myeloid differentiation primary response 88; cGAS: Cyclic GMP-AMP synthase; STING: Stimulator of interferon genes; IRF: Interferon Regulatory Factor.

Characteristic features of PAMPs are that they are important products of biosynthetic pathways in the lifecycle of the pathogen and they lack the ability to evade immune detection due to genetic mutability. Targeting viral proteins is inefficient since viral proteins evolve rapidly without being functionally compromised by mutation. From an evolutionary standpoint, it was a better

investment for an innate immune system to target viral nucleic acids instead leaving the adaptive immune system to target proteins (Medzhitov, 2007)

Nucleic acid detection is accomplished by an array of PRRs, each specializing in the detection of distinct nucleic acids. For instance,

- the cytosolic RIG-I-like receptors (RLR) RIG-I recognized 5'-triphosphorylated RNA and MDA5 recognize cytosolic long double stranded RNAs (dsRNAs) (Yoneyama et al., 2005),
- the endosomal Toll-like receptors (TLR) TLR3, TLR7/8 and TLR9 recognize endosomal dsRNA, ssRNA or unmethylated DNA respectively (Kawai and Akira, 2010),
- cGAS recognizes dsDNA (Sun et al., 2013).

PRR signals are transduced to transcription factor activity via adaptor molecules such as mitochondrial antiviral-signaling protein (MAVS), myeloid differentiation factor-88 (MYD88) or stimulator of IFN genes (STING).

- MAVS is the adaptor molecule for RLRs that detect cytosolic RNA (Kawai et al., 2005),
- Myd88 is the adaptor molecule for endosomal TLR 7/8/9. TRIF is the adaptor molecule for TLR3 (Medzhitov et al., 1998),
- STING is the adaptor molecule for cytosolic DNA detection by cGAS (Ishikawa et al., 2009).

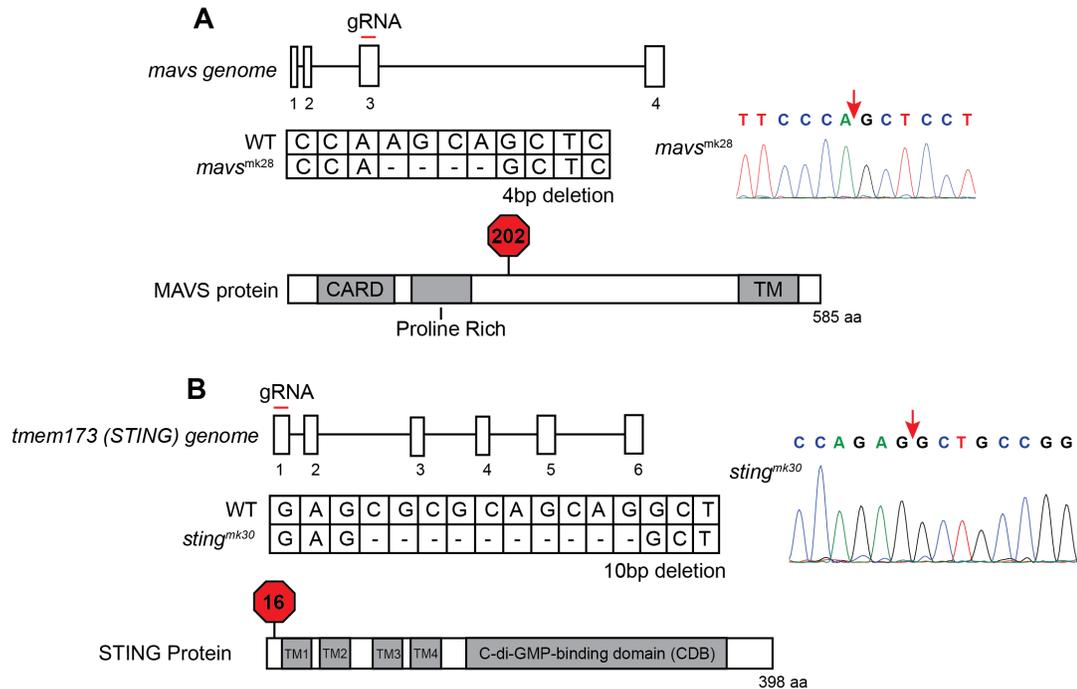
Activation of such adaptor molecules leads to phosphorylation of interferon (IFN) response factors 3 or 7 (IRF3/7). Phosphorylated dimers of IRF3/7 translocate to the nucleus, where they bind to and activate specific promoters, triggering expression of IFN as well as a subset of ISGs. These ISGs include IRFs, PRRs and antiviral effectors (Schneider et al., 2014) (**Figure 4.4**).

However, nucleic acids are not unique to viruses and thus aberrantly accumulated endogenous nucleic acids can trigger autoimmunity via the viral detection machinery that activates an interferon response. Several studies have reported that in addition to extracellular pathogens, PRRs also recognize PAMPs associated with cell-intrinsic stimuli including DNA damage, endogenous retroviral RNA and RNA-DNA hybrids (Chiappinelli et al., 2015; Hartlova et al., 2015; Mankan et al., 2014; Roulois et al., 2015). Recent studies have further linked global hypomethylation to activation of antiviral signaling pathways in zebrafish mutated for *dnmt1* and in cancer cells treated with the DNA methyltransferase inhibitor 5-azacytidine (Chernyavskaya et al., 2017; Chiappinelli et al., 2015; Roulois et al., 2015). In these studies, the interferon response was attributed to the upregulation of endogenous retroviral elements (ERVs) due to hypomethylation of these sequences. These parasitic transposons triggered an interferon response through viral mimicry detected by the MDA5-MAVS dsRNA detection machinery.

### **The innate immune response in *zbtb24* mutants is mediated by sensors of cytosolic RNA**

To clarify the origin of the interferon response in *zbtb24* mutants, I examined the major families of PRRs involved in innate immunity. Using CRISPR-Cas9, I introduced mutations in adaptor proteins that are activated via different PRRs to cause downstream activation of ISGs. These key adaptor proteins are conserved in structure and function in zebrafish (Li et al., 2017). Mutations in MAVS (*mavs<sup>mk28</sup>*) eliminates the C-terminal transmembrane domain that is required for interferon induction (Seth et al., 2005) (**Figure 4.5 A**) Mutation in

STING (*sting*<sup>mk30</sup>) takes out all domains (Figure 4.5 B). A mutant allele for Myd88 (*myd88*<sup>hu3568</sup>) required for signaling through most TLRs has previously been described (van der Vaart et al., 2013).

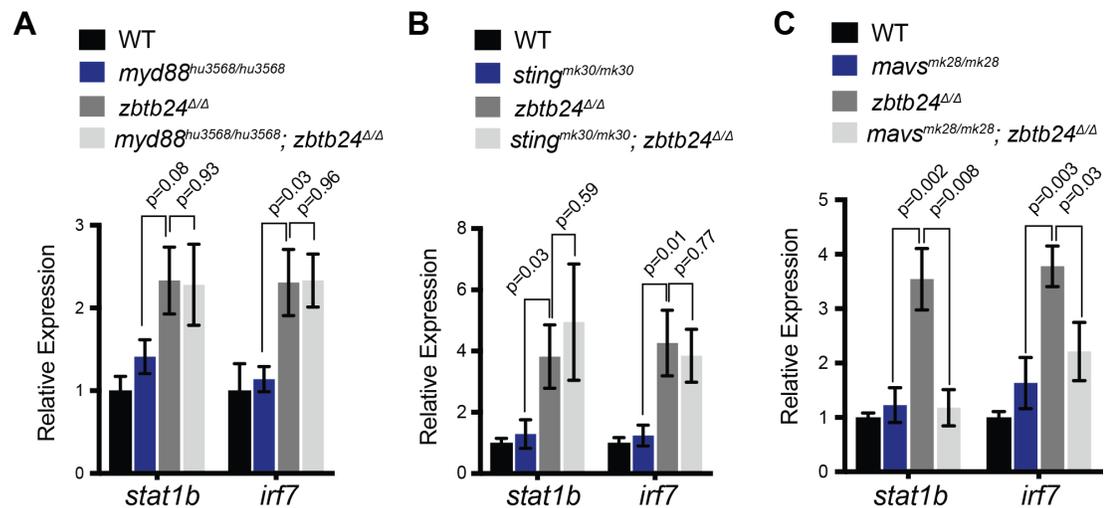


**Figure 4.5: Mutation of zebrafish orthologs of *mavs* and *sting*.**

(A) Schematic of mutation in *mavs* including position of targeted gRNA, sequence information of the mutation, and location of predicted STOP codon. (B) Schematic of mutation in *sting* including position of targeted gRNA, sequence information of the mutation, and location of predicted STOP codon.

As in prior experiments, significant increases of the ISGs, signal transducer and activator of transcription 1b (*stat1b*) and interferon regulatory factor (*irf7*) were observed in *zbtb24*<sup>Δ/Δ</sup> larvae at 3 wpf by qRT-PCR. Introduction of *myd88* or *sting* mutations had little impact on expression of these ISGs, as similar transcript levels were detected in *zbtb24*<sup>Δ/Δ</sup> single mutant animals compared to *myd88*<sup>hu3568/hu3568</sup>, *zbtb24*<sup>Δ/Δ</sup> (Figure 4.6 A) or *sting*<sup>mk30/mk30</sup>, *zbtb24*<sup>Δ/Δ</sup> double

mutants (**Figure 4.6 B**). Sustained ISG expression in these double mutants suggests limited roles for TLR and cGAS PRRs in mediating the interferon response in *zbtb24* mutants. In contrast to *myd88* and *sting*, mutation of *mavs* suppressed *stat1b* and *irf7* upregulation in *zbtb24*<sup>Δ/Δ</sup> mutant animals. Expression levels of *irf7* and *stat1b* were reduced 2- and 4-fold respectively in *mavs*<sup>mk28/mk28</sup>; *zbtb24*<sup>Δ/Δ</sup> double mutants when compared to *zbtb24*<sup>Δ/Δ</sup> single mutant zebrafish (**Figure 4.6 C**), indicating a requirement for *mavs* and suggesting a role for cytosolic RNA in the upregulation of these ISGs.



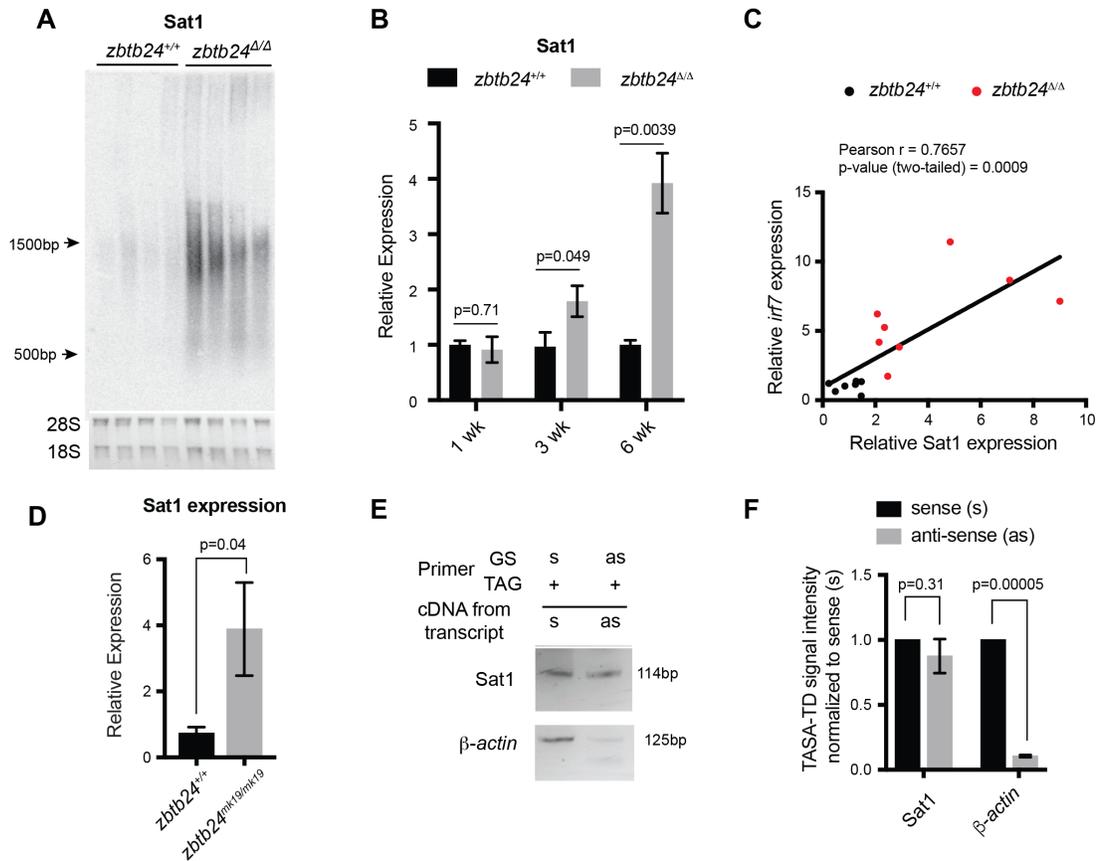
**Figure 4.6: Interferon response in *zbtb24* mutants is mediated by sensors of cytosolic RNA.**

(A) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 3 wpf. n = 4 biological replicates. (B) Expression of the ISGs *stat1b* and *irf7* in indicated genotypes at 3 wpf. n ≥ 7 biological replicates. (C) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 3 wpf. n ≥ 5 biological replicates. All error bars indicate SEM.

### **Satellite RNA is upregulated in *zbtb24* mutants**

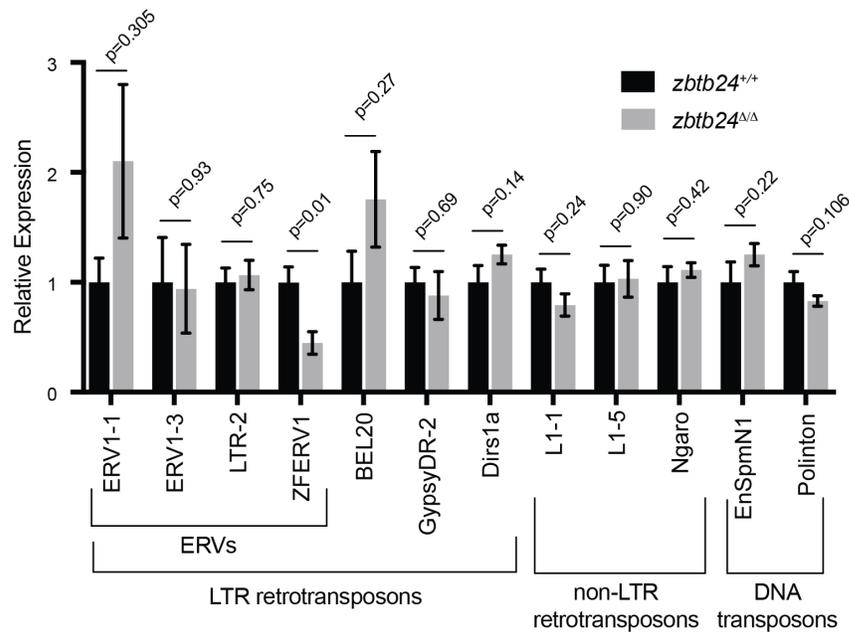
Given known roles for DNA methylation in transcriptional repression, I next tested whether loss of methylation at pericentromeric sequence resulted in increased levels of Sat1 transcripts that could trigger the RNA mediated interferon response.

Consistent with this model, strong derepression of Sat1 RNA from hypomethylated pericentromeres was noted in *zbtb24* mutant adults (**Figure 4.7 A and D**). Increases in Sat1 transcripts correlated with levels of *irf7* expression in adult zebrafish ( $r=0.77$ ), and upregulation of Sat1 transcripts coincided with the window of ISG induction during development (**Figure 4.7 B-C**). Both sense and antisense transcripts were detected in mutants using TAG-aided sense/antisense transcript detection (TASA-TD) strand-specific PCR (Henke et al., 2015), suggesting the potential for derepressed Sat1 transcripts to form double stranded RNAs (**Figure 4.7 D-E**). Whereas transcripts for other dispersed repetitive elements remained unchanged between mutants and wildtype (**Figure 4.8**).



**Figure 4.7: Pericentromeric Sat1 transcripts are upregulated in *zbtb24* mutants.**

(A) Northern blot analysis of Sat1 transcripts in *zbtb24*<sup>+/+</sup> vs *zbtb24*<sup>Δ/Δ</sup> zebrafish at 6 wpf. Each lane represents a biological replicate. The lower panel represents the cropped ethidium-bromide stained gel as loading control. (B) qRT-PCR for Sat1 transcripts in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 1, 3 and 6 wpf. Error bars indicate SEM of at least 4 biological replicates in each group. (C) Correlation between the expression of Sat1 and *irf7* in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> at 6 weeks (n=15). (D) qRT-PCR of Sat1 RNA in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> zebrafish at 6 wpf. Error bars indicate SEM from at least 3 biological replicates. (E) TASA-TD PCR amplified sense (s) and antisense (as) transcripts Sat1 (114 bp) and  $\beta$ -actin (125bp) from first strand *zbtb24*<sup>Δ/Δ</sup> cDNA at 6 wpf. PCR primers: gene-specific (GS); TAG. The products from TASA-TD PCR were run on the same gel, then cropped and presented. (F) Quantification of TASA-TD from panel E. Error bars indicate SD from 2 biological replicates.

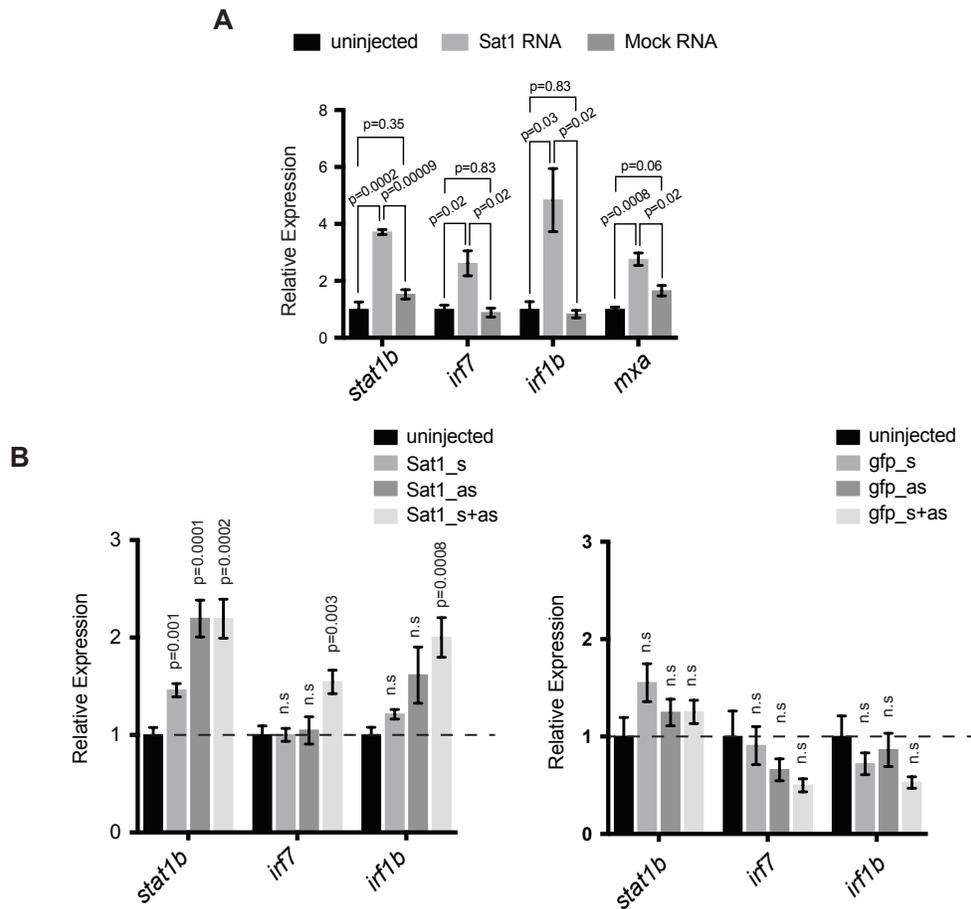


**Figure 4.8: Mutation in *zbtb24* upregulates *Sat1* transcripts but not transposons.**

qRT-PCR reveals similar levels of expression from transposable elements in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 6 wpf. Error bars indicate SEM from 4-8 biological replicates. ERVs: Endogenous Retroviruses; LTR: Long Terminal Repeats.

### Satellite RNA is sufficient to trigger an interferon response

To determine whether *Sat1* transcripts were sufficient to activate an innate immune response, in vitro synthesized RNA corresponding to *Sat1* sense and antisense transcripts were injected into wild-type embryos at the 1-cell stage. Expression of the ISGs *stat1b*, *irf7*, *irf1b* and *mxr* was then assessed at 8 hours post fertilization. Co-injection of sense and antisense *Sat1* RNA was sufficient to reproducibly cause a 2 to 4-fold upregulation in expression of these ISGs, whereas combined injection of sense and antisense control transcripts encoding a fragment of zebrafish  $\beta$ -actin or GFP had no effect on expression of these genes (**Figure 4.9**). Lower level upregulation of some, but not all ISGs was



**Figure 4.9: Sat1 RNAs are sufficient to elicit an interferon response.**

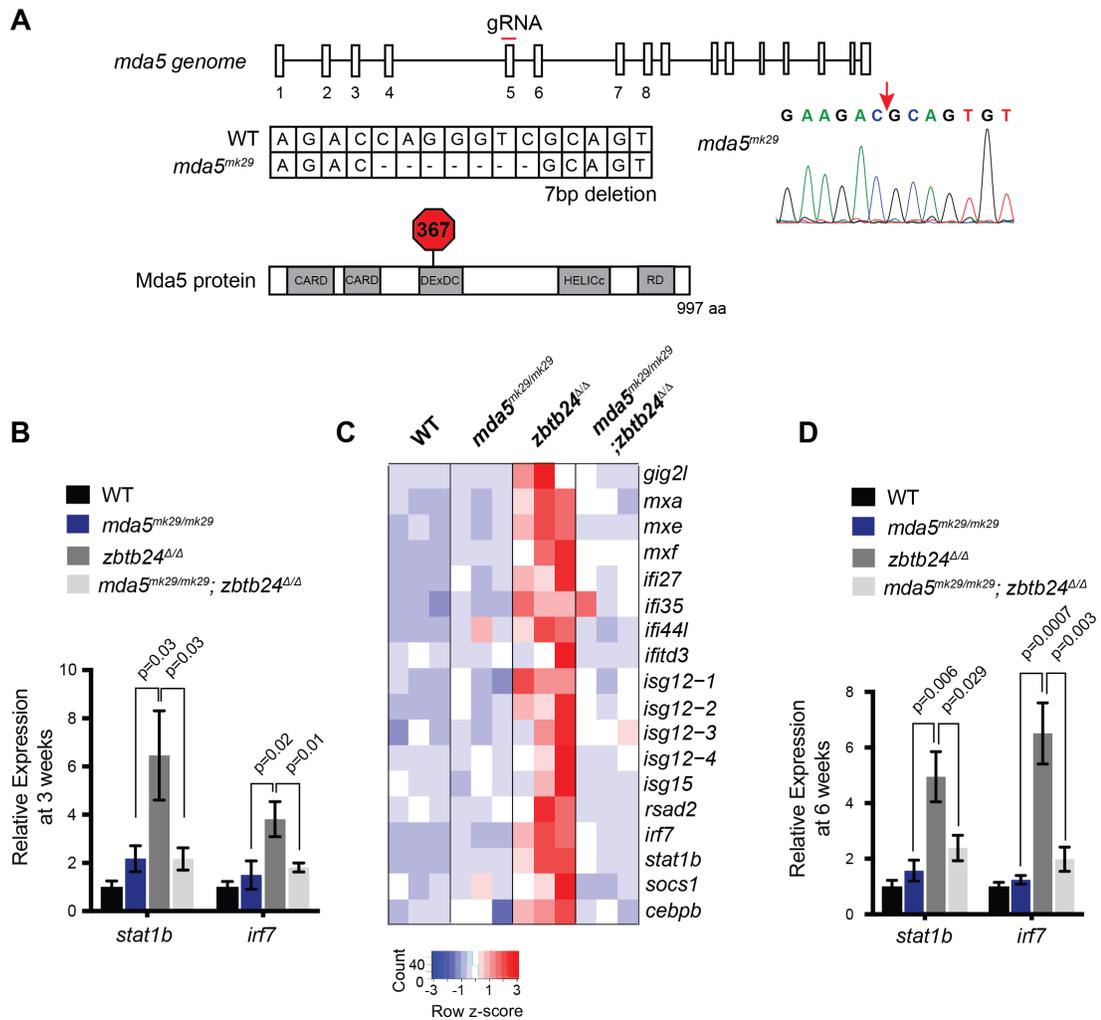
(A) Expression of interferon stimulated genes *stat1b*, *irf7*, *irf1b* and *mxs* in wildtype embryos injected with Sat1 or control RNA. 50pg of in vitro transcribed sense and antisense transcripts were injected into wildtype zebrafish embryos at the 1-cell stage. Total RNA was extracted at 8 hpf for qRT-PCR analysis. Error bars indicate SEM from at least 3 biological replicates with  $n=20$  embryos for each biological replicate. (B) Expression of interferon stimulated genes in wild-type embryos injected with sense (Sat1\_s), anti-sense (Sat1\_as) or an equimolar mixture of sense and anti-sense Sat1 transcripts (Sat1\_s+as). Transcripts of similar size encoding a fragment of GFP are used as a control. 50pg of in vitro transcribed transcripts were injected into wild-type zebrafish embryos at the 1-cell stage. Total RNA was extracted at 8 hours post fertilization for qRT-PCR analysis. Error bars indicate SEM from 4 to 8 biological replicates with  $n=20$  embryos for each biological replicate. p-values are adjusted for multiple comparisons using the Holm-Sidak method, tested against uninjected control. n.s: Not significant.

noted when sense or antisense Sat1 transcripts were individually injected into the embryo, suggesting that the response was primarily triggered by formation of Sat1 dsRNA (**Figure 4.9 B**). Collectively, these results functionally link the derepression of Sat1 transcripts to the activation of the innate immune response in *zbtb24* mutants.

### **The cytosolic RNA helicase Mda5 is required for the Interferon Response in *zbtb24* mutants**

Finally, I sought to identify the specific PRR required for the interferon response in *zbtb24* mutants. The RLR family of PRRs consists of two RNA helicases that signal through Mavs: Melanoma Differentiation-Associated protein 5 (Mda5) and Retinoic acid-inducible gene I (Rig-I). Rig-I binds 5' triphosphorylated RNA molecules, whereas Mda5 recognizes long double-stranded RNAs in the cytosol (Crowl et al., 2017). Given that 5' triphosphorylation of RNAs is a typical viral signature that is unlikely to be present on endogenous RNA transcripts, I reasoned that Mda5 was a more likely candidate for the receptor. To test the requirement for *mda5*, I generated a 7 base-pair deletion in this gene that disrupted the DEAD box helicase domain (**Figure 4.10 A**). This *mda5*<sup>mk29</sup> allele was then introduced onto the *zbtb24* mutant background, and expression of the ISGs *stat1b* and *irf7* was examined at 3 wpf and 6wpf (**Figure 4.10 B and D**). Homozygous mutation of *mda5* was sufficient to restore *stat1b* and *irf7* expression to wild-type levels in *zbtb24*<sup>Δ/Δ</sup> mutant larvae, suggesting that Mda5 is the primary PPR required for the response. This requirement was further validated by RNA-seq, which revealed that a broad panel of ISGs that showed

elevated expression in *zbtb24* single mutants were no longer upregulated in *mda5<sup>mk29/mk29</sup>; zbtb24<sup>Δ/Δ</sup>* double mutants (**Figure 4.10 C**)

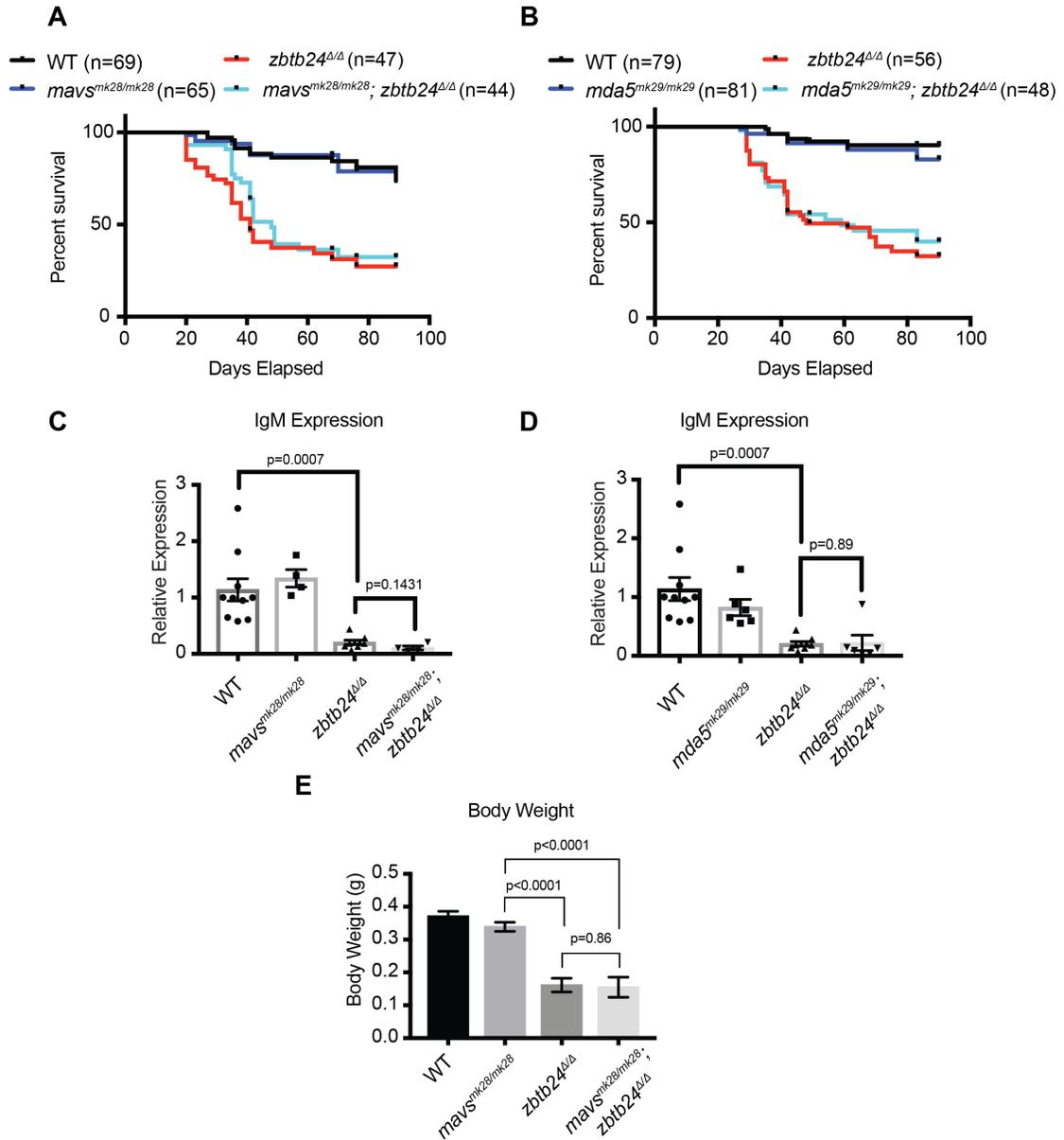


**Figure 4.10: Mutation of cytosolic dsRNA receptor Mda5 mitigates the interferon response in *zbtb24<sup>Δ/Δ</sup>* zebrafish.**

(A) Schematic of mutation in *mda5* including position of targeted gRNA, sequence information of the mutation, and location of predicted STOP codon. The site of deletion is indicated with a red arrow on the sequence trace. (B) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 3 wpf.  $n \geq 7$  biological replicates. (C) RNA-seq heatmap of interferon stimulated genes upregulated in *zbtb24<sup>Δ/Δ</sup>* zebrafish and rescued in *mda5<sup>mk29/mk29</sup>; zbtb24<sup>Δ/Δ</sup>* zebrafish at 3 wpf. Shown are Z-score normalized gene expression values. (D) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 6 wpf.  $n=6$  biological replicates.

## **Co-mutations in Interferon activation genes do not rescue ICF Syndrome Phenotypes in *zbtb24* mutants**

The interferon response reported in *zbt24* mutant zebrafish in this study has not been tested for in ICF syndrome human patients. To test the possibility that such an interferon response contributed to ICF Syndrome phenotypes, I assayed for key ICF syndrome phenotypic characteristics in *zbtb24* mutant zebrafish in the background of the *mda5* or *mavs* homozygous mutants. Such double mutants that demonstrated a rescue of interferon response did not show any change in life-span compared to *zbtb24* single mutants (**Figure 4.11 A-B**). Developmental growth defects and decreased body weight was similar in double mutants compared to *zbtb24* single mutants (**Figure 4.11 E**). I did not observe a change in expression of immunoglobulins among double mutants, suggesting that the interferon response was not causing immunodeficiencies (**Figure 4.11 C-D**). Collectively these data suggest that interferon response is unlikely to be the cause of ICF phenotypes in *zbtb24* mutant zebrafish, but rather, it likely represents an unappreciated consequence of pericentromeric methylation loss.

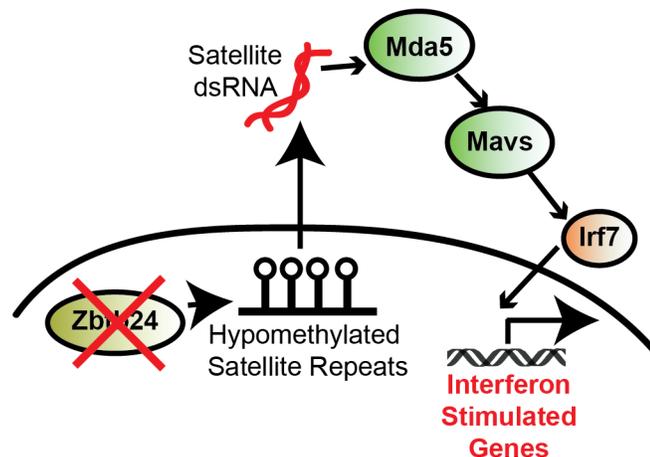


**Figure 4.11: Mitigation of interferon response in *zbtb24* mutants does not rescue ICF syndrome phenotypes.**

(A-B) Kaplan-Meier curve indicating survival among groups of zebrafish of the indicated genotype. (C-D) qRT-PCR analysis of IgM, zebrafish immunoglobulin in zebrafish of the indicated genotype measured at 6 wpf. (E) Average weight of 5-month-old adult zebrafish of the indicated genotype.

## DISCUSSION

In this study, I take advantage of the progressive Sat1 methylation loss in *zbtb24* mutants to identify activation of interferon signaling as one of the earliest *in vivo* consequences of pericentromeric hypomethylation. This phenotype cannot be attributed to defects in adaptive immunity, as the zebrafish adaptive immune system is not functional until roughly 4 wpf (Trede et al., 2004). This study implicates Mda5-Mavs RLR signaling in the activation of the innate immune system in *zbtb24* mutants and suggests a cytosolic satellite RNAs as a trigger for the response (**Figure 4.12**).



**Figure 4.12: Model for the activation of interferon response in *zbtb24* mutants.**

Loss of *Zbtb24* function causes hypomethylation of pericentromeric Sat1 repeats, which leads to depression of associated Sat1 transcripts. These pericentromeric transcripts are recognized by the dsRNA helicase Mda5 which signals through Mavs and Irf7 to upregulate ISGs. Auto-regulatory feedback implicates *irf7* as both an ISG and a key downstream effector of Mda5/Mavs signaling.

DNA hypomethylation has previously been associated with immune dysfunction related diseases. Systemic lupus erythematosus (SLE) is one of the most commonly studied autoimmune diseases. Increased apoptosis results in

accumulation of hypomethylated DNA fragments that induce SLE-like symptoms in non-susceptible mice (Wen et al., 2007). Over 25% of CpG dinucleotides in the human genome reside within Alu elements, the most abundant human repeats (Xie et al., 2009). Accumulation of RNA from Alu elements induces age-related macular degeneration by activating an inflammatory response (Tarallo et al., 2012). The role of abundant 5mC at satellite repeats has not been clear. This study identifies a previously unappreciated role for DNA methylation at pericentromeric satellite repeats in preventing autoimmunity.

Induction of an interferon response has been reported in the context of global hypomethylation in cancer cell lines treated with the DNA methyltransferase inhibitor 5-azacytidine and in zebrafish mutated for the maintenance DNA methyltransferase machinery (Chernyavskaya et al., 2017; Chiappinelli et al., 2015; Roulois et al., 2015). In each of these cases induction of the interferon response was attributed to massive derepression of endogenous retroviral elements. In this study, I functionally link the derepression of Sat1 transcripts to the activation of the innate immune response in *zbtb24* mutants. These results are distinguished from these earlier studies in that I identify hypomethylation of pericentromeric sequences and subsequent derepression of associated satellite transcripts as a previously unappreciated trigger of innate immunity. Immunostimulatory motifs have been noted in pericentromeric RNAs derived from mouse and humans, and transcripts derived from these repeats have been observed in *p53* null mouse fibroblasts following global methylation loss (Leonova et al., 2013; Tanne et al., 2015). However, while these studies suggest the potential for pericentromeric hypomethylation to drive an interferon

response in diverse vertebrate species, experimental evidence in support of this model has been lacking. Here I demonstrate a causative link between derepression of pericentromeric RNAs and the interferon response, and identify a requirement for Mda5/Mavs in mediating the response. These findings suggest that aberrant upregulation of both sense and antisense transcripts derived from pericentromeric repeats creates an abundance of double stranded RNAs within the cytosol, which mimic features of double stranded RNA viruses. This finding raises the possibility that this pathway may also recognize additional endogenous long dsRNAs that lack viral origin.

## Chapter 5 : PERICENTROMERIC DNA HYPOMETHYLATION AND GENOME INSTABILITY

### PREFACE

Loss of DNA hypomethylation from pericentromeric repeats has been observed in several tumors such as Breast adenocarcinomas (Narayan et al., 1998), (Tsuda et al., 2002), Wilms Tumors (Qu et al., 1999b) , Ovarian Epithelial Tumors (Qu et al., 1999a), Glioblastoma (Fanelli et al., 2008) and Urothelial Carcinomas (Nakagawa et al., 2005).

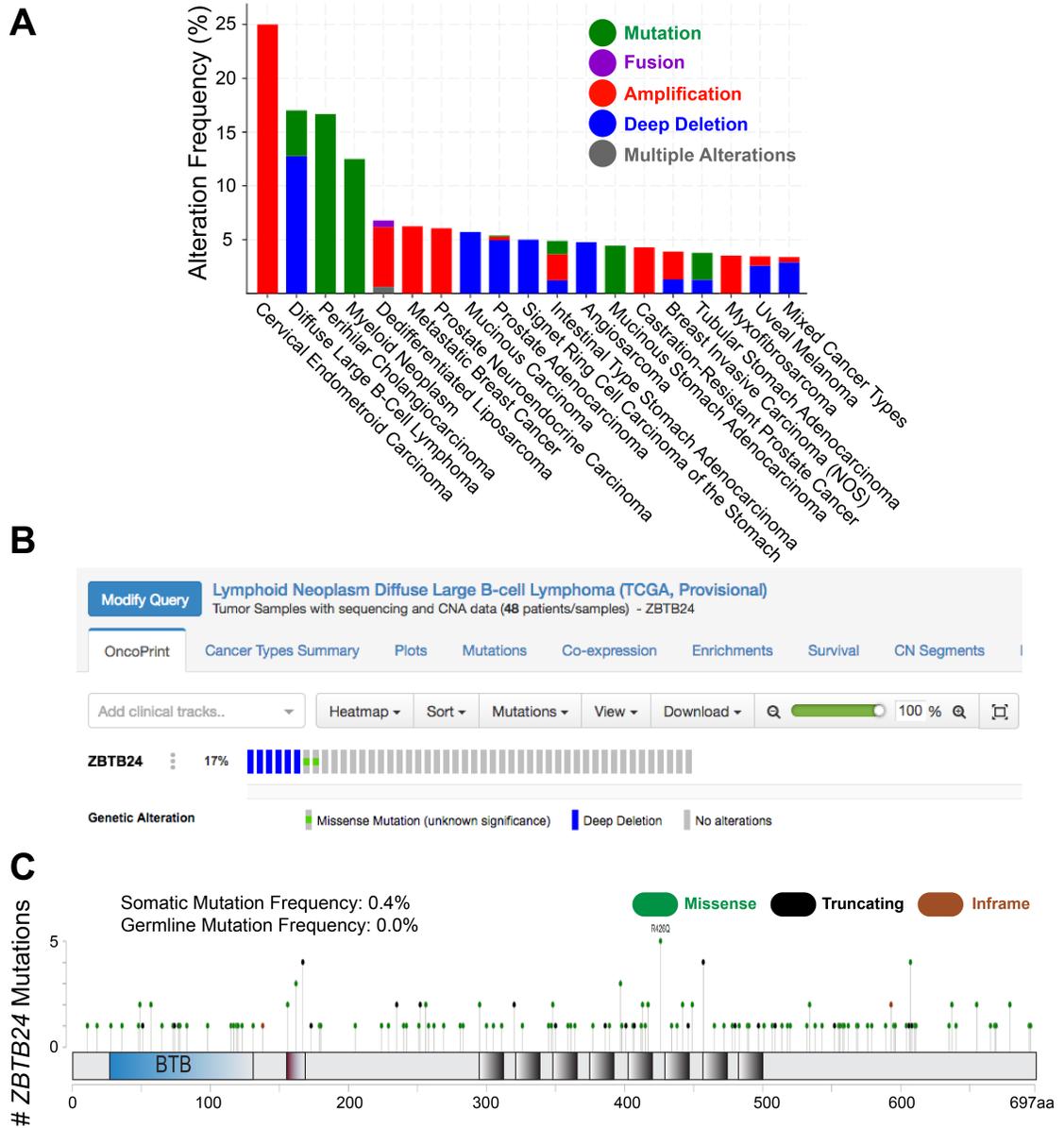
*Dnmt1* mutant mice with global genome hypomethylation including hypomethylation of satellite DNA develop aggressive thymic tumors (Gaudet et al., 2003). While this study suggests a causal role for global hypomethylation in cancer progression, the specific contribution of pericentromeric DNA hypomethylation in cancer development is unclear. One hypothesis is that pericentromeric DNA hypomethylation makes the pericentromeric region particularly vulnerable for chromosomal rearrangements which contributes to genome instability. However, a study that compared incidence of satellite hypomethylation to karyotypic instability in tumors reported incidences of hypomethylated pericentromeres without chromosomal rearrangements (Ehrlich et al., 2003). This suggests that hypomethylated pericentromeres contribute, but are not sufficient, to cause genomic instability.

Another hypothesis is that expression of non-coding satellite RNAs from hypomethylated satellite RNAs could contribute to tumor development via acting

as a sink for various anti-tumor components. Aberrant expression of satellite RNA has been observed in several cancers (Ting et al., 2011). In cancer cell lines, hypomethylated HSATII DNA sequestered polycomb group complex PRC1, which is commonly misregulated in cancers. In these cells, HSATII RNA accumulated the methyl-DNA binding protein, MeCP2, which affected broader cancer associated expression patterns (Hall et al., 2017). More recently, aberrant overexpression of satellite RNAs in mammary glands induced tumor formation in mice by sequestering BRCA1-associated protein networks required for the stabilization of DNA replication forks (Zhu et al., 2018). Collectively, these studies indicate a role of satellite RNAs emerging from hypomethylated satellite DNA in cancer.

Cytogenetic chromosomal abnormalities in mitogen-stimulated lymphocytes from peripheral blood of ICF patients have been frequently reported. These include, chromosome breaks, whole-arm deletions, multi-branched chromosomes or decondensation in chromosomes 1, 9 and 16 (Ehrlich, 2009). Although, similar chromosome anomalies are not observed in primary tissues from affected individuals. Despite these chromosomal aberrations, very few incidences of cancer have been reported in ICF Syndrome patients. One ICF2 patient developed a classical Hodgkin's lymphoma (Schuetz et al., 2007). It is likely, that pericentromeric DNA hypomethylation in ICF patients predisposes them to cancer but severely shortened lifespans preclude detection of malignancies in ICF patients.

Recently, cells with mutations in ICF genes CDCA7 and HELLS showed compromised classical non-homologous end joining (NHEJ). Cells displayed



**Figure 5.1: Gene mutation analysis of *ZBTB24* in human cancers.**

Mutations of *ZBTB24* gene in human cancers were analyzed using the cBioPortal Database. **(A)** Frequencies of *ZBTB24* alterations in human cancers. **(B)** Genetic alterations in *ZBTB24* in Lymphoid Neoplasm Diffuse Large B-cell Lymphoma. **(C)** Distribution of mutations identified in *ZBTB24* in human cancers.

increased apoptosis, abnormal chromosome segregation, centrosome amplification and significant accumulation of  $\gamma$ H2AX (Unoki et al., 2018). Although less prominent, cells with *zbtb24* mutations also showed similar defects. This defect in NHEJ could contribute to increased genome instability in ICF syndrome.

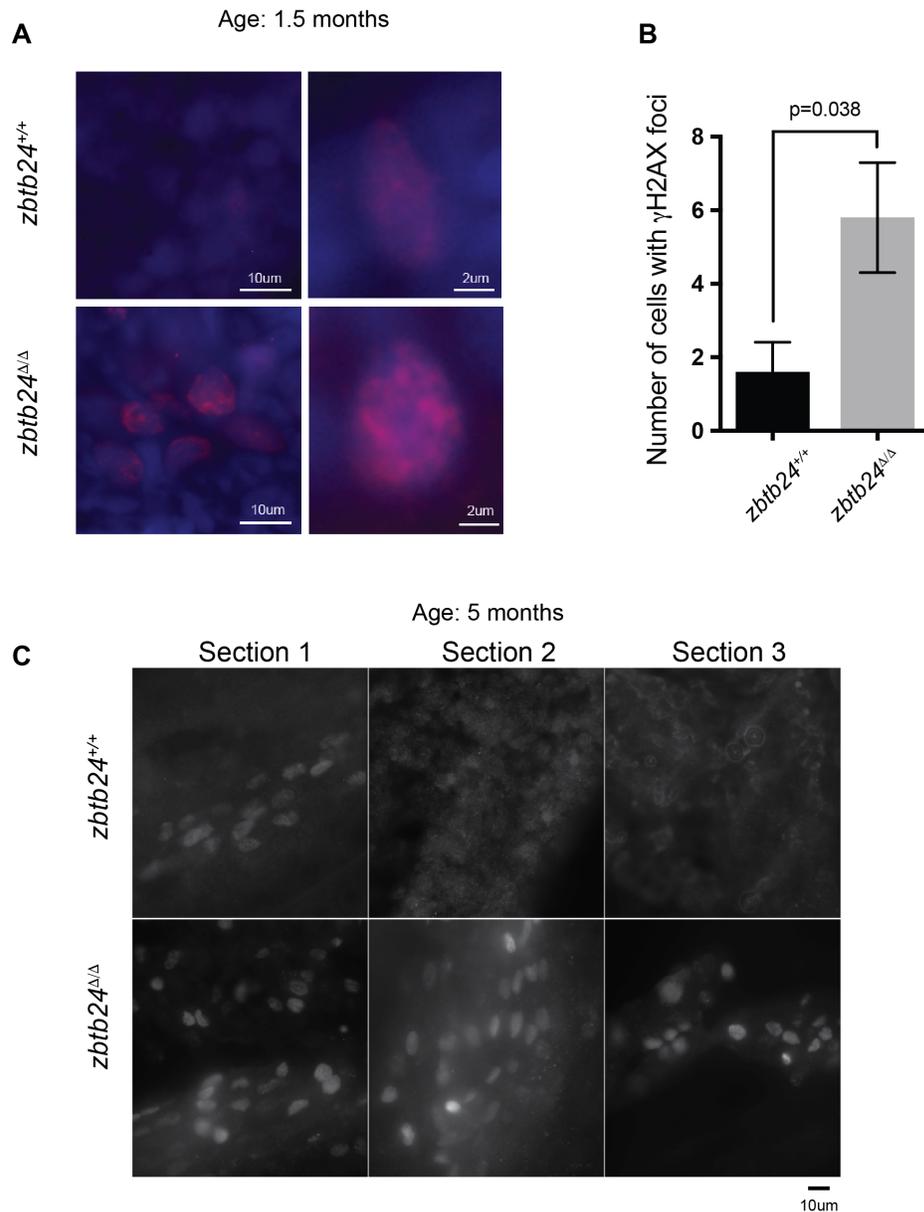
### **Somatic mutations in *ZBTB24* in cancer**

Mutation analysis using cBioPortal for mutations in *ZBTB24* in human cancers revealed that the gene was altered in ~1% of queried samples in 238 studies (**Figure 5.1 A**) and indicated ~13% frequency of deep deletions in Diffuse Large B-Cell Lymphoma (**Figure 5.1 A-B**). Somatic mutations in *ZBTB24* do not appear to show any particular hotspot mutation site (**Figure 5.1 C**). Intriguingly, *ZBTB24* maps to chromosome 6q21, a region which is frequently deleted in many types of cancer, including lymphoid malignancies, melanoma, prostate cancer, fibroadenomas, and carcinoma of breast.

Given the contribution of pericentromeric hypomethylation to genomic instability is ambiguous in cancer and ICF syndrome, I investigated any potential link between pericentromeric hypomethylation and tumorigenesis in the disease model of ICF Syndrome with mutations in zebrafish *zbtb24*.

### **EVIDENCE OF DNA DAMAGE AND CANCER INCIDENCE**

To test for increased incidence of genome instability in *zbtb24* mutants, I performed immunohistochemistry on sagittal cryosections for fixed *zbtb24* mutants along with their wildtype siblings as controls using an antibody against

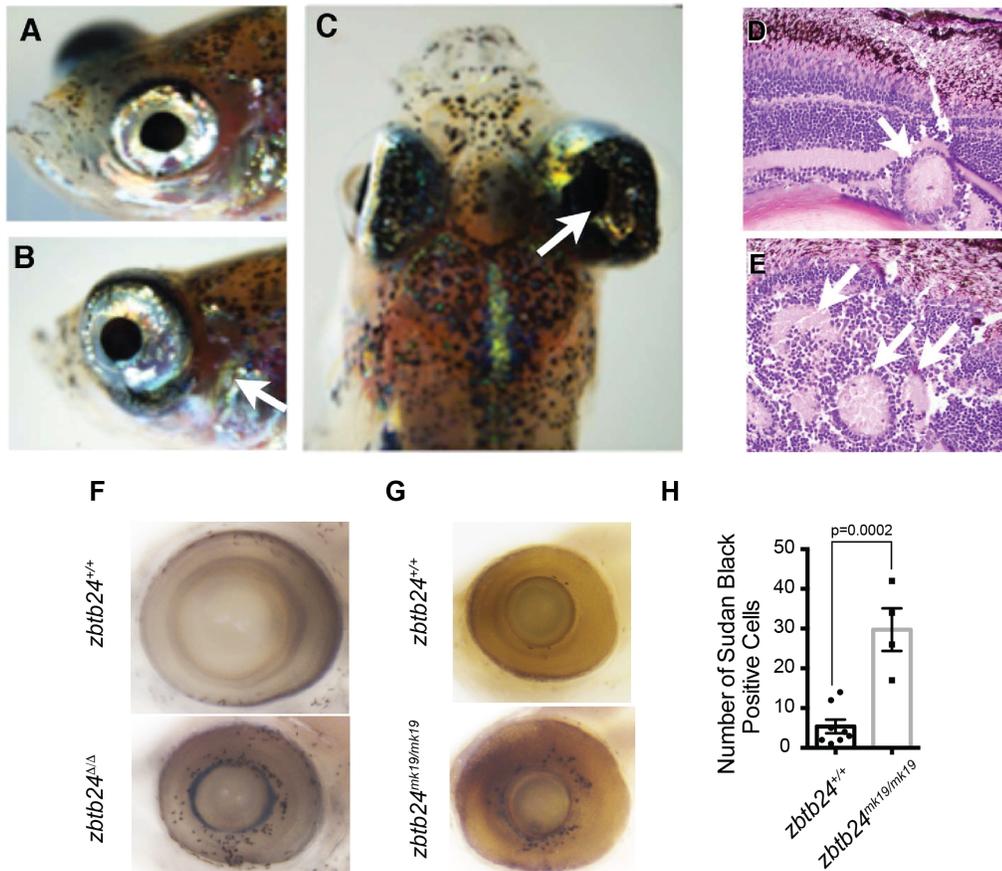


**Figure 5.2: Increased incidence of DNA damage in *zbtb24* mutants.**

(A) Immunohistochemistry of  $\gamma$ -H2AX (red) merged with DAPI (blue) on cryosection of *zbtb24*<sup>+/+</sup> vs *zbtb24* <sup>$\Delta/\Delta$</sup>  at 1.5 months (B) Quantification of number of cells with  $\gamma$ -H2AX foci in 5 representative regions in a cryosection of *zbtb24* <sup>$\Delta/\Delta$</sup>  at 1.5 months. The equivalent tissue in *zbtb24*<sup>+/+</sup> is used as a control for counting (C) Immunohistochemistry of  $\gamma$ -H2AX on three cryosections of *zbtb24*<sup>+/+</sup> vs *zbtb24* <sup>$\Delta/\Delta$</sup>  at 5 months.

yH2ax to assay for any incidence of DNA damage. Preliminary experiments revealed increased DNA damage across tissues in *zbtb24* mutants at 1.5 months (**Figure 5.2 A-B**) and 5 months (**Figure 5.2 C**). This finding provided an initial indication of genome instability in *zbtb24* mutants.

Obvious solid tumors have not been observed in *zbtb24* mutants, though systematic analysis for liquid tumors in *zbtb24* mutants remains to be performed. To further investigate whether *zbtb24* mutants are pre-disposed to cancer, I crossed them onto a homozygous mutant background for the tumor suppressor, *tp53* (Berghmans et al., 2005). In preliminary experiments, I observed development of ocular tumors starting at 2 months post fertilization in 5 out of 16 genotyped *zbtb24<sup>Δ/Δ</sup>; tp53<sup>-/-</sup>* adult zebrafish (**Figure 5.3 A-E**). None of the *tp53<sup>-/-</sup>* siblings (wild-type for *zbtb24*) raised in this cohort developed any ocular tumors in the time period of data collection (2-8 months post fertilization). Since, the *zbtb24* mutants are predisposed to chronic inflammation and given that the association between chronic inflammation and cancer is well-accepted (Mantovani et al., 2008), I tested for signs of inflammation in the eyes of *zbtb24* mutants at 3 weeks post fertilization. I picked this stage since it represented the earliest stage in *zbtb24* mutants with incidence of inflammatory phenotypes (initially observed in the gut). Sudan black staining for neutrophils revealed increased inflammation in the eyes of *zbtb24* mutants (**Figure 5.3 F-H**). The persistence of eye inflammation later in development in *zbtb24* mutants remains to be tested. Together, these results provide preliminary evidence of genome instability and cancer in *zbtb24* mutants.



**Figure 5.3: Increased incidence of eye tumors in *zbtb24*<sup>Δ/Δ</sup>; *tp53*<sup>-/-</sup> zebrafish.**

(A-C) Representative brightfield images of *zbtb24*<sup>Δ/Δ</sup>; *tp53*<sup>-/-</sup> zebrafish that develop eye tumors (white arrow) starting ~2 months. (D-E) H&E staining of a cross-section of the ocular tumors observed in *zbtb24*<sup>Δ/Δ</sup>; *tp53*<sup>-/-</sup> zebrafish (F) Sudan black staining in the eye of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> at 3 weeks. (G) Sudan black staining in the eye of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> at 3 weeks. (H) Quantification of Sudan black positive cells in the eyes of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>mk19/mk19</sup> at 3 weeks.

## DISCUSSION

Preliminary analysis reveals increased incidence of DNA damage in *zbtb24*<sup>Δ/Δ</sup> and eye tumors in *zbtb24*<sup>Δ/Δ</sup>; *tp53*<sup>-/-</sup> adult zebrafish. The incidence of eye tumors correlates with eye inflammatory phenotypes observed in *zbtb24*<sup>Δ/Δ</sup>. Analysis of

Sudan black staining in juvenile zebrafish mutants revealed highest levels of neutrophil staining in the gastrointestinal tract, a highly proliferative tissue during juvenile development, and in the proliferative region of the eye. It is likely that hypomethylation of pericentromeric satellite repeats accumulates the most in rapidly proliferating cells and contributes to inflammation. If tumors observed in *zbtb24*<sup>Δ/Δ</sup>; *tp53*<sup>-/-</sup> zebrafish are indeed associated with inflammation, these fish are likely to have an increased risk for intestinal cancers as well. This is an area of future investigation.

Several studies have previously implicated a protective role for an interferon response against cancer. For instance, suppressing interferon signaling promoted melanoma development with 100% penetrance in a mouse model with oncogenic mutations in BRAF (Katlinskaya et al., 2016). In this study, I demonstrate the activation of an interferon response is a primary consequence of pericentromeric DNA hypomethylation. Given the link between interferon activation and tumor suppression, it is likely that inhibiting the interferon response in *zbtb24* mutants could further promote accumulation of genome instability and/or the onset of malignant phenotypes.

## Chapter 6 : SUMMARY AND PERSPECTIVES

### Animal model of ICF Syndrome

I have identified *zbtb24* homozygous mutant zebrafish as a faithful animal model of ICF syndrome. This viable animal model of ICF syndrome recapitulates key phenotypic hallmarks of the disease including slow growth, facial anomalies, immunoglobulin deficiencies, reduced lifespan and pericentromeric DNA hypomethylation. Using this animal model, I have identified a link between progressive pericentromeric DNA hypomethylation and progressive development of ICF syndrome growth defects. In addition, I have identified intestinal inflammation as an important yet underappreciated feature of ICF syndrome. Given that previous attempts to model ICF syndrome have resulted in perinatal or embryonic lethality (Geiman et al., 2001; Ueda et al., 2006; Wu et al., 2016), this zebrafish model provides an important new resource for understanding ICF disease etiology during juvenile and adult life stages. In particular, *zbtb24* mutant zebrafish will be useful for understanding phenotypes such as immunoglobulin deficiency and its connection to pericentromeric DNA hypomethylation (if any), which have not been observed in mouse models and are difficult to study in cell culture systems. Regular intravenous infusions of immunoglobulin are the most common treatment offered to ICF patients currently. *zbtb24* mutant zebrafish can be used as a platform for drug discovery to treat the different clinical manifestations of ICF syndrome.

## **Function of Zbtb24**

Previous studies have suggested that ZBTB24 is a transcription factor that acts to regulate DNA methylation through transcriptional control of the ICF gene CDCA7 (Thompson et al., 2018; Wu et al., 2016). Consistent with this model, I observe near complete loss of *cdca7* expression in *zbtb24* mutants. Collectively, my findings corroborate previous research that implicate ZBTB24 as a transcription factor and expand on its role in long-term maintenance of 5mC at pericentromeric satellite repeats. While, the progressive loss of 5mC from Sat1 repeats in *zbtb24* mutants indicates a role for Zbtb24 in maintenance methylation, its role in *de novo* establishment at these sites is unclear. Curiously, methylation levels in sperm from *zbtb24* mutants and wildtype animals appeared comparable. This opens avenues of future research on understanding the differences between the control of pericentromeric methylation in germ and somatic cells.

## **Pericentromeric hypomethylation, genome instability and the interferon response**

A major finding of this dissertation is that an innate immune response is an early consequence of pericentromeric hypomethylation. These results support a model in which derepression of transcripts from hypomethylated pericentromeres triggers activates the innate immune system through the Mda5/Mavs viral RNA recognition pathway. These findings identify roles for pericentromeric RNA as a trigger of autoimmunity and reveal important functions for pericentromeric methylation in suppressing the generation of these immunostimulatory transcripts. Hypomethylation of pericentromeric sequences

is observed in abnormal cell contexts including cancer and senescence. Massive increases in pericentromeric transcripts and upregulation of interferon genes have both been noted in cancer (Cheon et al., 2014; Ting et al., 2011). Our data raise the possibility that pericentromeric hypomethylation and subsequent derepression of associated RNAs represents an important but underappreciated trigger of autoimmunity in a variety of disease states.

## **MATERIALS AND METHODS**

### **Zebrafish husbandry**

Zebrafish husbandry and care were conducted in full accordance with animal care and use guidelines with approval by the Institutional Animal Care and Use Committees at Memorial Sloan Kettering Cancer Center and the University of Georgia. Zebrafish were raised under standard conditions at 28° C. Wild-type lines were of the AB background. Mutant alleles are summarized in Appendix 1.

### **TALEN and CRISPR mutagenesis**

TALEN sequences were selected using Targeter 2.0 software (Doyle et al., 2012). TAL repeat assembly was achieved using the Golden Gate assembly method, and assembled repeats were integrated into the GoldyTALEN scaffold (Bedell et al., 2012; Cermak et al., 2011). Assembled vectors served as templates for in vitro mRNA transcription using the T3 mMessage mMachine kit (Ambion) according to manufacturer's instructions. 50–100pg mRNA was injected into wild-type embryos at the one-cell stage. Injected embryos were raised to adulthood and F1 progeny were screened for germline transmission of mutations as previously described (Li et al., 2015). Primers used for detection of mutations and subsequent genotyping are included in Appendix 1.

Target selection for CRISPR/Cas9 mediated mutagenesis was performed using CHOPCHOP (Labun et al., 2016). sgRNA templates were generated either by cloning into pT7-gRNA as described by (Jao et al., 2013) or using the oligo-based approach described in (Gagnon et al., 2014) and (Burger et al., 2016).

Template oligos are listed in Appendix 3. sgRNAs were in vitro transcribed from their respective templates using T7 RNA polymerase (Promega) as per manufacturer protocol. Cas9 RNA was in vitro transcribed from the pT3TS-nls-zCas9-nls plasmid (Jao et al., 2013) using the T3 mMessage mMachine Kit (Ambion). For mutagenesis, 200-400ng of sgRNA and ~500ng of Cas9 mRNA were co-injected into wild-type embryos at the one-cell stage. Injected embryos were raised to adulthood, and F1 progeny were screened for germline transmission of mutations as previously described (Li et al., 2015). Primers used for detection of mutations and genotyping are included in Appendix 1.

### **Zebrafish imaging and length measurements**

All bright field imaging of zebrafish larvae and adult was performed using Olympus MVX10 with CellSens Standard software. Standard-length was documented using ImageJ as defined in (Parichy et al., 2009). Photoshop (Adobe) adjustments to brightness and contrast were equally applied to all images of whole zebrafish in order to improve visualization.

### **FACS Analysis of Whole Kidney Marrow**

Adult zebrafish at 6 months were sacrificed with a combination of tricaine (Sigma-Aldrich, CAS number 886-86-2) and rapid chilling. Whole kidneys were dissected using forceps and placed in 0.9× PBS/5% FCS. Manual disaggregation using a P1000 pipette resulted in single cell suspensions. Cells were filtered over a 40 µm nylon mesh filter, and resuspended in PBS/FCS to give a final concentration of 100,000 cells/µl. FACS sorting of single cells were

analyzed for forward/side scatter profiles. FACS data were analyzed using FloJo software.

### **Morpholino oligonucleotide (MO) knockdown**

One cell wild-type zebrafish embryos were injected with 3.5ng of *zbtb24* ATG blocking MO (5'- gagtaaggatcatcATGTCTGCCCTT-3') or Splice site blocking MO (obtained from the Zon Lab, Harvard Univ.).

### **Whole Mount in situ Hybridization (WISH)**

WISH was performed as described previously (Thisse and Thisse, 2008).

### **DNA Methylation Analysis**

For Southern blot analysis, 1 µg of purified total genomic DNA was digested with the indicated methylation sensitive restriction enzyme, fractionated by electrophoresis through a 0.9% agarose gel and transferred to nylon membrane. Probes were PCR amplified using primers in Appendix 2 and radiolabeled with <sup>32</sup>P-dCTP using Rediprime™ II Random Prime Labelling System (Amersham) according to manufacturer protocol. Hybridization signals were imaged and analyzed using a Typhoon phosphorimager (GE Life Sciences). Signal intensities were measured using ImageJ. Methylation changes at Sat1 was quantified as a ratio of the intensity of the unmethylated / methylated blot regions as indicated in the respective blot. HypCH4IV was selected for Sat1 methylation analysis over the more traditional Msp1/HpaII isoschizomer pair because Sat1 sequences lack the CCGG sites that are recognized by these enzymes.

## **Enhanced Reduced Representation Bisulfite Sequencing (ERRBS)**

50 ng of high quality genomic DNA was prepared from fin tissue from adult zebrafish at 24 wpf as previously described (Garrett-Bakelman et al., 2015). DNA was digested with Msp1. Bisulphite conversion rates (calculated using non-CpG methylation conversion rates) ranged from 99.6 to 99.7% for all samples (Figure 2-figure supplement 2C). Amplified libraries were sequenced on the HiSeq2000 platform for 50 cycles single end read runs. ERRBS data were filtered for sequence adapters, limited to the first 29 bp of the read (Boyle et al., 2012), and mapped to the zebrafish genome (danRer7) using BSMAP (v 2.90) (Xi and Li, 2009). Methylation scores were calculated as the number of unconverted reads divided by the number of total reads at each CpG site. DMRs with at least a 0.2 change in methylation were determined using DSS (delta=0.2, p.threshold=0.01) (Park and Wu, 2016). Sat1 sequences are deficient in Msp1 sites, and are therefore not included in ERRBS data.

## **RNA Expression Analysis**

For qRT-PCR, total RNA was isolated using Trizol (Invitrogen) and precipitated with isopropanol. RNA used for assaying expression of repeat sequences subsequently was treated with DNase using TURBO DNA-free™ Kit (Ambion) prior to analyses. RNA was converted to cDNA using GoScript™ Reverse Transcriptase Kit (Promega) and Real Time PCR was performed using an Applied Biosystems 7500 PCR Machine. Analysis was performed using the  $2^{-\Delta\Delta Ct}$  method, with relative mRNA levels of all transcripts normalized to  $\beta$ -actin1. All primer sequences are listed in Appendix 2.

For Northern blot analysis, total RNA was extracted with using Trizol (Invitrogen). 2 µg of RNA was subjected to electrophoresis on 1% agarose gel and transferred to Amersham Hybond-N+ membrane (GE Healthcare). The membrane was probed with <sup>32</sup>P-dCTP radiolabeled Sat1 DNA probe at 42°C. Hybridization signals were imaged and analyzed using a Typhoon phosphorimager (GE Life Sciences).

TAG-aided sense/antisense transcript detection (TASA-TD) strand-specific PCR was performed as described by (Henke et al., 2015). Oligos used are listed in Supplemental File 3.

### **Transcriptome sequencing**

After RiboGreen quantification and quality control by Agilent BioAnalyzer, 500 ng of total RNA underwent polyA selection and TruSeq library preparation according to instructions provided by Illumina (TruSeq Stranded mRNA LT Kit), with 8 cycles of PCR. Samples were barcoded and run on a HiSeq 2500 High Output in a 50bp/50bp paired end run, using the TruSeq SBS v4 Kit (Illumina). An average of 45.3 million paired reads was generated per sample. The percent of mRNA bases averaged 62.8%.

For single-mutant RNA-seq analysis presented in Fig 3, reads were mapped to the Zebrafish genome (danRer7) using the rnaStar aligner v2.5.0a (Dobin et al., 2013). I used the two-pass mapping method outlined in (Engstrom et al., 2013). The first mapping pass used a list of known annotated junctions from Ensemble. Novel junctions found in the first pass were then added to the known junctions

and a second mapping pass was done (on the second pass the RemoveNoncanonical flag was used). Expression counts (counts per million, cpm) were computed from the mapped reads using HTSeq v0.5.3 (Anders et al., 2015) and Ensemble D.rerio v79 gene annotations. Normalization and differential expression was performed using DESeq (Anders and Huber, 2010).

For RNA-seq analysis presented in Fig 6, raw RNA-seq FASTQ reads were trimmed for adapters and preprocessed to remove low-quality reads using Trimmomatic v0.33 (arguments: LEADING:3 TRAILING:3 MINLEN:36) (Bolger et al., 2014) prior to mapping to the *Danio rerio* GRCz10 reference genome assembly. Reads were mapped using TopHat v2.1.1 (Kim et al., 2013) supplied with a reference General Features File (GFF) to the *Danio rerio* GRCz10 reference genome assembly, and with the following arguments: -i 10 -l 5000 --library-type fr-firststrand. Gene expression was estimated using Cuffquant (a tool from Cufflinks v2.2.1), with following arguments --library-type fr-firststrand. Expression level were normalized in FPKM units by Cuffnorm (a tool from Cufflinks v2.2.1), with following arguments --library-type fr-firststrand.

### **RNA synthesis and injections**

Sat1 RNA and dsRed RNAs were *in vitro* transcribed using Riboprobe® *in vitro* transcription systems (Promega). Oligos to amplify the DNA template for *in vitro* transcription are included in Appendix 3. Sense and anti-sense transcripts were transcribed *in vitro* using the T3 and T7 RNA polymerases respectively. RNA was purified illustra MicroSpin G-50 Columns (GE Healthcare) and 50 ng of

sense and antisense RNA was co-injected into zebrafish embryos at the 1-cell stage.

### **Histology**

For Hematoxylin and Eosin (H&E) staining, adult zebrafish were fixed in 10% Neutral Buffered Formalin for 48 hours. Zebrafish were then decalcified in 0.5 M EDTA for 24h. After decalcification, fish were incubated overnight in 70% Ethanol before embedding in paraffin blocks. Sections were stained with H&E according to standard procedures.

### **Sperm Count**

Adult zebrafish at 8 months were sacrificed with a combination of tricaine and rapid chilling. Whole testis was dissected using forceps and crushed in 100 ul of PBS. For determining sperm-count, sperm samples were diluted 1:20 for each fish. 10ul of the diluted sample was then loaded onto a hemocytometer (Bright-Line, Hauser Scientific) for counting. The volume over the central counting area is  $0.1 \text{ mm}^3$  or 0.1 microliter. Average number of sperm counted over the central counting area was multiplied by 10000 to obtain the number of sperm/ml of the diluted sample. The obtained value was multiplied by the dilution factor to obtain the final sperm count.

### **Chromatin Immunoprecipitation (ChIP)**

ChIP was performed as described in Lindeman et al. 2009 with modifications. Briefly, zebrafish juveniles at 1 month were euthanized using tricaine. Chromatin was prepared from euthanized fish by lysing flash frozen samples

using an automated pulverizer (Covaris) and crosslinking using 1% Formaldehyde for 5 mins. Chromatin shearing was performed using a Covaris S220 sonicator using the following conditions: 1 ml tubes with total chromatin from each fish in buffer containing 1% SDS were sonicated using peak intensity power of 140, duty factor of 5.0 and 200 cycles per burst, for 14 minutes for *zbtb24*<sup>+/+</sup> and 6 minutes for *zbtb24*<sup>Δ/Δ</sup>. Shearing was monitored using 1% agarose gel. To provide standardized input for each ChIP experiment, chromatin was diluted to A260=0.2. For each ChIP, 2 μg antibody per 10 μl Dynabeads and 100 μl chromatin was incubated overnight at 4 °C. Following antibodies were used in this study: anti-H3K9me3 antibody (abcam ab8898), H3K27me3 (Millipore 07-449) and IgG control (abcam ab15008). After elution, ChIP DNA and input controls were purified using QIAquick PCR purification kit (Qiagen). Eluted DNA was analyzed by qPCR using primers targeting Sat1.

### **5-aza-dC Treatment**

Zebrafish embryos were treated with 5-aza-dC (Sigma-Aldrich) to the final concentration of 25uM or 50uM within the first 2 hours post fertilization, when zebrafish are sensitive to 5-aza-dC treatments as described in Martin et al., 1999. At 24hpf, total RNA was collected for expression analysis. At 24hpf, genomic DNA was also collected and digested with methylation sensitive enzyme, HpaII, to test for global DNA hypomethylation.

### **Immunofluorescence**

Whole zebrafish at the desired stage of analysis were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) overnight at 4°C room

temperature, rehydrated with PBS and and cryoprotected by overnight saturation in 30% sucrose. 25 µm sections were cut using a cryostat and collected on SuperFrost slides (Thermo Fisher). anti-H3K9me3 antibody (abcam ab8898), H3K27me3 (Millipore 07-449), anti-H3K4me3 (Millipore 07-473) and anti-γ-H2AX (Genetex, GTX127342) were diluted 1:200 in PBS, 10% sheep serum, 1% DMSO, 0.1% Tween-20. Alexa-596 goat anti-rabbit secondary antibodies (abcam ab150088) were diluted 1:1000. Images were acquired using an AxioCam camera on a Zeiss microscope.

### **Microgavage**

Microgavage of zebrafish juveniles at 3 weeks post fertilization was performed as described in (Cocchiaro and Rawls, 2013)

### **Sudan Black staining**

Juveniles at appropriate stage (3 weeks), rinsed in PBS and incubated in 0.03% Sudan Black (Sigma-Aldrich) in 70% ethanol and 0.1% phenol for 30mins-3 hrs minutes (depending on the uptake of the stain). After staining, the fixed samples are washed extensively in 70% ethanol in water, then progressively rehydrated to PBS and 0.1% Tween 20 (PBT). After rehydrating, samples are washed in 1% KOH and 1% H<sub>2</sub>O<sub>2</sub> to remove pigmentation for imaging.

### **Statistical Analysis**

The Student unpaired 2-tailed t-test was used for statistical analysis unless specified otherwise. Statistical analysis was performed using GraphPad PRISM software.

## REFERENCES

- Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol* 19, 229-244.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol* 11, R106.
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166-169.
- Anderson, R.M., Bosch, J.A., Goll, M.G., Hesselson, D., Dong, P.D., Shin, D., Chi, N.C., Shin, C.H., Schlegel, A., Halpern, M., *et al.* (2009). Loss of Dnmt1 catalytic activity reveals multiple roles for DNA methylation during pancreas development and regeneration. *Dev Biol* 334, 213-223.
- Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Herault, Y., Guillou, F., and Bourc'his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* 354, 909-912.
- Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krug, R.G., 2nd, Tan, W., Penheiter, S.G., Ma, A.C., Leung, A.Y., *et al.* (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature* 491, 114-118.
- Beisel, C., and Paro, R. (2011). Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet* 12, 123-135.
- Berghmans, S., Murphey, R.D., Wienholds, E., Neubergh, D., Kutok, J.L., Fletcher, C.D., Morris, J.P., Liu, T.X., Schulte-Merker, S., Kanki, J.P., *et al.* (2005). tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. *Proc Natl Acad Sci U S A* 102, 407-412.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120.
- Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760-1764.
- Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science* 294, 2536-2539.
- Boyle, P., Clement, K., Gu, H., Smith, Z.D., Ziller, M., Fostel, J.L., Holmes, L., Meldrim, J., Kelley, F., Gnirke, A., *et al.* (2012). Gel-free multiplexed reduced

representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol* 13, R92.

Briolat, V., Jouneau, L., Carvalho, R., Palha, N., Langevin, C., Herbomel, P., Schwartz, O., Spaink, H.P., Levraud, J.P., and Boudinot, P. (2014). Contrasted innate responses to two viruses in zebrafish: insights into the ancestral repertoire of vertebrate IFN-stimulated genes. *J Immunol* 192, 4328-4341.

Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L.M., Catena, R., Jinek, M., *et al.* (2016). Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* 143, 2025-2037.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 39, e82.

Chedin, F., Lieber, M.R., and Hsieh, C.L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A* 99, 16916-16921.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. *Molecular and Cellular Biology* 23, 5594-5605.

Cheon, H., Borden, E.C., and Stark, G.R. (2014). Interferons and their stimulated genes in the tumor microenvironment. *Semin Oncol* 41, 156-173.

Chernyavskaya, Y., Mudbhary, R., Zhang, C., Tokarz, D., Jacob, V., Gopinath, S., Sun, X., Wang, S., Magnani, E., Madakashira, B.P., *et al.* (2017). Loss of DNA methylation in zebrafish embryos activates retrotransposons to trigger antiviral signaling. *Development* 144, 2925-2939.

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A., *et al.* (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell* 162, 974-986.

Choo, K.H.A. (1997). *The centromere* (Oxford ; New York: Oxford University Press).

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell* 112, 407-421.

Cocchiaro, J.L., and Rawls, J.F. (2013). Microgavage of zebrafish larvae. *J Vis Exp*, e4434.

Conrad, M.A., Dawany, N., Sullivan, K.E., Devoto, M., and Kelsen, J.R. (2017). Novel ZBTB24 Mutation Associated with Immunodeficiency, Centromere Instability, and Facial Anomalies Type-2 Syndrome Identified in a Patient with Very Early Onset Inflammatory Bowel Disease. *Inflamm Bowel Dis* 23, 2252-2255.

Corley-Smith, G.E., Lim, C.J., and Brandhorst, B.P. (1996). Production of androgenetic zebrafish (*Danio rerio*). *Genetics* 142, 1265-1276.

Crowl, J.T., Gray, E.E., Pestal, K., Volkman, H.E., and Stetson, D.B. (2017). Intracellular Nucleic Acid Detection in Autoimmunity. *Annu Rev Immunol* 35, 313-336.

Cruickshanks, H.A., McBryan, T., Nelson, D.M., Vanderkraats, N.D., Shah, P.P., van Tuyn, J., Singh Rai, T., Brock, C., Donahue, G., Dunican, D.S., *et al.* (2013). Senescent cells harbour features of the cancer epigenome. *Nat Cell Biol* 15, 1495-1506.

de Greef, J.C., Wang, J., Balog, J., den Dunnen, J.T., Frants, R.R., Straasheijm, K.R., Aytekin, C., van der Burg, M., Duprez, L., Ferster, A., *et al.* (2011). Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Am J Hum Genet* 88, 796-804.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev* 25, 1010-1022.

Dejardin, J. (2015). Switching between Epigenetic States at Pericentromeric Heterochromatin. *Trends Genet* 31, 661-672.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21.

Doyle, E.L., Booher, N.J., Standage, D.S., Voytas, D.F., Brendel, V.P., Vandyk, J.K., and Bogdanove, A.J. (2012). TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res* 40, W117-122.

Ehrlich, M. (2003). The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clinical Immunology* 109, 17-28.

Ehrlich, M. (2009). DNA hypomethylation in cancer cells. *Epigenomics* 1, 239-259.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* 10, 2709-2721.

Ehrlich, M., Hopkins, N.E., Jiang, G., Dome, J.S., Yu, M.C., Woods, C.B., Tomlinson, G.E., Chintagumpala, M., Champagne, M., Dillerg, L., *et al.* (2003). Satellite DNA hypomethylation in karyotyped Wilms tumors. *Cancer Genet Cytogenet* 141, 97-105.

Ehrlich, M., Jackson, K., and Weemaes, C. (2006). Immunodeficiency, centromeric region instability, facial anomalies syndrome (ICF). *Orphanet J Rare Dis* 1, 2.

Ekker, M., Fritz, A., and Westerfield, M. (1992). Identification of two families of satellite-like repetitive DNA sequences from the zebrafish (*Brachydanio rerio*). *Genomics* 13, 1169-1173.

Engstrom, P.G., Steijger, T., Sipos, B., Grant, G.R., Kahles, A., Ratsch, G., Goldman, N., Hubbard, T.J., Harrow, J., Guigo, R., *et al.* (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat Methods* 10, 1185-1191.

Erukashvily, N.I., Donev, R., Waisertreiger, I.S., and Podgornaya, O.I. (2007). Human chromosome 1 satellite 3 DNA is decondensed, demethylated and transcribed in senescent cells and in A431 epithelial carcinoma cells. *Cytogenet Genome Res* 118, 42-54.

Esteller, M. (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* 21, 5427-5440.

Fanelli, M., Caprodossi, S., Ricci-Vitiani, L., Porcellini, A., Tomassoni-Ardori, F., Amatori, S., Andreoni, F., Magnani, M., De Maria, R., Santoni, A., *et al.* (2008). Loss of pericentromeric DNA methylation pattern in human glioblastoma is associated with altered DNA methyltransferases expression and involves the stem cell compartment. *Oncogene* 27, 358-365.

Feng, S., Cokus, S.J., Zhang, X., Chen, P.Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., *et al.* (2010). Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A* 107, 8689-8694.

Gagnon, J.A., Valen, E., Thyme, S.B., Huang, P., Akhmetova, L., Pauli, A., Montague, T.G., Zimmerman, S., Richter, C., and Schier, A.F. (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* 9, e98186.

Garrett-Bakelman, F.E., Sheridan, C.K., Kacmarczyk, T.J., Ishii, J., Betel, D., Alonso, A., Mason, C.E., Figueroa, M.E., and Melnick, A.M. (2015). Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. *J Vis Exp*, e52246.

Garrido-Ramos, M.A. (2017). Satellite DNA: An Evolving Topic. *Genes (Basel)* 8.

Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003). Induction of tumors in mice by genomic hypomethylation. *Science* 300, 489-492.

Geiman, T.M., Tessarollo, L., Anver, M.R., Kopp, J.B., Ward, J.M., and Muegge, K. (2001). Lsh, a SNF2 family member, is required for normal murine development. *Biochim Biophys Acta* 1526, 211-220.

Gill, R.M., Gabor, T.V., Couzens, A.L., and Scheid, M.P. (2013). The MYC-associated protein CDCA7 is phosphorylated by AKT to regulate MYC-dependent apoptosis and transformation. *Mol Cell Biol* 33, 498-513.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 74, 481-514.

Goll, M.G., and Halpern, M.E. (2011). DNA methylation in zebrafish. *Prog Mol Biol Transl Sci* 101, 193-218.

Grafodatskaya, D., Choufani, S., Ferreira, J.C., Butcher, D.T., Lou, Y., Zhao, C., Scherer, S.W., and Weksberg, R. (2010). EBV transformation and cell culturing destabilizes DNA methylation in human lymphoblastoid cell lines. *Genomics* 95, 73-83.

Hall, L.L., Byron, M., Carone, D.M., Whitfield, T.W., Pouliot, G.P., Fischer, A., Jones, P., and Lawrence, J.B. (2017). Demethylated HSATII DNA and HSATII RNA Foci Sequester PRC1 and MeCP2 into Cancer-Specific Nuclear Bodies. *Cell Rep* 18, 2943-2956.

Hartlova, A., Erttmann, S.F., Raffi, F.A., Schmalz, A.M., Resch, U., Anugula, S., Lienenklaus, S., Nilsson, L.M., Kroger, A., Nilsson, J.A., *et al.* (2015). DNA damage primes the type I interferon system via the cytosolic DNA sensor STING to promote anti-microbial innate immunity. *Immunity* 42, 332-343.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293, 1098-1102.

Henke, C., Strissel, P.L., Schubert, M.T., Mitchell, M., Stolt, C.C., Faschingbauer, F., Beckmann, M.W., and Strick, R. (2015). Selective expression of sense and antisense transcripts of the sushi-ichi-related

retrotransposon--derived family during mouse placentogenesis. *Retrovirology* 12, 9.

Houwing, S., Berezikov, E., and Ketting, R.F. (2008). Zili is required for germ cell differentiation and meiosis in zebrafish. *EMBO J* 27, 2702-2711.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., *et al.* (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498-503.

Ishikawa, H., Ma, Z., and Barber, G.N. (2009). STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* 461, 788-792.

Jagannathan, M., Cummings, R., and Yamashita, Y.M. (2018). A conserved function for pericentromeric satellite DNA. *Elife* 7.

Jao, L.E., Wente, S.R., and Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A* 110, 13904-13909.

Jenness, C., Giunta, S., Muller, M.M., Kimura, H., Muir, T.W., and Funabiki, H. (2018). HELLS and CDCA7 comprise a bipartite nucleosome remodeling complex defective in ICF syndrome. *Proc Natl Acad Sci U S A* 115, E876-E885.

Jiang, L., Zhang, J., Wang, J.J., Wang, L., Zhang, L., Li, G., Yang, X., Ma, X., Sun, X., Cai, J., *et al.* (2013). Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* 153, 773-784.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13, 484-492.

Jones, P.A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068-1070.

Joseph, A., Mitchell, A.R., and Miller, O.J. (1989). The organization of the mouse satellite DNA at centromeres. *Exp Cell Res* 183, 494-500.

Katlinskaya, Y.V., Katlinski, K.V., Yu, Q., Ortiz, A., Beiting, D.P., Brice, A., Davar, D., Sanders, C., Kirkwood, J.M., Rui, H., *et al.* (2016). Suppression of Type I Interferon Signaling Overcomes Oncogene-Induced Senescence and Mediates Melanoma Development and Progression. *Cell Rep* 15, 171-180.

Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol* 11, 373-384.

Kawai, T., Takahashi, K., Sato, S., Coban, C., Kumar, H., Kato, H., Ishii, K.J., Takeuchi, O., and Akira, S. (2005). IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat Immunol* 6, 981-988.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14, R36.

Labun, K., Montague, T.G., Gagnon, J.A., Thyme, S.B., and Valen, E. (2016). CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Research* 44, W272-W276.

Laird, P.W. (2010). Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet* 11, 191-203.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. *Current Biology* 13, 1192-1200.

Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71, 865-873.

Leonova, K.I., Brodsky, L., Lipchick, B., Pal, M., Novototskaya, L., Chenchik, A.A., Sen, G.C., Komarova, E.A., and Gudkov, A.V. (2013). p53 cooperates with DNA methylation and a suicidal interferon response to maintain epigenetic silencing of repeats and noncoding RNAs. *Proc Natl Acad Sci U S A* 110, E89-98.

Li, C., Lan, Y., Schwartz-Orbach, L., Korol, E., Tahiliani, M., Evans, T., and Goll, M.G. (2015). Overlapping Requirements for Tet2 and Tet3 in Normal Development and Hematopoietic Stem Cell Emergence. *Cell Rep* 12, 1133-1143.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915-926.

Li, Y., Esain, V., Teng, L., Xu, J., Kwan, W., Frost, I.M., Yzaguirre, A.D., Cai, X., Cortes, M., Majenburg, M.W., *et al.* (2014). Inflammatory signaling regulates

embryonic hematopoietic stem and progenitor cell production. *Genes Dev* 28, 2597-2612.

Li, Y., Li, Y., Cao, X., Jin, X., and Jin, T. (2017). Pattern recognition receptors in zebrafish provide functional and evolutionary insight into innate immune signaling pathways. *Cell Mol Immunol* 14, 80-89.

Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., and Jones, P.A. (2002). Cooperativity between DNA Methyltransferases in the Maintenance Methylation of Repetitive Elements. *Molecular and Cellular Biology* 22, 480-491.

Liang, J., Yan, R., Chen, G., Feng, J., Wu, W.W., Ren, W., Zhu, C., Zhao, Y., Gao, X.M., and Wang, J. (2016). Downregulation of ZBTB24 hampers the G0/1-to S-phase cell-cycle transition via upregulating the expression of IRF-4 in human B cells. *Genes Immun* 17, 276-282.

Lopez-Flores, I., and Garrido-Ramos, M.A. (2012). The repetitive DNA content of eukaryotic genomes. *Genome Dyn* 7, 1-28.

Mankan, A.K., Schmidt, T., Chauhan, D., Goldeck, M., Honing, K., Gaidt, M., Kubarenko, A.V., Andreeva, L., Hopfner, K.P., and Hornung, V. (2014). Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. *EMBO J* 33, 2937-2946.

Mantovani, A., Allavena, P., Sica, A., and Balkwill, F. (2008). Cancer-related inflammation. *Nature* 454, 436-444.

Marjoram, L., Alvers, A., Deerhake, M.E., Bagwell, J., Mankiewicz, J., Cocchiari, J.L., Beerman, R.W., Willer, J., Sumigray, K.D., Katsanis, N., *et al.* (2015). Epigenetic control of intestinal barrier function and inflammation in zebrafish. *Proc Natl Acad Sci U S A* 112, 2770-2775.

McCurley, A.T., and Callard, G.V. (2008). Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. *BMC Mol Biol* 9, 102.

Medzhitov, R. (2007). Recognition of microorganisms and activation of the immune response. *Nature* 449, 819-826.

Medzhitov, R., Preston-Hurlburt, P., Kopp, E., Stadlen, A., Chen, C., Ghosh, S., and Janeway, C.A., Jr. (1998). MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. *Mol Cell* 2, 253-258.

Mudbhary, R., and Sadler, K.C. (2011). Epigenetics, development, and cancer: zebrafish make their mark. *Birth Defects Res C Embryo Today* 93, 194-203.

Myant, K., and Stancheva, I. (2008). LSH cooperates with DNA methyltransferases to repress transcription. *Mol Cell Biol* 28, 215-226.

Nakagawa, T., Kanai, Y., Ushijima, S., Kitamura, T., Kakizoe, T., and Hirohashi, S. (2005). DNA hypomethylation on pericentromeric satellite regions significantly correlates with loss of heterozygosity on chromosome 9 in urothelial carcinomas. *J Urol* 173, 243-246.

Narayan, A., Ji, W., Zhang, X.Y., Marrogi, A., Graff, J.R., Baylin, S.B., and Ehrlich, M. (1998). Hypomethylation of pericentromeric DNA in breast adenocarcinomas. *Int J Cancer* 77, 833-838.

Nitta, H., Unoki, M., Ichiyanagi, K., Kosho, T., Shigemura, T., Takahashi, H., Velasco, G., Francastel, C., Picard, C., Kubota, T., *et al.* (2013). Three novel ZBTB24 mutations identified in Japanese and Cape Verdean type 2 ICF syndrome patients. *J Hum Genet* 58, 455-460.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247-257.

Page, D.M., Wittamer, V., Bertrand, J.Y., Lewis, K.L., Pratt, D.N., Delgado, N., Schale, S.E., McGue, C., Jacobsen, B.H., Doty, A., *et al.* (2013). An evolutionarily conserved program of B-cell development and activation in zebrafish. *Blood* 122, e1-11.

Parichy, D.M., Elizondo, M.R., Mills, M.G., Gordon, T.N., and Engeszer, R.E. (2009). Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Dev Dyn* 238, 2975-3015.

Park, Y., and Wu, H. (2016). Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics* 32, 1446-1453.

Phillips, R.B., and Reed, K.M. (2000). Localization of repetitive DNAs to zebrafish (*Danio rerio*) chromosomes by fluorescence in situ hybridization (FISH). *Chromosome Res* 8, 27-35.

Plohl, M., Mestrovic, N., and Mravinac, B. (2014). Centromere identity from the DNA point of view. *Chromosoma* 123, 313-325.

Potok, M.E., Nix, D.A., Parnell, T.J., and Cairns, B.R. (2013). Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* 153, 759-772.

Qu, G., Dubeau, L., Narayan, A., Yu, M.C., and Ehrlich, M. (1999a). Satellite DNA hypomethylation vs. overall genomic hypomethylation in ovarian epithelial tumors of different malignant potential. *Mutat Res* 423, 91-101.

- Qu, G.-z., Grundy, P.E., Narayan, A., and Ehrlich, M. (1999b). Frequent Hypomethylation in Wilms Tumors of Pericentromeric DNA in Chromosomes 1 and 16. *Cancer Genetics and Cytogenetics* 109, 34-39.
- Rai, K., Jafri, I.F., Chidester, S., James, S.R., Karpf, A.R., Cairns, B.R., and Jones, D.A. (2010). Dnmt3 and G9a cooperate for tissue-specific development in zebrafish. *J Biol Chem* 285, 4110-4121.
- Rai, K., Nadauld, L.D., Chidester, S., Manos, E.J., James, S.R., Karpf, A.R., Cairns, B.R., and Jones, D.A. (2006). Zebra fish Dnmt1 and Suv39h1 regulate organ-specific terminal differentiation during development. *Mol Cell Biol* 26, 7077-7085.
- Robertson, K.D. (2005). DNA methylation and human disease. *Nat Rev Genet* 6, 597-610.
- Roulois, D., Loo Yau, H., Singhanian, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J., *et al.* (2015). DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell* 162, 961-973.
- Schneider, W.M., Chevillotte, M.D., and Rice, C.M. (2014). Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol* 32, 513-545.
- Schuetz, C., Barbi, G., Barth, T.F., Hoenig, M., Schulz, A., Moeller, P., Smeets, D., de Greef, J.C., van der Maarel, S.M., Vogel, W., *et al.* (2007). ICF syndrome: high variability of the chromosomal phenotype and association with classical Hodgkin lymphoma. *Am J Med Genet A* 143A, 2052-2057.
- Seth, R.B., Sun, L., Ea, C.K., and Chen, Z.J. (2005). Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* 122, 669-682.
- Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., *et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450, 908-912.
- Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R., and Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479, 74-79.
- Siggs, O.M., and Beutler, B. (2012). The BTB-ZF transcription factors. *Cell Cycle* 11, 3358-3369.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat Rev Genet* 14, 204-220.

Sun, L., Wu, J., Du, F., Chen, X., and Chen, Z.J. (2013). Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* 339, 786-791.

Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 9, 465-476.

Suzuki, T., Fujii, M., and Ayusawa, D. (2002). Demethylation of classical satellite 2 and 3 DNA with chromosomal instability in senescent human fibroblasts. *Exp Gerontol* 37, 1005-1014.

Tanne, A., Muniz, L.R., Puzio-Kuter, A., Leonova, K.I., Gudkov, A.V., Ting, D.T., Monasson, R., Cocco, S., Levine, A.J., Bhardwaj, N., *et al.* (2015). Distinguishing the immunostimulatory properties of noncoding RNAs expressed in cancer cells. *Proc Natl Acad Sci U S A* 112, 15154-15159.

Tao, Y., Xi, S., Shan, J., Maunakea, A., Che, A., Briones, V., Lee, E.Y., Geiman, T., Huang, J., Stephens, R., *et al.* (2011). Lsh, chromatin remodeling family member, modulates genome-wide cytosine methylation patterns at nonrepeat sequences. *Proc Natl Acad Sci U S A* 108, 5626-5631.

Tarallo, V., Hirano, Y., Gelfand, B.D., Dridi, S., Kerur, N., Kim, Y., Cho, W.G., Kaneko, H., Fowler, B.J., Bogdanovich, S., *et al.* (2012). DICER1 loss and Alu RNA induce age-related macular degeneration via the NLRP3 inflammasome and MyD88. *Cell* 149, 847-859.

Tazi, J., and Bird, A. (1990). Alternative chromatin structure at CpG islands. *Cell* 60, 909-920.

Thijssen, P.E., Ito, Y., Grillo, G., Wang, J., Velasco, G., Nitta, H., Unoki, M., Yoshihara, M., Suyama, M., Sun, Y., *et al.* (2015). Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome. *Nat Commun* 6, 7870.

Thisse, C., and Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc* 3, 59-69.

Thompson, J.J., Kaur, R., Sosa, C.P., Lee, J.H., Kashiwagi, K., Zhou, D., and Robertson, K.D. (2018). ZBTB24 is a transcriptional regulator that coordinates with DNMT3B to control DNA methylation. *Nucleic Acids Res.*

Ting, D.T., Lipson, D., Paul, S., Brannigan, B.W., Akhavanfard, S., Coffman, E.J., Contino, G., Deshpande, V., Iafrate, A.J., Letovsky, S., *et al.* (2011).

Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* 331, 593-596.

Trede, N.S., David M. Langenau, David Traver, A. Thomas Look, and Zon, a.L.I. (2004). The Use of Zebrafish to Understand Immunity. *Immunity* 20, 367-379.

Tsien, F., Fiala, E.S., Youn, B., Long, T.I., Laird, P.W., Weissbecker, K., and Ehrlich, M. (2002). Prolonged culture of normal chorionic villus cells yields ICF syndrome-like chromatin decondensation and rearrangements. *Cytogenet Genome Res* 98, 13-21.

Tsuda, H., Takarabe, T., Kanai, Y., Fukutomi, T., and Hirohashi, S. (2002). Correlation of DNA hypomethylation at pericentromeric heterochromatin regions of chromosomes 16 and 1 with histological features and chromosomal abnormalities of human breast carcinomas. *Am J Pathol* 161, 859-866.

Ueda, Y., Okano, M., Williams, C., Chen, T., Georgopoulos, K., and Li, E. (2006). Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. *Development* 133, 1183-1192.

Unoki, M., Funabiki, H., Velasco, G., Francastel, C., and Sasaki, H. (2018). CDCA7 and HELLS mutations undermine nonhomologous end joining in centromeric instability syndrome. *J Clin Invest*.

van der Vaart, M., van Soest, J.J., Spaik, H.P., and Meijer, A.H. (2013). Functional analysis of a zebrafish myd88 mutant identifies key transcriptional components of the innate immune system. *Dis Model Mech* 6, 841-854.

Velasco, G., Grillo, G., Touleimat, N., Ferry, L., Ivkovic, I., Ribierre, F., Deleuze, J.F., Chantalat, S., Picard, C., and Francastel, C. (2018). Comparative methylome analysis of ICF patients identifies heterochromatin loci that require ZBTB24, CDCA7 and HELLS for their methylated state. *Hum Mol Genet*.

von Bernuth, H., Ravindran, E., Du, H., Frohler, S., Strehl, K., Kramer, N., Issa-Jahns, L., Amulic, B., Ninnemann, O., Xiao, M.S., *et al.* (2014). Combined immunodeficiency develops with age in Immunodeficiency-centromeric instability-facial anomalies syndrome 2 (ICF2). *Orphanet J Rare Dis* 9, 116.

Vourc'h, C., and Biamonti, G. (2011). Transcription of Satellite DNAs in Mammals. *Prog Mol Subcell Biol* 51, 95-118.

Ward, W.S., and Coffey, D.S. (1991). DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biol Reprod* 44, 569-574.

Wen, Z.K., Xu, W., Xu, L., Cao, Q.H., Wang, Y., Chu, Y.W., and Xiong, S.D. (2007). DNA hypomethylation is crucial for apoptotic DNA to induce systemic

lupus erythematosus-like autoimmune disease in SLE-non-susceptible mice. *Rheumatology (Oxford)* 46, 1796-1803.

White, R.J., Collins, J.E., Sealy, I.M., Wali, N., Dooley, C.M., Digby, Z., Stemple, D.L., Murphy, D.N., Billis, K., Hourlier, T., *et al.* (2017). A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife* 6.

Wu, H., Thijssen, P.E., de Klerk, E., Vonk, K.K., Wang, J., den Hamer, B., Aytekin, C., van der Maarel, S.M., and Daxinger, L. (2016). Converging disease genes in ICF syndrome: ZBTB24 controls expression of CDCA7 in mammals. *Hum Mol Genet* 25, 4041-4051.

Wu, S.F., Zhang, H., and Cairns, B.R. (2011). Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Res* 21, 578-589.

Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* 10, 232.

Xiao, W., Custard, K.D., Brown, R.C., Lemmon, B.E., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (2006). DNA methylation is critical for Arabidopsis embryogenesis and seed viability. *Plant Cell* 18, 805-814.

Xie, H., Wang, M., Bonaldo Mde, F., Smith, C., Rajaram, V., Goldman, S., Tomita, T., and Soares, M.B. (2009). High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum. *Nucleic Acids Res* 37, 4331-4340.

Xiong, Z., and Laird, P.W. (1997). COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res* 25, 2532-2534.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402, 187-191.

Yoneyama, M., Kikuchi, M., Matsumoto, K., Imaizumi, T., Miyagishi, M., Taira, K., Foy, E., Loo, Y.M., Gale, M., Jr., Akira, S., *et al.* (2005). Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *J Immunol* 175, 2851-2858.

Zhu, Q., Hoong, N., Aslanian, A., Hara, T., Benner, C., Heinz, S., Miga, K.H., Ke, E., Verma, S., Soroczynski, J., *et al.* (2018). Heterochromatin-Encoded Satellite RNAs Induce Breast Cancer. *Mol Cell* 70, 842-853 e847.

Zilberman, D., and Henikoff, S. (2007). Genome-wide analysis of DNA methylation patterns. *Development* 134, 3959-3965.

Zimmerman, A.M., Moustafa, F.M., Romanowski, K.E., and Steiner, L.A. (2011). Zebrafish immunoglobulin IgD: unusual exon usage and quantitative expression profiles with IgM and IgZ/T heavy chain isotypes. *Mol Immunol* 48, 2220-2223.

## APPENDIX 1: LIST OF MUTANT ALLELES

Name of allele	Mutation in	Reference	Mutation	Genotyping Primers	Restriction Enzyme for genotyping
<i>mk22</i>	<i>zbtb24</i>	This publication	~8kb large deletion	F(P1):AGTCCTCGCTCTGCACTCAG R(P3):CTCTTGGCGGTGAAACACTT	N/A
<i>mk19</i>	<i>zbtb24</i>	This publication	8bp deletion	F(P1):AGTCCTCGCTCTGCACTCAG R(P2):TCTCGTCCACCAACACGAC	Fnu4HI
<i>mk28</i>	<i>mavs</i>	This publication	4bp deletion	F:ACAGCAGGTGGAGCAAGTTT R:TGAGGTGGAGATGGGAGATT	Fnu4HI
<i>mk29</i>	<i>mda5</i>	This publication	7bp deletion	F:CTGCTCTGGAGGAGAAGAACAT R:TCCAATATCATGCAGCCATAAG	BsaJI
<i>mk30</i>	<i>tmem173 (STING)</i>	This publication	10bp deletion	F:TTTCTGTGTGGCTCTGTCAAGT R:AGCGATAATTCCAGCTCTTTCA	BssHII
<i>hu3568</i>	<i>myd88</i>	van der Vaart et al., 2013	Point mutation	F:GAGGCGATTCCAGTAACAGC R:GAAGCGAACAAAGAAAGCAA	MseI
<i>s904</i>	<i>dnmt1</i>	Anderson et al., 2009	1bp frame-shift	Genotyped based on GFP labelling as described in Goll et al., 2009.	N/A

## APPENDIX 2: LIST OF PRIMERS

### qPCR

Name	Sequence
q_zbtb24_F	TCTGCACTCAGCAACACACA
q_zbtb24_R	CGGTGAACAATGCTGAGAAA
q_lgZ_F	AAAGCAACGATACCAAAGTG (Page et al., 2013)
q_lgZ_R	AACAGCTTGCAAGACAATTC (Page et al., 2013)
q_lgD_F	GACACATTAGCCCATCAGCA (Page et al., 2013)
q_lgD_R	CTGGAGAGCAGCAAAGGAT (Page et al., 2013)
q_lgM_F	GAAGCCTCCAATTCTGTTGG (Zimmerman et al., 2011)
q_lgM_R	CCGGGCTAAACACATGAAG (Zimmerman et al., 2011)
q_stat1b_F	GCTGCTCCAATGAAGGTTCC (Li et al., 2014)
q_stat1b_R	GCAACGGGTCTTGCAACAG (Li et al., 2014)
q_irf7_F	AAAGTGGGCAGTACGAAGGT
q_irf7_R	TCCATTTTGCTTTGTCGTTAG
q_irf1b_F	CGGATGAAGAGTCTGTCTCA (Briolat et al., 2014)
q_irf1b_R	TGCCGAGTGAGCTCTAAGAT (Briolat et al., 2014)
q_mxa_F	CTGCTTTTCCCAGAACTTCG
q_mxa_R	CCCCGGTACTTGACTTCGTA
q_ifi27_F	CTACTTACAGCCGCGTAGC
q_ifi27_R	GCAGCTGACATCATTGAGGA
q_ifi44l_F	CAAGCTGGAGCAGGAAAGTC
q_ifi44l_R	ACCAGGGGCAAAGTTTTCTT
q_lgals9_F	GGGCAGAAGAAAAACATGGA

q_ lgals9_R	CAGTGAAGCGGTGCTTGTA
q_gbp1_F	CCGCAGTGAGAAAGAAGGAC
q_gbp1_R	TGATCCGTTACATTCTCCA
q_CRP_F	GGCTCAATCCAAAAGGAACA
q_CRP_R	ATGACCTTCTCCCTCAAGCA
q_il22ra2_F	CTACCCTCAGTGGGAAACCA
q_il22ra2_R	TCAGTGGAGTTCTCGGAGGT
q_il1b_F	GACTTCGCAGCACAAAATGA
q_il1b_R	CACTTCACGCTCTTGGATGA
q_CCL20_F	ATATGGGCCGCTGAACTATG
q_CCL20_R	GGCATCTATACGGCACACCT
q_caspbl_F	CCTCGAGTCGAACTTCTGG
q_caspbl_R	ATGATCCCTCGTGGTCTCTG
q_tnfa_F	CAGGGCAATCAACAAGATGG (Marjoram et al., 2015)
q_tnfa_R	TGGTCCTGGTCATCTCTCCA (Marjoram et al., 2015)
q_bactin1_F	CGAGCAGGAGATGGGAACC (McCurley and Callard, 2008)
q_bactin1_R	CAACGGAAACGCTCATTGC (McCurley and Callard, 2008)
q_Sat1_F	GTCTCTGACTGAGTTTGCATTAC
q_Sat1_R	ACATTCTGAATTGGACGTTGA
q_ERV1-1_F	GTGTTCCGGAGAAAGTGGAA
q_ERV1-1_R	ACCCTCGTGCAGTGGTTTAG
q_ERV1-3_F	ATCACTATCCCGTGGCTGAG
q_ERV1-3_R	ATGTCCTCCACTCGCTTGAG

q_LTR-2_F	GGTGTCTGTTAGAATGCCCTTGAC (Houwing et al., 2008)
q_LTR-2_R	GGTTATACCTGTGGGTCACGTG (Houwing et al., 2008)
q_ZFERV1_F	CAAAACTGGGGTTTGGGAAGA
q_ZFERV1_R	CCCTGCTCCATTGTCTCAGT
q_BEL20_F	GTGTCACTTCCCCAAGTCGT
q_BEL20_R	GAAAGTGCCTCCAGAAGTGC
q_GypsyDR-2_F	GAAATCACCTGTGCATTTAC (Houwing et al., 2008)
q_GypsyDR-2_R	ATGCAGACATTGGGTAAAGC (Houwing et al., 2008)
q_Dirs1a_F	GGGTGCGTCACGCTTGC (Houwing et al., 2008)
q_Dirs1a_R	GTAACCTCGAACGTTCCCC (Houwing et al., 2008)
q_L1-1_F	AAATGCTTGGACATGGAAGG
q_L1-1_R	TCTGCTGCATCTTGGAACTG
q_L1-5_F	GCACAAAGGACAAATTCCTGGAC (Houwing et al., 2008)
q_L1-5_R	GTCCACGTTTAGTATTACAGTTGC (Houwing et al., 2008)
q_Ngaro_F	GGAGCGATCGAGACCTACC (Houwing et al., 2008)
q_Ngaro_R	CAATCATATCACGTGCTCCTCTCG (Houwing et al., 2008)
q_EnSpmN1_F	GATTGGCCATTGTGTTACATGC (Houwing et al., 2008)
q_EnSpmN1_R	GCTGTGACTGTCATAGGTTTACC (Houwing et al., 2008)

q_Polinton_F	CCTGACAATGTTGTCAGCCTG (Houwing et al., 2008)
q_Polinton_R	CATGAAAGCTAAGGGTATAACTCTG (Houwing et al., 2008)

### COBRA

Name	Sequence
Satbis_F	GAGTTTGTATTATTGTTTATT
Satbis_R	CACTTACAACCTTACTTTAAC

### Primers for probe synthesis for southern blotting

Name	Sequence
Sat1_F	TGTTTTAGACAACATTTTCATGCAC
Sat1_R	AGTCAGCCAGCAGAGAGGTC
DANA_F	GGCGACGCAGTGGCGCAGTGGG (Anderson et al., 2009)
DANA_R	TTTTCTTTTTGGCTTAGTCCC (Anderson et al., 2009)
SINE_HE1_F	TGGCTCAGTGGTTAGCACTG
SINE_HE1_R	TTTATCAGGGGTCGCCACAG
L1-10_F	ATGGAAGAGGAAGGCAAGGT
L1-10_R	GCAGGCCAATTAGTCTCAGG
Kolobok_F	AGTGAAGCACAGTTGAGCGA
Kolobok_R	ACACTCCTGGATCAGTCGGA
ERV1-3-LTR_F	TAAACATGTGTAGTGGAACTTACAGC
ERV1-3-LTR_R	TAAAGAGGCGCTCTCTGTGGT
ERV4-DR1_F	GTGGAAAACAGGGTCATTGG

ERV4-DR1_R	CACGAAGGCATAAATGCAAA
Gypsy21_LTR_F	TTTAAATCTTAAAATTGCCTAAAGGT
Gypsy21_LTR_R	TTGCACGCGATGACAACC
ZFERV2-LTR_F	TGTCCCATAGGGTGTGGTT
ZFERV2-LTR_R	CAGTGCATGTCCAAAATGG

### APPENDIX 3: LIST OF OLIGOS

#### Oligos for mutagenesis

Name	Sequence
mavs_cc_F (oligo based)	GAAATTAATACGACTCACTATAGGACATGTCAGG <u>AGCTGCTTGT</u> TTTAGAGCTAGAAAT
mda5_cc_F (oligo based)	GAAATTAATACGACTCACTATAGGTGATAAACAC <u>TGCGACCCGT</u> TTTAGAGCTAGAAAT
Invariant reverse oligo	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTG ATAACGGACTAGCCTTATTTAACTTGCTATTTCT AGCTCTAAAAC
STING_cc_F (cloning based)	TAGGCAGCCTGCTGCGCGCTCT
STING_cc_R (cloning based)	AAACAGAGCGCGCAGCAGGCTG

gRNA target sequence is underlined

#### Oligos for in vitro transcription of Sat1

Name	Sequence
Sat1_FT3	TTGAAATTAACCCTCACTAAAGGGAGACTGGTTTTATTA CATTCTGAATTGG
Sat1_RT7	TTGATAATACGACTCACTATAGGGAGATCCAGCCATAAA ATGCATCA
$\beta$ -actin_FT3	TTGAAATTAACCCTCACTAAAGGGAGAACTTTGAGCTCC TCCACACG
$\beta$ -actin_RT7	TTGATAATACGACTCACTATAGGGAGAACTCGGTGATGA CGTTCTCG
GFP_FT3	TTGAAATTAACCCTCACTAAAGGGAGATATATCATGGCC GACAAGCA
GFP_RT7	TTGATAATACGACTCACTATAGGGAGAGAACTCCAGCA GGACCATGT

**Oligos for TASA-TD**

<b>Name</b>	<b>Sequence</b>
b-actin sense TAG	GCACACGACGACAGACGACGACGCCACCAACGGAAACGCTC ATTGC
b-actin anti- sense TAG	GCACACGACGACAGACGACGACGCCACCGAGCAGGAGATGG GAACC
sat1 sense TAG	GCACACGACGACAGACGACGACGCACACATTCTGAATTGGA CGTTGA
sat1 antisense TAG	GCACACGACGACAGACGACGACGCACGTCTCTGACTGAGTT TGCATTAC
b-actin sense	CGAGCAGGAGATGGGAACC
b-actin antisense	CAACGGAAACGCTCATTGC
sat1 sense	GTCTCTGACTGAGTTTGCATTAC
sat1 antisense	ACATTCTGAATTGGACGTTGA
TAG primer	GCACACGACGACAGACGACGCAC

## **APPENDIX 4: *eLife* PAPER**

Reprinted from Rajshekar, S. *et al.* Pericentromeric hypomethylation elicits an interferon response in an animal model of ICF syndrome. *Elife* **7**, doi:10.7554/eLife.39658 (2018).

# Pericentromeric hypomethylation elicits an interferon response in an animal model of ICF syndrome

Srivarsha Rajshekar<sup>1,2,3</sup>, Jun Yao<sup>2</sup>, Paige K Arnold<sup>4</sup>, Sara G Payne<sup>5</sup>,  
Yinwen Zhang<sup>3</sup>, Teresa V Bowman<sup>5</sup>, Robert J Schmitz<sup>6</sup>, John R Edwards<sup>7</sup>,  
Mary Goll<sup>2,6\*</sup>

<sup>1</sup>Program in Biochemistry and Structural Biology, Cell and Developmental Biology, and Molecular Biology, Weill Cornell Graduate School of Medical Sciences, Cornell University, New York, United States; <sup>2</sup>Developmental Biology Program, Memorial Sloan Kettering Cancer Center, New York, United States; <sup>3</sup>Institute of Bioinformatics, University of Georgia, Athens, United States; <sup>4</sup>Louis V. Gerstner Jr. Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, United States; <sup>5</sup>Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, New York, United States; <sup>6</sup>Department of Genetics, University of Georgia, Georgia, United States; <sup>7</sup>Department of Medicine, Center for Pharmacogenomics, Washington University in St. Louis School of Medicine, Missouri, United States

**Abstract** Pericentromeric satellite repeats are enriched in 5-methylcytosine (5mC). Loss of 5mC at these sequences is common in cancer and is a hallmark of Immunodeficiency, Centromere and Facial abnormalities (ICF) syndrome. While the general importance of 5mC is well-established, the specific functions of 5mC at pericentromeres are less clear. To address this deficiency, we generated a viable animal model of pericentromeric hypomethylation through mutation of the ICF-gene *ZBTB24*. Deletion of zebrafish *zbtb24* caused a progressive loss of 5mC at pericentromeres and ICF-like phenotypes. Hypomethylation of these repeats triggered derepression of pericentromeric transcripts and activation of an interferon-based innate immune response. Injection of pericentromeric RNA is sufficient to elicit this response in wild-type embryos, and mutation of the MDA5-MAVS dsRNA-sensing machinery blocks the response in mutants. These findings identify activation of the innate immune system as an early consequence of pericentromeric hypomethylation, implicating derepression of pericentromeric transcripts as a trigger of autoimmunity.

**Editorial note:** This article has been through an editorial process in which the authors decide how to respond to the issues raised during peer review. The Reviewing Editor's assessment is that all the issues have been addressed (see decision letter).

DOI: <https://doi.org/10.7554/eLife.39658.001>

\*For correspondence:  
Mary.Goll@uga.edu

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 17

**Received:** 30 June 2018

**Accepted:** 04 November 2018

**Published:** 28 November 2018

**Reviewing editor:** Deborah Bourc'his, Institut Curie, France

© Copyright Rajshekar et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

In vertebrate genomes, the majority of cytosine residues within CpG dinucleotides are methylated at the 5 position of the cytosine ring (5-methylcytosine, 5mC) (*Suzuki and Bird, 2008*). 5mC is established by the de novo DNA methyltransferases of the Dnmt3 family, and is propagated by the maintenance DNA methyltransferase, Dnmt1 (*Goll and Bestor, 2005*). In mice, frogs and zebrafish, mutation or morpholino-mediated depletion of *Dnmt1* results in extensive genome-wide methylation

**eLife digest** Cells package DNA into structures called chromosomes. When cells divide, each chromosome duplicates, and a structure called a centromere initially holds the copies together. The sequences of DNA on either side of the centromeres are often highly repetitive. In backboned animals, this DNA normally also has extra chemical modifications called methyl groups attached to it. The role that these methyl groups play in this region is not known, although in other DNA regions they often stop the DNA being 'transcribed' into molecules of RNA.

The cells of people who have a rare human genetic disorder called ICF syndrome, lack the methyl groups near the centromere. The methyl groups may also be lost in old and cancerous cells.

Researchers often use 'model' animals to investigate the effects of DNA modifications. But, until now, there were no animal models that lose methyl groups from the DNA around centromeres in the same way as seen in ICF syndrome.

Rajshekar et al. have developed a new zebrafish model for ICF syndrome that loses the methyl groups around its centromeres over time. Studying the cells of these zebrafish showed that when the methyl groups are missing, the cell starts to transcribe the DNA sequences around the centromeres. The resulting RNA molecules appear to be mistaken by the cell for viral RNA. They activate immune sensors that normally detect RNA viruses, which triggers an immune response.

The new zebrafish model can now be used in further studies to help researchers to understand the key features of ICF syndrome. Future work could also investigate whether the loss of methyl groups around the centromeres plays a role in other diseases where the immune system attacks healthy tissues.

DOI: <https://doi.org/10.7554/eLife.39658.002>

loss and embryonic lethality (*Anderson et al., 2009; Lei et al., 1996; Rai et al., 2006; Stancheva and Meehan, 2000*). In these species, global methylation deficiencies are linked to a variety of adverse outcomes including deregulation of gene expression, derepression of transposons, elevated levels of DNA damage and increased genome instability during mitosis (*Smith and Meissner, 2013*). Recent studies have further linked global hypomethylation to activation of antiviral signaling pathways in zebrafish mutated for *dnmt1* and in cancer cells treated with the DNA methyltransferase inhibitor 5-azacytidine (*Chernyavskaya et al., 2017; Chiappinelli et al., 2015; Roulois et al., 2015*). While these studies reinforce the general importance of DNA methylation in vertebrate development and tissue homeostasis, the extensive genome-wide loss of methylation in these models makes it difficult to assign significance to methylation deficiencies at any particular subclass of sequence.

The pericentromeric satellite sequences that juxtapose chromosome centromeres represent an essential structural component of chromosomes and a significant source of 5mC in vertebrate genomes. These highly repetitive sequences appear particularly susceptible to methylation loss in cancer and senescent cells, although the consequences of this hypomethylation are not well understood (*Erukashvily et al., 2007; Fanelli et al., 2008; Nakagawa et al., 2005; Narayan et al., 1998; Qu et al., 1999; Suzuki et al., 2002; Tsuda et al., 2002*). In contrast to global hypomethylation, loss of 5mC at pericentromeric repeats is compatible with human development. Individuals with the rare, autosomal recessive disorder Immunodeficiency, Centromere and Facial anomalies (ICF) syndrome show extensive hypomethylation of pericentromeric repeats, while methylation across the rest of the genome is relatively intact (*Tuck-Muller et al., 2000; Velasco et al., 2018; Weisenberger et al., 2005*). Affected individuals usually die in late childhood or early adulthood, and exhibit variable symptoms including immunoglobulin deficiency, facial dysmorphism, growth retardation and a generalized failure to thrive (*Ehrlich et al., 2008*). Chromosome anomalies including whole-arm deletions and multiradial chromosomes have also been reported in mitogen-stimulated lymphocytes from ICF-patients. However, similar chromosome anomalies are not observed in primary tissues from affected individuals (*Ehrlich, 2003*).

Homozygosity mapping and whole-exome sequencing have separately implicated four genes in ICF syndrome: DNA Methyltransferase 3B (*DNMT3B*, ICF type-1), Zinc-finger and BTB domain containing 24 (*ZBTB24*, ICF type-2), Cell division cycle associated 7 (*CDCA7*, ICF type-3) and Helicase,

lymphoid-specific (*HELLS*, ICF type-4) (*de Greef et al., 2011; Thijssen et al., 2015; Xu et al., 1999*). Most of the described mutations in *DNMT3B* cause amino acid substitutions within the C-terminal catalytic domain, suggesting they may be hypomorphic. In contrast, the majority of mutations in *ZBTB24*, *CDCA7* and *HELLS* are predicted to cause loss of function. Mechanistically, *ZBTB24*, *CDCA7* and *HELLS* are thought to converge in a singular pathway that facilitates *DNMT3B* access to pericentromeric DNA (*Jenness et al., 2018; Wu et al., 2016*).

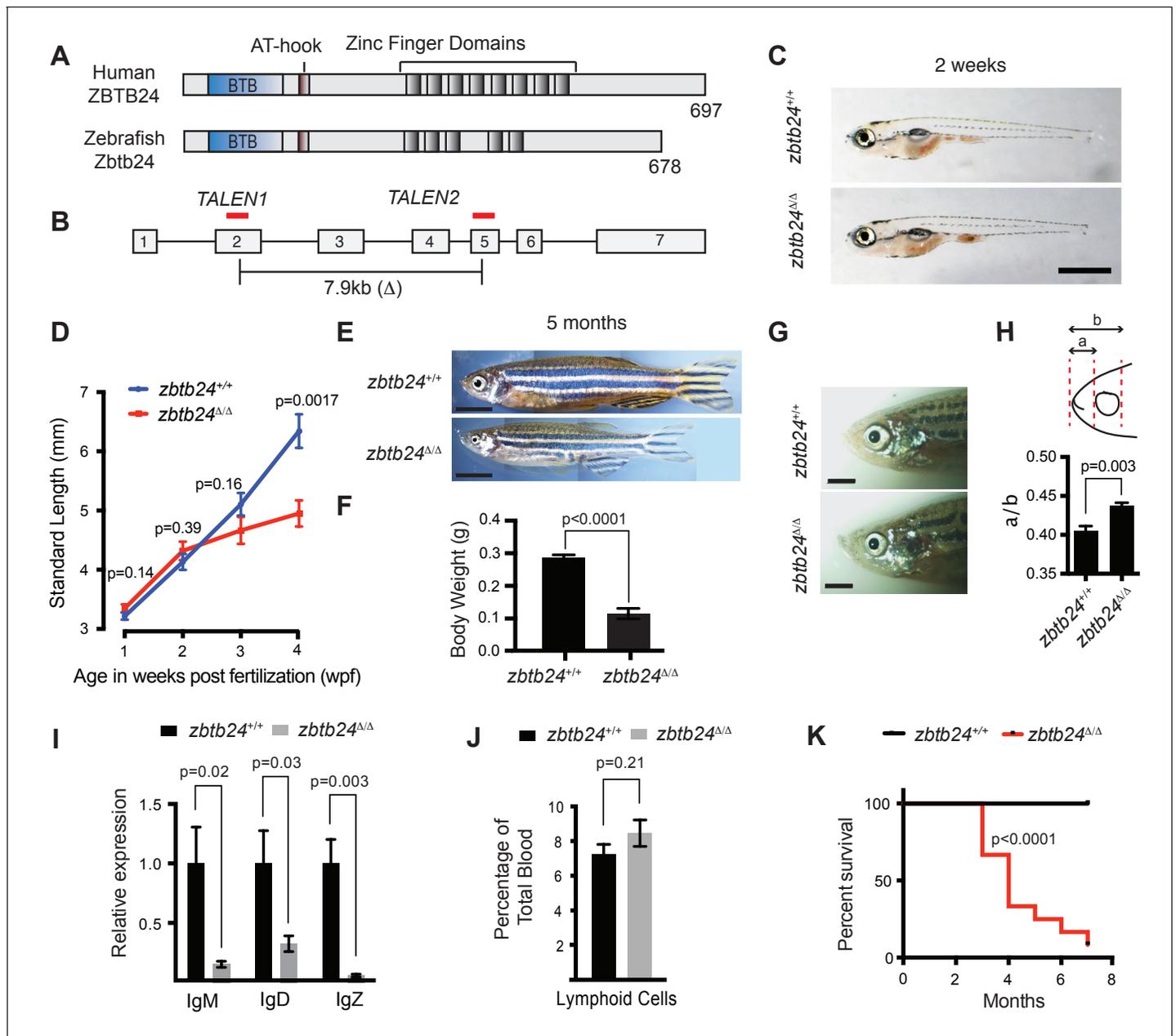
To date, most studies of pericentromeric 5mC loss have been performed using transformed B-cell lines derived from ICF patients carrying mutations in *DNMT3B* (*Ehrlich et al., 2008*). Attempts to generate viable mouse models of pericentromeric hypomethylation through mutation of ICF genes have had limited success. Mice harboring ICF-like mutations in *Dnmt3b* exhibit some characteristics of ICF syndrome including small size and facial anomalies. However, most mice die within 24 hr of birth (*Ueda et al., 2006*). Global methylation profiles were not assessed in these mutants; but significant hypomethylation was reported at both pericentromeric repeats and retroviral sequences. Similar perinatal lethality was observed following deletion of the mouse *HELLS* orthologue. In this case, mutations were accompanied by roughly 50% reductions in 5mC, and methylation loss was noted at pericentromeres, retroviruses and some single copy sequences (*Tao et al., 2011*). Deletion of the mouse *Zbtb24* gene was reported to cause embryonic lethality; but methylation changes in these mutants have not been investigated (*Wu et al., 2016*).

Here, we describe a viable model of pericentromeric methylation loss, generated through mutation of the zebrafish *zbtb24* gene. Homozygous mutant adults exhibited key phenotypic hallmarks of ICF syndrome including hypomethylation of pericentromeric satellite repeats. Hypomethylation of these repeats was first detected in mutants at 2 weeks post fertilization (wpf) and became more severe as animals matured. This progressive methylation loss allowed us to investigate the primary consequences of pericentromeric hypomethylation in the context of a vertebrate animal. Using this model, we link derepression of transcripts from hypomethylated pericentromeres to activation of an interferon-based innate immune response, and we demonstrate that this response is mediated through the MDA5-MAVS RNA sensing machinery. Our findings provide insights into the earliest consequences of pericentromeric hypomethylation, demonstrating an unappreciated function for methylation of pericentromeric repeats in protecting against autoimmunity.

## Results

### Mutation of zebrafish *zbtb24* causes ICF syndrome-like phenotypes

The zebrafish genome encodes a single, well-conserved orthologue of *ZBTB24*, which we mutated using TAL effector nucleases (TALENs) (*Figure 1A* and *Figure 1—figure supplements 1* and *2*). The recovered 7.9 kb deletion allele (*zbtb24<sup>mk22</sup>*; here after referred to as *zbtb24<sup>Δ</sup>*), eliminates coding sequence between exons 2 and 5 (*Figure 1B*). Animals that were homozygous for this deletion lacked detectable *zbtb24* transcripts, suggesting *zbtb24<sup>Δ</sup>* is a null allele (*Figure 1—figure supplement 2D*). *Zbtb24<sup>Δ/Δ</sup>* embryos were born to heterozygous parents at the expected Mendelian ratios and had no obvious morphological abnormalities during the first two weeks of development (*Figure 1C*). Phenotypes that were reminiscent of ICF syndrome emerged as animals matured. Consistent with the small stature observed in ICF syndrome, by 3–4 weeks post fertilization (wpf), *zbtb24<sup>Δ/Δ</sup>* mutant zebrafish were smaller than wild-type siblings raised under identical conditions, and this size reduction persisted into adulthood (*Figure 1D–F*). As adults, *zbtb24<sup>Δ/Δ</sup>* mutants exhibited facial anomalies that were characterized by a quantifiable elongation of the snout (*Figure 1G–H*). We also noted evidence of hypogammaglobulinemia in the presence of normal lymphoid cell numbers, which is an immunological hallmark of ICF syndrome (*Figure 1I–J*). Significant death was noted among homozygous mutants at 4 months of age and fewer than 10% of *zbtb24<sup>Δ/Δ</sup>* animals survived beyond 8 months (*Figure 1K*). Attempts to recover fertilized embryos by intercrossing or outcrossing *zbtb24<sup>Δ/Δ</sup>* adults were unsuccessful, suggesting that animals were sterile (*Figure 1—figure supplement 3A*). Gonadal morphology in *zbtb24<sup>Δ/Δ</sup>* mutants appeared overtly similar to wild-type siblings in histological sections (*Figure 1—figure supplement 3B–C*). However, testes size and sperm count were severely reduced in *zbtb24* mutants, providing one potential explanation for impaired male fertility (*Figure 1—figure supplement 3D–G*). Similar ICF-like phenotypes were observed in zebrafish that were homozygous for a second independently-isolated mutant allele of



**Figure 1.** Mutation of *zbtb24* causes ICF syndrome-like phenotypes in zebrafish. (A) Schematic of human and zebrafish Zbtb24 proteins. The BTB/POZ domain is indicated in blue and C2H2-type zinc fingers in dark grey. (B) Schematic of zebrafish *zbtb24* gene. Location of TALEN target sequences are indicated in red (not to scale). Brackets indicate the region deleted by the *zbtb24*<sup>mk22(Δ)</sup> allele. (C) Representative images of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> zebrafish at 2 wpf. Scale bar: 1 mm. (D) Standard length measurements for *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> zebrafish at 1, 2, 3 and 4 wpf ( $n \geq 6$  for each group). (E) Representative images of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> zebrafish at 5 months. Scale bar: 5 mm. (F) Average weight of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> zebrafish at 5 months ( $n = 5$  for each group). (G) Representative images of facial abnormalities in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> adults at 6 months. Scale bar: 2 mm. (H) Schematic and quantification of facial abnormalities in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> zebrafish ( $n = 5$  for each group). (I) Abundance of IgM, IgD and IgZ transcripts in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> zebrafish at 6 weeks post fertilization ( $n = 5$  for each group). (J) Quantification of lymphoid cell populations in total blood isolated from *zbtb24*<sup>+/+</sup> or *zbtb24*<sup>ΔΔ</sup> kidney marrow from adults, measured by Forward/Side scatter flow cytometry ( $n = 11$  for each group). (K) Kaplan-Meier curve indicating survival among groups of *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>ΔΔ</sup> zebrafish ( $n = 12$  for each group). All error bars indicate standard error of the mean (SEM).

DOI: <https://doi.org/10.7554/eLife.39658.003>

The following figure supplements are available for figure 1:

**Figure supplement 1.** Zbtb24 conservation in vertebrate species.

DOI: <https://doi.org/10.7554/eLife.39658.004>

Figure 1 continued on next page

Figure 1 continued

**Figure supplement 2.** TALEN design for introducing mutations at the endogenous *zbtb24* zebrafish gene.

DOI: <https://doi.org/10.7554/eLife.39658.005>

**Figure supplement 3.** Fertility and gonad analysis in *zbtb24*<sup>Δ/Δ</sup> mutants.

DOI: <https://doi.org/10.7554/eLife.39658.006>

**Figure supplement 4.** A second mutant allele of *zbtb24* recapitulates key features of ICF Syndrome.

DOI: <https://doi.org/10.7554/eLife.39658.007>

*zbtb24* (*zbtb24*<sup>mk19</sup>) (Figure 1—figure supplement 4). Taken together, these findings identify *zbtb24* homozygous mutant zebrafish as a faithful animal model of ICF syndrome phenotypes.

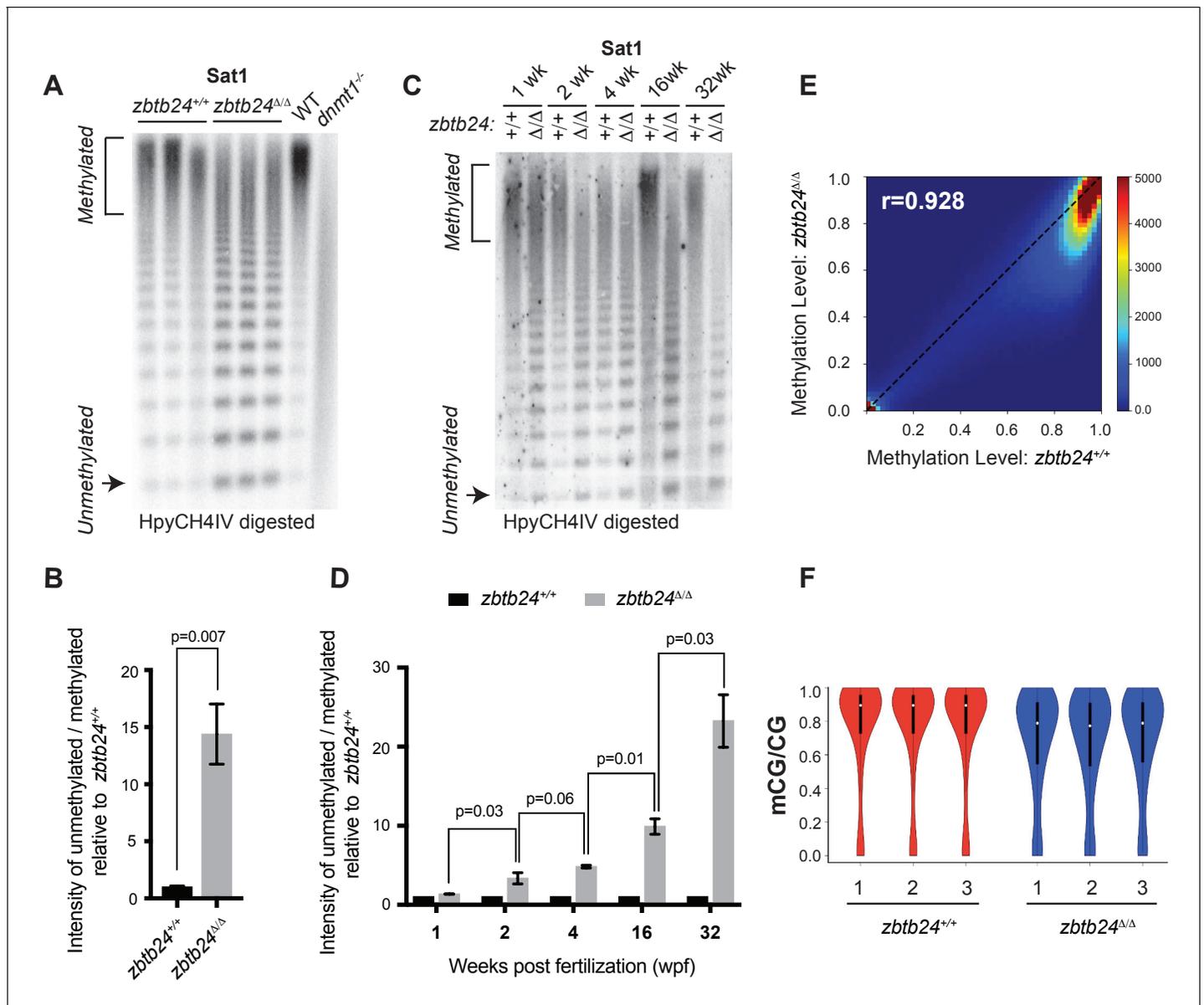
### Progressive methylation loss at pericentromeric repeats in *zbtb24* mutants

Pericentromeric satellite type-1 (Sat1) repeats are found on all zebrafish chromosomes and comprise 5–8% of the zebrafish genome (Phillips and Reed, 2000). As expected, we found that Sat1 sequences from wild-type adults were resistant to digestion with the methylation sensitive restriction enzyme HpyCH4IV, indicating that these pericentromeric repeats were heavily methylated. In contrast to wildtype, Sat1 sequences from *zbtb24*<sup>Δ/Δ</sup> and *zbtb24*<sup>mk19/mk19</sup> mutant adults were readily digested with HpyCH4IV, indicating extensive loss of methylation at these repeats (Figure 2A–B and Figure 2—figure supplement 1A). Comparable Sat1 methylation deficiencies were observed when DNA was isolated from dissected adult brain, skin, muscle and fin, suggesting that these sequences were similarly hypomethylated in most adult somatic tissues (Figure 2—figure supplement 1B and Figure 2—figure supplement 3A). Methylation levels at Sat1 repeats appeared normal in remaining sperm extracted from *zbtb24*<sup>Δ/Δ</sup> mutant adults, suggesting methylation loss may be restricted to somatic tissues (Figure 2—figure supplement 1C–D).

Somewhat unexpectedly, we found that pericentromeric methylation loss in *zbtb24*<sup>Δ/Δ</sup> mutants was progressive. While extensive hypomethylation of Sat1 sequences was detected in adults lacking *zbtb24*, similar hypomethylation was not observed in mutants at 1 wpf (Figure 2C–D). At 2 wpf, *zbtb24*<sup>Δ/Δ</sup> mutants exhibited roughly 3-fold increases in HpyCH4IV digestion, and sensitivity to digestion became increasingly pronounced in older animals (Figure 2C–D). By 32 weeks, Sat1 sequences from *zbtb24* mutants exhibited a 23-fold increase in HpyCH4IV digestion compared to wildtype, suggesting a greater than 95% reduction in methylation of these repetitive sequence blocks. Histone H3 lysine nine trimethylation levels were unaffected at Sat1 sequences in *zbtb24*<sup>Δ/Δ</sup> mutant adults (Figure 2—figure supplement 2).

### *Zbtb24* mutants exhibit modest reductions in 5mC at non-pericentromeric sequences

To clarify whether other sequences were also hypomethylated in *zbtb24* mutants, we performed Enhanced Reduced Representation Bisulfite Sequencing (ERRBS) using genomic DNA isolated from the fins of three 6-month-old *zbtb24*<sup>Δ/Δ</sup> mutant adults and three wild-type siblings (Garrett-Bakelman et al., 2015). At this stage, Sat1 sequences from isolated fins were 20-fold more sensitive to HpyCH4IV in *zbtb24* mutants compared to controls, indicating extensive loss of DNA methylation at pericentromeric repeats (Figure 2—figure supplement 3A–B). We then used ERRBS data to interrogate the methylation status of 979,971 non-pericentromeric CpG sites across the genome in the same tissue samples. Our analysis revealed a strong correlation between genome wide 5mC levels in wild-type and *zbtb24*<sup>Δ/Δ</sup> mutant adults (Pearson's correlation value of 0.928), although overall methylation levels appeared reduced by ~10% at all methylated sequence features in mutants (Figure 2E–F and Figure 2—figure supplement 4). Reductions consisted primarily of small-magnitude changes in 5mC across the genome, with only 1.3% (13,205) of examined CpG dinucleotides exhibiting methylation differences of greater than 20%. Consistent with these findings, at a threshold of 20% change (p-value<0.01), only 55 differentially methylated regions (DMRs) were identified between wild-type and *zbtb24*<sup>Δ/Δ</sup> adults (Supplementary file 4). Methylation levels at endogenous retroviruses and other transposable elements were also examined by methylation sensitive restriction digest. All tested elements were similarly resistant to digestion in *zbtb24*<sup>Δ/Δ</sup> mutant adults and



**Figure 2.** Mutation of *zbtb24* causes progressive methylation loss at pericentromeric satellite repeats. (A) Southern blot of genomic DNA digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish Sat1 sequence. Each lane represents DNA isolated from one adult individual of the indicated genotype. DNA from *dnmt1*<sup>-/-</sup> zebrafish larvae at 7 days post fertilization and their phenotypically wild-type siblings (WT) provides a positive control. (B) Quantification of methylation changes at Sat1 sequences in panel A. Error bars indicate SEM from the three biological replicates. (C) Southern blot of genomic DNA digested with 5mC-sensitive restriction enzyme HpyCH4IV and probed with zebrafish Sat1 sequence. Genomic DNA was isolated from *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> animals at 1, 2, 4, 16 and 32 wpf as indicated. (D) Quantification of methylation changes at Sat1 sequences in panel C. Data represent averages from two independent experiments. Error bars represent the standard deviation (SD). (E) Correlation heatmap of CpG methylation levels in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> as assessed by ERRBS (Data reflects three biological replicates of each genotype). The density of CpGs increases from blue to dark red. (F) Violin Plots indicating overall CpG methylation levels in fins from adult *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish.

DOI: <https://doi.org/10.7554/eLife.39658.008>

The following figure supplements are available for figure 2:

**Figure supplement 1.** *Zbtb24* mutation causes methylation loss at pericentromeric repeats.

DOI: <https://doi.org/10.7554/eLife.39658.009>

**Figure supplement 2.** Repressive histone modifications are unaffected in *zbtb24*<sup>Δ/Δ</sup> mutants.

DOI: <https://doi.org/10.7554/eLife.39658.010>

**Figure supplement 3.** *Zbtb24* mutants exhibit modest reductions in 5mC at non-pericentromeric sequences.

DOI: <https://doi.org/10.7554/eLife.39658.011>

Figure 2 continued on next page

Figure 2 continued

**Figure supplement 4.** DNA Methylation levels at different genomic classes.

DOI: <https://doi.org/10.7554/eLife.39658.012>

**Figure supplement 5.** Methylation at interspersed repeats is unaffected in *zbtb24* mutants.

DOI: <https://doi.org/10.7554/eLife.39658.013>

wild-type siblings (**Figure 2—figure supplement 5**). Collectively, these data indicate that pericentromeres are a predominant site of methylation loss in *zbtb24*<sup>Δ/Δ</sup> mutants.

### Mutation of *zbtb24* causes activation of innate immune response genes

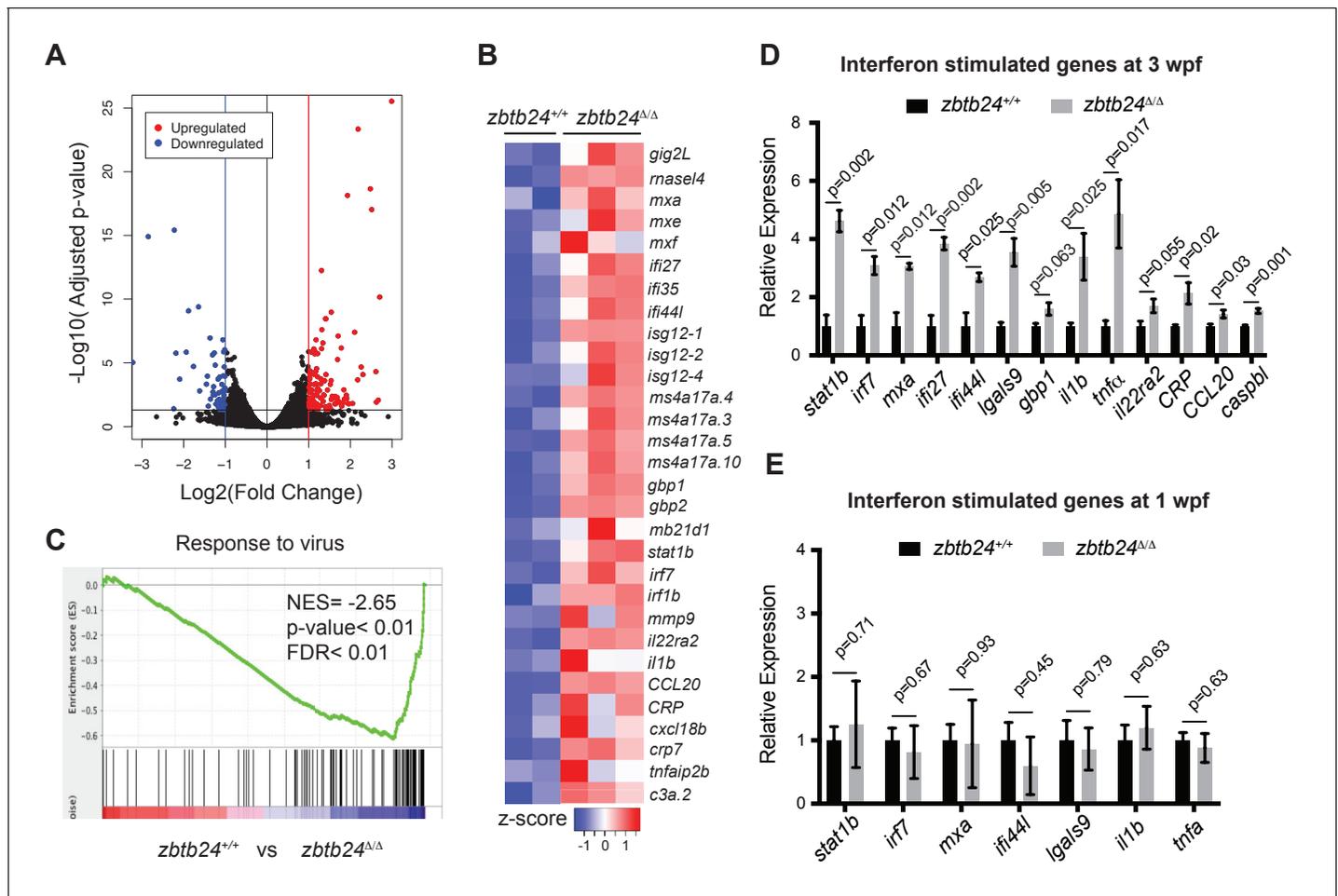
To gain insights into the early consequences of methylation loss in *zbtb24* mutants, we performed transcriptome analysis on RNA isolated from wild-type and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 2 wpf. At this stage, *zbtb24*<sup>Δ/Δ</sup> mutants remain morphologically indistinguishable from wildtype, but show clear hypomethylation of pericentromeric sequences. RNA-seq identified 58 genes that were downregulated by more than 2-fold in *zbtb24*<sup>Δ/Δ</sup> larvae at 2 wpf, while 119 were upregulated by 2-fold or more (**Figure 3A**). No gene enrichment signature was observed among downregulated genes. However, roughly 30% of upregulated genes were associated with activation of the innate immune system. In particular, we noted that upregulated transcripts included those associated with interferon stimulated genes (ISGs) and inflammatory cytokines (**Figure 3B**). Consistent with these observations, Gene Set Enrichment Analysis (GSEA) identified significant enrichment of genes involved in viral response, a key function of innate immune pathways (**Figure 3C**). Upregulation of ISGs was also observed in *zbtb24*<sup>Δ/Δ</sup> and *zbtb24*<sup>mk19/mk19</sup> mutants by qRT-PCR at 3 wpf, whereas the same genes were expressed at wild-type levels at 1 wpf (**Figure 3D–E** and **Figure 3—figure supplement 1**). No immune-related genes (and only one gene differentially upregulated in the RNA-Seq) were found within 100 kb of identified DMRs, suggesting that direct loss of methylation at these sequences was unlikely to cause the response (**Figure 2—figure supplement 3D** and **Supplementary file 4**). Consistent with previous studies, we found that global methylation depletion using the DNA methyltransferase inhibitor 5-azacytidine also resulted in upregulation of immune response genes (**Figure 3—figure supplement 2**).

### The innate immune response in *zbtb24* mutants is mediated by sensors of cytosolic RNA

The innate immune system represents an ancient defense system in which pathogen-associated molecular patterns (PAMPs) are recognized by pattern recognition receptors (PRRs). These PRRs induce signaling cascades that drive the production of interferons and other inflammatory cytokines with antiviral and immune modulating functions (**Schneider et al., 2014**). In addition to extracellular pathogens, PRRs also recognize PAMPs associated with cell-intrinsic stimuli including DNA damage, endogenous retroviral RNA and RNA-DNA hybrids (**Chiappinelli et al., 2015; Härtlova et al., 2015; Mankan et al., 2014; Roulois et al., 2015**).

To clarify the origin of the response in *zbtb24* mutants, we examined the major families of PRRs involved in innate immunity. These include the Toll-like receptors (TLRs), which have broad functions in detecting PAMPs, the RIG-I like receptors (RLRs), which are involved in the detection of cytosolic RNA and cGAMP synthase (cGAS), which functions as a cytosolic sensor of DNA and RNA/DNA hybrids (**Crowl et al., 2017**). Mutations in key mediator proteins required to propagate interferon signaling from each PRR family were introduced onto the *zbtb24* mutant background and we tested the effect on ISG expression. Mutations in the zebrafish orthologs of *mitochondrial antiviral-signaling protein* (*mavs*), which is an intermediate in RLR signaling and stimulator of interferon genes (*sting*), which is involved in cGAS signaling were generated using CRISPR/Cas9 technology (**Figure 4—figure supplement 1A–B**). The mutant allele of *Myeloid differentiation primary response 88* (*myd88*), which is required for signaling through most TLRs, was previously described (**van der Vaart et al., 2013**).

As in prior experiments, significant increases of the ISGs, *signal transducer and activator of transcription 1b* (*stat1b*) and *interferon regulatory factor* (*irf7*) were observed in *zbtb24*<sup>Δ/Δ</sup> larvae at 3 wpf by qRT-PCR (**Figure 4A–C**). Introduction of *myd88* or *sting* mutations had little impact on



**Figure 3.** Mutation of *zbtb24* leads to activation of innate immune response genes. (A) Volcano plot representation of differential gene expression in *zbtb24*<sup>+/+</sup> vs *zbtb24*<sup>Δ/Δ</sup> zebrafish at 2 wpf. Blue and red points mark genes with >2 fold downregulation or upregulation respectively. (B) RNA-seq heatmap showing innate immune genes upregulated in *zbtb24*<sup>Δ/Δ</sup> mutant compared to *zbtb24*<sup>+/+</sup> siblings. Shown are Z-score normalized gene expression values. (C) GSEA of a set of genes involved in Response to Virus in zebrafish comparing *zbtb24*<sup>+/+</sup> vs *zbtb24*<sup>Δ/Δ</sup>. NES, normalized enrichment score; FDR, false discovery rate. (D) qRT-PCR demonstrating upregulated interferon and inflammatory response genes in *zbtb24*<sup>Δ/Δ</sup> mutants at 3 wpf. Expression levels are reported relative to  $\beta$ -actin. Error bars indicate SEM from at least 3 independent biological replicates with n = 8 total animals for each replicate. (E) qRT-PCR analysis reveals similar expression of interferon genes in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> larvae at 1 wpf. Error bars represent SEM from at least five biological replicates.

DOI: <https://doi.org/10.7554/eLife.39658.014>

The following figure supplements are available for figure 3:

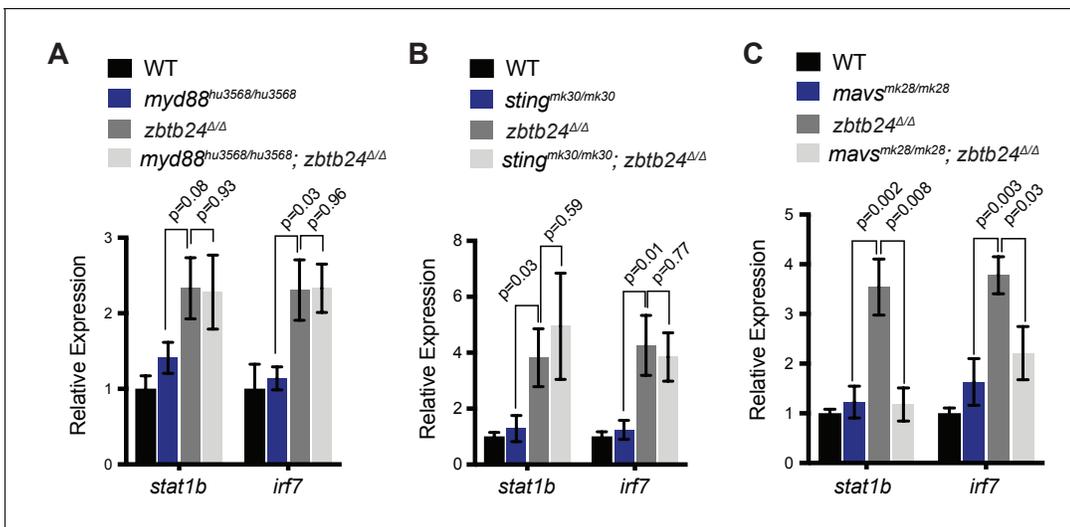
**Figure supplement 1.** Mutation in *zbtb24* leads to activation of innate immune response pathways.

DOI: <https://doi.org/10.7554/eLife.39658.015>

**Figure supplement 2.** Activation of interferon stimulated genes upon treating zebrafish embryos with 5aza-cytidine.

DOI: <https://doi.org/10.7554/eLife.39658.016>

expression of these ISGs, as similar transcript levels were detected in *zbtb24*<sup>Δ/Δ</sup> single mutant animals compared to *myd88*<sup>hu3568/hu3568</sup>; *zbtb24*<sup>Δ/Δ</sup> or *sting*<sup>mk30/mk30</sup>; *zbtb24*<sup>Δ/Δ</sup> double mutants (Figure 4A–B). Sustained ISG expression in these double mutants suggests limited roles for TLR and cGAS PRRs in mediating the interferon response in *zbtb24* mutants. In contrast to *myd88* and *sting*, mutation of *mavs* suppressed *stat1b* and *irf7* upregulation in *zbtb24*<sup>Δ/Δ</sup> mutant animals. Expression levels of *irf7* and *stat1b* were reduced 2- and 4-fold respectively in *mavs*<sup>mk28/mk28</sup>; *zbtb24*<sup>Δ/Δ</sup> double mutants when compared to *zbtb24*<sup>Δ/Δ</sup> single mutant zebrafish, indicating a requirement for *mavs* in



**Figure 4.** Interferon response in *zbtb24* mutants is mediated by sensors of cytosolic RNA. (A) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 3 wpf.  $n = 4$  biological replicates. (B) Expression of the ISGs *stat1b* and *irf7* in indicated genotypes at 3 wpf.  $n \geq 7$  biological replicates. (C) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 3 wpf.  $n \geq 5$  biological replicates. All error bars indicate SEM.

DOI: <https://doi.org/10.7554/eLife.39658.017>

The following figure supplement is available for figure 4:

**Figure supplement 1.** Mutation of zebrafish orthologs of *mavs*, *sting*, and *mda5*.

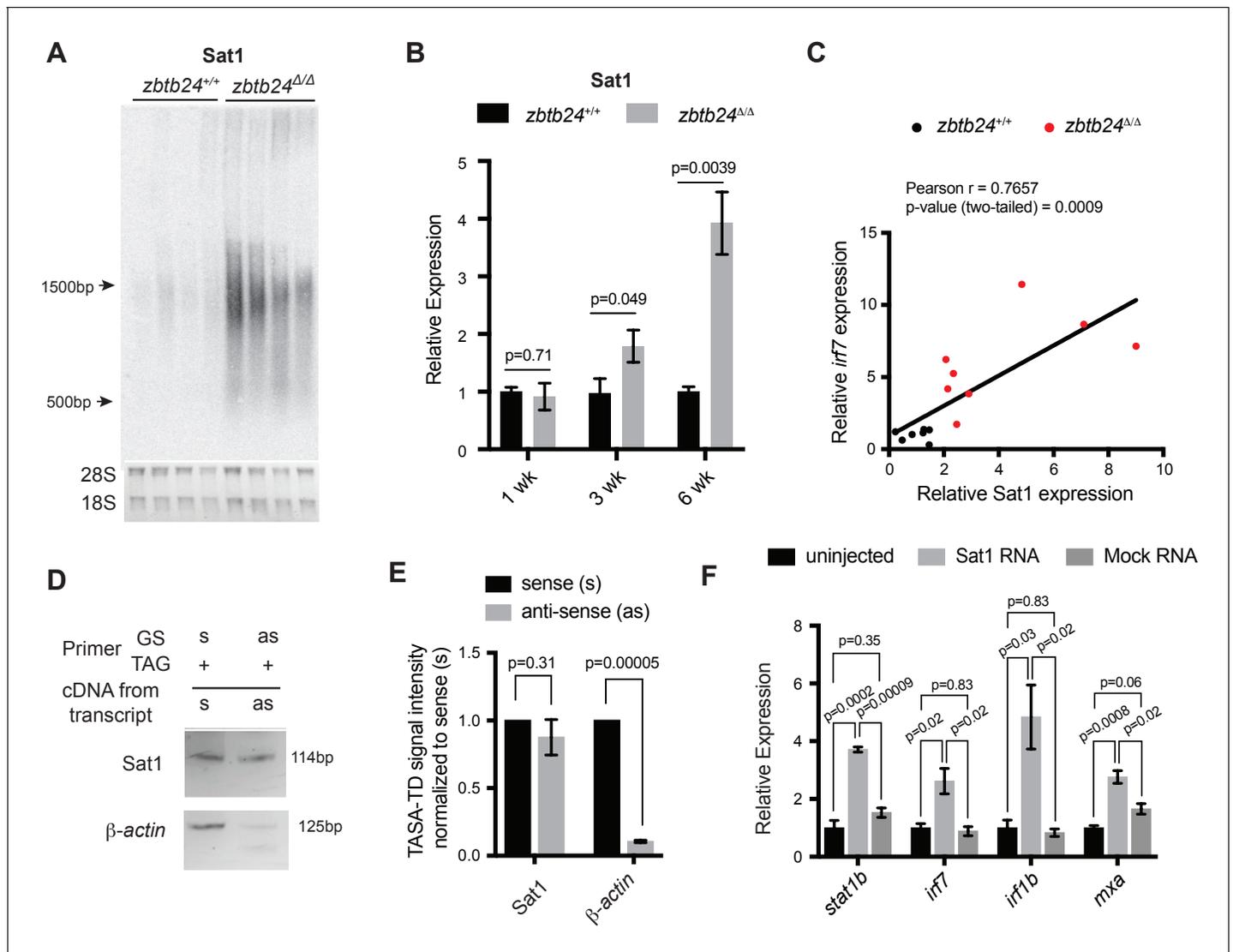
DOI: <https://doi.org/10.7554/eLife.39658.018>

the upregulation of these ISGs (Figure 4C). This finding implicates RLR signaling in the activation of the innate immune system in *zbtb24* mutants and suggests a cytosolic RNA trigger for the response.

### Pericentromeric RNA transcripts are sufficient to trigger the interferon response in *zbtb24* mutants

Given known roles for DNA methylation in transcriptional repression, we next tested whether loss of methylation at pericentromeric sequence resulted in increased levels of Sat1 transcripts that could trigger the RNA mediated interferon response. Consistent with this model, strong derepression of Sat1 RNA from hypomethylated pericentromeres was noted in *zbtb24* mutant adults (Figure 5A and Figure 5—figure supplement 1A), whereas transcripts for other dispersed repetitive elements remained unchanged between mutants and wildtype (Figure 5—figure supplement 1B). Increases in Sat1 transcripts correlated with levels of *irf7* expression in adult zebrafish ( $r = 0.77$ ), and upregulation of Sat1 transcripts coincided with the window of ISG induction during development (Figure 5B–C). Both sense and antisense transcripts were detected in mutants using TAG-aided sense/antisense transcript detection (TASA-TD) strand-specific PCR (Henke et al., 2015), suggesting the potential for derepressed Sat1 transcripts to form double stranded RNAs (Figure 5D–E).

To determine whether Sat1 transcripts were sufficient to activate an innate immune response, in vitro synthesized RNA corresponding to Sat1 sense and antisense transcripts were injected into wild-type embryos at the 1 cell stage. Expression of the ISGs *stat1b*, *irf7*, *irf1b* and *mxr* was then assessed at 8 hr post fertilization. Co-injection of sense and antisense Sat1 RNA was sufficient to reproducibly cause a 2 to 4-fold upregulation in expression of these ISGs, whereas combined injection of sense and antisense control transcripts encoding a fragment of zebrafish  $\beta$ -actin or GFP had no effect on expression of these genes (Figure 5F and Figure 5—figure supplement 2). Lower level upregulation of some, but not all ISGs was noted when sense or antisense Sat1 transcripts were individually injected into the embryo, suggesting that the response was primarily triggered by formation of Sat1 dsRNA (Figure 5—figure supplement 2). Collectively, these results functionally link the derepression of Sat1 transcripts to the activation of the innate immune response in *zbtb24* mutants.



**Figure 5.** Pericentromeric transcripts are sufficient to induce the innate immune response in *zbtb24* mutants. (A) Northern blot analysis of Sat1 transcripts in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 6 wpf. Each lane represents a biological replicate. The lower panel represents the cropped ethidium-bromide stained gel as loading control. (B) qRT-PCR for Sat1 transcripts in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> zebrafish at 1, 3 and 6 wpf. Error bars indicate SEM of at least four biological replicates in each group. (C) Correlation between the expression of Sat1 and *irf7* in *zbtb24*<sup>+/+</sup> and *zbtb24*<sup>Δ/Δ</sup> at 6 weeks (n = 15). (D) TASA-TD PCR amplified sense (s) and antisense (as) transcripts Sat1 (114 bp) and  $\beta$ -actin (125 bp) from first strand *zbtb24*<sup>Δ/Δ</sup> cDNA. PCR primers: gene-specific (GS); TAG. The products from TASA-TD PCR were run on the same gel, then cropped and presented. This panel is representative of two independent biological replicates. (E) Quantification of TASA-TD from panel D. Error bars indicate SD from two biological replicates. (F) Expression of interferon stimulated genes in wild-type embryos injected with Sat1 or control RNA encoding a similar-sized fragment of  $\beta$ -actin. 50 pg of in vitro transcribed sense and antisense transcripts were injected into wild-type zebrafish embryos at the 1 cell stage. Total RNA was extracted at 8 hr post fertilization for qRT-PCR analysis. Error bars indicate SEM from at least three biological replicates with n = 20 embryos for each biological replicate.

DOI: <https://doi.org/10.7554/eLife.39658.019>

The following figure supplements are available for figure 5:

**Figure supplement 1.** Mutation in *zbtb24* upregulates Sat1 transcripts but not transposons.

DOI: <https://doi.org/10.7554/eLife.39658.020>

**Figure supplement 2.** Effect of injecting in vitro transcribed Sat1 RNAs on expression of interferon stimulated genes.

DOI: <https://doi.org/10.7554/eLife.39658.021>

## The cytosolic RNA helicase MDA5 is required for the interferon response in *zbtb24* mutants

Finally, we sought to identify the specific PRR required for the interferon response in *zbtb24* mutants. The RLR family of PRRs consists of two RNA helicases that signal through Mavs: Melanoma Differentiation-Associated protein 5 (Mda5) and Retinoic acid-inducible gene I (Rig-I). Rig-I binds 5' triphosphorylated RNA molecules, whereas Mda5 has been implicated in the recognition of long double-stranded RNAs in the cytosol (Crowl et al., 2017). Given that 5' triphosphorylation of RNAs is a typical viral signature that is unlikely to be present on endogenous RNA transcripts, we reasoned that Mda5 was a more likely candidate for the receptor. To test the requirement for *mda5*, we generated a seven base-pair deletion in this gene that disrupted the DEAD box helicase domain (Figure 4—figure supplement 1C). This *mda5*<sup>mk29</sup> allele was then introduced onto the *zbtb24* mutant background, and expression of the ISGs *stat1b* and *irf7* was examined at 3 wpf and 6 wpf. Homozygous mutation of *mda5* was sufficient to restore *stat1b* and *irf7* expression to wild-type levels in *zbtb24*<sup>Δ/Δ</sup> mutant larvae, suggesting that Mda5 is the primary PRR required for the response (Figure 6A and B). This requirement was further validated by RNA-seq, which revealed that a broad panel of ISGs that showed elevated expression in *zbtb24* single mutants were no longer upregulated in *mda5*<sup>mk29/mk29</sup>; *zbtb24*<sup>Δ/Δ</sup> double mutants (Figure 6C).

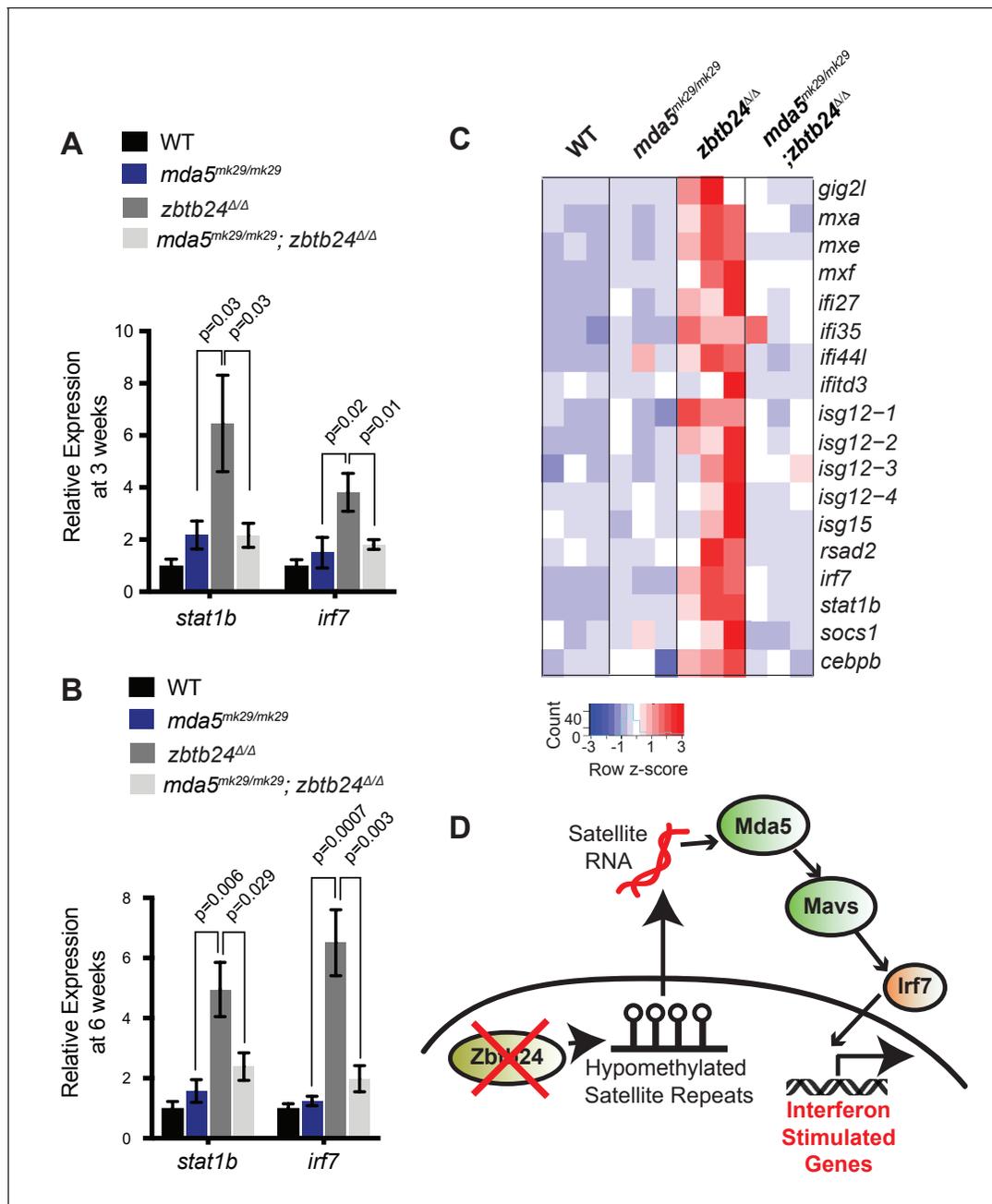
Taken together, these results support a model in which derepression of transcripts from hypomethylated pericentromeres triggers activation of the innate immune system through the Mda5/Mavs viral RNA recognition pathway (Figure 6D). These findings identify roles for pericentromeric RNA as a trigger of autoimmunity and reveal important functions for pericentromeric methylation in suppressing the generation of these immunostimulatory transcripts. Based on these results, we propose that induction of the innate immune system is one of the earliest in vivo consequences of pericentromeric methylation loss.

## Discussion

In this study, we describe a viable animal model of ICF syndrome which recapitulates key phenotypic hallmarks of the disease including slow growth, facial anomalies, immunoglobulin deficiencies and reduced lifespan. Given that previous attempts to model ICF syndrome have resulted in perinatal or embryonic lethality (Geiman et al., 2001; Ueda et al., 2006; Wu et al., 2016), this zebrafish model provides an important new resource for understanding ICF disease etiology during juvenile and adult life stages. In particular, *zbtb24* mutant zebrafish will be useful for understanding phenotypes such as immunoglobulin deficiency, which have not been observed in mouse models and are difficult to study in cell culture systems.

As in ICF syndrome, *zbtb24* mutant adult zebrafish exhibited extensive loss of methylation at pericentromeric sequences. For highly repetitive sequences, methylation sensitive restriction digest followed by Southern blot remains the most effective way to assess methylation levels. By this approach, we observed increases in HpyCH4IV digestion that are consistent with up to 95% reductions in methylation at Sat1 pericentromeric repeats in *zbtb24* mutants. While similar hypomethylation was observed in all adult somatic tissues that we examined, we unexpectedly observed that methylation levels in sperm from *zbtb24* mutants and wildtype animals appeared comparable. This finding raises the possibility that different pathways act to control pericentromeric methylation in germ and somatic cells.

Methylation levels at pericentromeric Sat1 sequences could not be quantified by ERRBS, as this technique relies on MspI restriction digest to enrich for CpG containing sequences, and zebrafish Sat1 repeats are lacking in this restriction site. Nonetheless, ERRBS analysis suggested that the general methylation landscape in human ICF syndrome and in *zbtb24* mutant zebrafish is similar. Methylome analysis of primary blood from ICF patients identified methylation changes of greater than 20% at roughly 3% of examined CpG dinucleotides. Significant changes in methylation of retroviruses and other dispersed repeats were not observed in these patients (Velasco et al., 2018). Consistent with these findings, our ERRBS analysis revealed methylation changes of greater than 20% at roughly 1.3% of assayed CpG dinucleotides and found methylation of dispersed repeats to be similar between wildtype and in *zbtb24* mutant zebrafish. The low-level methylation changes outside of the pericentromeres observed in ICF syndrome and our mutants raise the possibility that *zbtb24* may have additional modest roles in maintaining methylation at non pericentromeric sequences. One



**Figure 6.** Mutation of cytosolic RNA receptor Mda5 mitigates the interferon response in *zbtb24<sup>Δ/Δ</sup>* zebrafish. (A) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 3 wpf.  $n \geq 7$  biological replicates. (B) Expression of interferon signaling genes *stat1b* and *irf7* in indicated genotypes at 6 wpf.  $n = 6$  biological replicates. (C) RNA-seq heatmap of interferon stimulated genes upregulated in *zbtb24<sup>Δ/Δ</sup>* zebrafish and rescued in *mda5<sup>mk29/mk29</sup>; zbtb24<sup>Δ/Δ</sup>* zebrafish at 3 wpf. Shown are Z-score normalized gene expression values. (D) Model for the activation of interferon response in *zbtb24* mutants. Loss of Zbtb24 function causes hypomethylation of pericentromeric Sat1 repeats, which leads to derepression of associated Sat1 transcripts. These pericentromeric transcripts are recognized by the RNA helicase Mda5 which signals through Mavs and Irf7 to upregulate ISGs. Autoregulatory feedback implicates *irf7* as both an ISG and a key downstream effector of Mda5/Mavs signaling.

DOI: <https://doi.org/10.7554/eLife.39658.022>

The following figure supplement is available for figure 6:

**Figure supplement 1.** Expression of ICF genes *cdca7* and *hells* in *zbtb24* mutants (A) Expression levels of *cdca7* and *hells* in the RNA Seq data set reported in **Figure 6C** ( $n = 3$  biological replicates for each group).

DOI: <https://doi.org/10.7554/eLife.39658.023>

important caveat of ERRBS analysis is that CpG poor sequences can be under represented, leaving open the possibility that additional DMRs in CpG poor regions of the genome were overlooked by our approach.

The progressive loss of 5mC we observe in somatic tissues between larval and adult stages implicates *Zbtb24* in regulating the long-term maintenance of methylation at pericentromeric repeats. We are unaware of any developmental or methylation milestones that can account for the onset of hypomethylation around 2 wpf. Rather, we speculate that the onset of methylation loss at this stage partly reflects the need to deplete maternally loaded *zbtb24* prior to unmasking of the *zbtb24* mutant phenotype and partly reflects the culmination of minor methylation losses due to lower fidelity maintenance over many rounds of cell division. We note that the onset of ICF-like growth defects in *zbtb24* mutant zebrafish emerged in the weeks following Sat1 methylation loss. In at least one case of ICF syndrome type 2, growth reductions and immunodeficiency were also reported to develop with age, raising the possibility that similar progressive methylation loss may impact ICF etiology in humans (*von Bernuth et al., 2014*). It is also possible that *Zbtb24* functions in both maintenance and establishment of pericentromeric methylation, but that requirements for establishment are masked by maternally deposited RNA in *zbtb24* mutant zebrafish lines. Unfortunately, *zbtb24* homozygous mutant zebrafish are sterile, preventing the generation of the maternal-zygotic mutants required to address this question.

Previous studies have suggested that ZBTB24 is a transcription factor that may act to regulate DNA methylation through transcriptional control of the ICF gene *CDCA7* (*Wu et al., 2016*). Consistent with this model, we observe near complete loss of *cdca7* expression in *zbtb24* mutants in our RNA-seq data set and by qRT-PCR (*Figure 6—figure supplement 1*). A more recent study in cultured human cells proposed that ZBTB24 binding might be directly involved in recruiting DNMT3B to promote gene body methylation through recognition of AGGTCCTGGCAG motifs in human cells (*Thompson et al., 2018*). Analysis using Find Individual Motif Occurrences (FIMO) (*Grant et al., 2011*), did not reveal this motif in the promoter or gene body of zebrafish *cdca7* or at Sat1 sequences.

In the current study, we take advantage of the progressive Sat1 methylation loss in *zbtb24* mutants to identify activation of interferon signaling as one of the earliest in vivo consequences of pericentromeric hypomethylation. This phenotype cannot be attributed to defects in adaptive immunity, as the zebrafish adaptive immune system is not functional until roughly 4 wpf (*Trede et al., 2004*). Induction of an interferon response has been reported in the context of global hypomethylation in cancer cell lines treated with the DNA methyltransferase inhibitor 5-azacytidine and in zebrafish mutated for the maintenance DNA methyltransferase machinery (*Chernyavskaya et al., 2017; Chiappinelli et al., 2015; Roulois et al., 2015*). In each of these cases induction of the interferon response was attributed to massive derepression of endogenous retroviral elements.

Our results are distinguished from these earlier studies in that we identify hypomethylation of pericentromeric sequences and subsequent derepression of associated satellite transcripts as a previously unappreciated trigger of innate immunity. Immunostimulatory motifs have been noted in pericentromeric RNAs derived from mouse and humans, and transcripts derived from these repeats have been observed in p53 null mouse fibroblasts following global methylation loss (*Leonova et al., 2013; Tanne et al., 2015*). However, while these studies suggest the potential for pericentromeric hypomethylation to drive an interferon response in diverse vertebrate species, experimental evidence in support of this model has been lacking. Here we demonstrate a causative link between derepression of pericentromeric RNAs and the interferon response, and identify a requirement for Mda5/Mavs in mediating the response. Our findings suggest that aberrant Sat1 transcripts derived from pericentromeric repeats trigger this response, and that these transcripts may mimic features of double stranded RNA viruses in the cytosol. This finding raises the possibility that this pathway may also recognize additional endogenous RNAs that lack viral origin.

While mutation of *mda5/mavs* rescued the interferon response in *zbtb24* mutants, *mda5/mavs* mutation had little impact on other ICF phenotypes observed in *zbtb24* mutants. Therefore, we find it unlikely that the interferon response drives ICF etiology. Rather this response represents an additional consequence of pericentromeric hypomethylation. Hypomethylation of pericentromeric sequences is compatible with human viability and is observed in abnormal cell contexts including cancer and senescence. Massive increases in pericentromeric transcripts and upregulation of interferon genes have both been noted in cancer (*Cheon et al., 2014; Ting et al., 2011*). Our data raise

the possibility that pericentromeric hypomethylation and subsequent derepression of associated RNAs represents an important but underappreciated trigger of autoimmunity in a variety of disease states.

## Materials and methods

### Zebrafish husbandry

Zebrafish husbandry and care were conducted in full accordance with animal care and use guidelines with approval by the Institutional Animal Care and Use Committees at Memorial Sloan Kettering Cancer Center and the University of Georgia. Zebrafish were raised under standard conditions at 28°C. Wild-type lines were of the AB background. All mutant alleles are summarized in *Supplementary file 1*.

### TALEN and CRISPR mutagenesis

TALEN sequences were selected using Targeter 2.0 software (Doyle *et al.*, 2012). TAL repeat assembly was achieved using the Golden Gate assembly method, and assembled repeats were integrated into the GoldyTALEN scaffold (Bedell *et al.*, 2012; Cermak *et al.*, 2011). Assembled vectors served as templates for in vitro mRNA transcription using the T3 mMessage mMachine kit (Ambion) according to manufacturer's instructions. 50–100 pg mRNA was injected into wild-type embryos at the one-cell stage. Injected embryos were raised to adulthood and F1 progeny were screened for germline transmission of mutations as previously described (Li *et al.*, 2015). Primers used for detection of mutations and subsequent genotyping are included in *Supplementary file 1*.

Target selection for CRISPR/Cas9 mediated mutagenesis was performed using CHOPCHOP (Labun *et al.*, 2016). sgRNA templates were generated either by cloning into pT7-gRNA as described by Jao *et al.* (2013) or using the oligo-based approach described in Gagnon *et al.*, 2014 and Burger *et al.* (2016). All template oligos are listed in *Supplementary file 3*. sgRNAs were in vitro transcribed from their respective templates using T7 RNA polymerase (Promega) as per manufacturer protocol. Cas9 RNA was in vitro transcribed from the pT3TS-nls-zCas9-nls plasmid (Jao *et al.*, 2013) using the T3 mMessage mMachine Kit (Ambion). For mutagenesis, 200–400 ng of sgRNA and ~500 ng of Cas9 mRNA were co-injected into wild-type embryos at the one-cell stage. Injected embryos were raised to adulthood, and F1 progeny were screened for germline transmission of mutations as previously described (Li *et al.*, 2015). Primers used for detection of mutations and subsequent genotyping are included in *Supplementary file 1*.

### Zebrafish imaging and length measurements

All bright field imaging of zebrafish larvae and adult was performed using Olympus MVX10 with CellSens Standard software. Standard-length was documented using ImageJ as defined in Parichy *et al.* (2009). Photoshop (Adobe) adjustments to brightness and contrast were equally applied to all images of whole zebrafish in order to improve visualization.

### FACS analysis of whole kidney marrow

Adult zebrafish at 6 months were sacrificed with a combination of tricaine (Sigma-Aldrich, CAS number 886-86-2) and rapid chilling. Whole kidneys were dissected using forceps and placed in 0.9 × PBS/5% FCS. Manual disaggregation using a P1000 pipette resulted in single cell suspensions. Cells were filtered over a 40 µm nylon mesh filter, and resuspended in PBS/FCS to give a final concentration of 100,000 cells/µl. FACS sorting of single cells were analyzed for forward/side scatter profiles. FACS data were analyzed using FloJo software.

### Histology

For Hematoxylin and Eosin (H and E) staining, adult zebrafish were fixed in 10% Neutral Buffered Formalin for 48 hr. Zebrafish were then decalcified in 0.5 M EDTA for 24 hr. After decalcification, fish were incubated overnight in 70% Ethanol before embedding in paraffin blocks. Sections were stained with H and E according to standard procedures.

## Sperm count

Adult zebrafish at 8 months were sacrificed with a combination of tricaine and rapid chilling. Whole testis was dissected using forceps and crushed in 100  $\mu$ l of PBS. For determining sperm-count, sperm samples were diluted 1:20 for each fish. 10  $\mu$ l of the diluted sample was then loaded onto a hemocytometer (Bright-Line, Hauser Scientific) for counting. The volume over the central counting area is 0.1  $\text{mm}^3$  or 0.1 microliter. Average number of sperm counted over the central counting area was multiplied by 10000 to obtain the number of sperm/ml of the diluted sample. The obtained value was multiplied by the dilution factor to obtain the final sperm count.

## DNA methylation analysis

For Southern blot analysis, 1  $\mu$ g of purified total genomic DNA was digested with the indicated methylation sensitive restriction enzyme, fractionated by electrophoresis through a 0.9% agarose gel and transferred to nylon membrane. Sperm DNA was isolated from sperm samples collected by crushing dissected testes in PBS. Probes were PCR amplified using primers in **Supplementary file 2** and radiolabeled with  $^{32}\text{P}$ -dCTP using Rediprime<sup>TM</sup> II Random Prime Labelling System (Amersham) according to manufacturer protocol. Hybridization signals were imaged and analyzed using a Typhoon phosphorimager (GE Life Sciences). Signal intensities were measured using ImageJ. Methylation changes at Sat1 was quantified as a ratio of the intensity of the unmethylated/methylated blot regions as indicated in the respective blot.

HypCH4IV was selected for Sat1 methylation analysis over the more traditional MspI/HpaII isoschizomer pair because Sat1 sequences lack the CCGG sites that are recognized by these enzymes.

## Chromatin immunoprecipitation (ChIP)

ChIP was performed as described in Lindeman et al. 2009 with modifications. Briefly, zebrafish juveniles at 1 month were euthanized using tricaine. Chromatin was prepared from euthanized fish by lysing flash frozen samples using an automated pulverizer (Covaris) and crosslinking using 1% Formaldehyde for 5 mins. Chromatin shearing was performed using a Covaris S220 sonicator using the following conditions: 1 ml tubes with total chromatin from each fish in buffer containing 1% SDS were sonicated using peak intensity power of 140, duty factor of 5.0 and 200 cycles per burst, for 14 min for *zbtb24*<sup>+/+</sup> and 6 min for *zbtb24* <sup>$\Delta/\Delta$</sup> . Shearing was monitored using 1% agarose gel. To provide standardized input for each ChIP experiment, chromatin was diluted to A260 = 0.2. For each ChIP, 2  $\mu$ g antibody per 10  $\mu$ l Dynabeads and 100  $\mu$ l chromatin was incubated overnight at 4°C. Following antibodies were used in this study: anti-H3K9me3 antibody (abcam ab8898), anti-H3K27me3 (Millipore 07-449), anti-H3 (abcam 1791) and IgG control (abcam ab15008). After elution, ChIP DNA and input controls were purified using QIAquick PCR purification kit (Qiagen). Eluted DNA was analyzed by qPCR using primers targeting Sat1 (**Supplementary file 2**).

## Enhanced Reduced Representation Bisulfite Sequencing (ERRBS)

50 ng of high quality genomic DNA was prepared from fin tissue from 6-month-old adult zebrafish DNA was digested with MspI and bisulfite converted using the EZ DNA methylation kit (zymo) as in **Garrett-Bakelman et al. (2015)**. Bisulphite conversion rates (calculated using non-CpG methylation conversion rates) ranged from 99.6% to 99.7% for all samples (**Figure 2—figure supplement 2C**). Amplified libraries were sequenced on the Hiseq2500 platform using a minimum of single-read 51 cycles. ERRBS data were filtered for sequence adapters, limited to the first 29 bp of the read (**Boyle et al., 2012**), and mapped to the zebrafish genome (danRer7) using BSmap (v 2.90) (**Xi and Li, 2009**). Other than limiting to the first 29 bp all other BSmap parameters were the defaults. Methylation scores were calculated as the number of unconverted reads divided by the number of total reads at each CpG site. DMRs were called as described in **Park and Wu (2016)**. DMRs with at least a 0.2 change in methylation were determined using DSS (delta = 0.2, p.threshold = 0.01). CallDMR function in DSS was used with default parameters except for p.threshold and delta as specified. Sat1 sequences are deficient in MspI sites, and are therefore not included in ERRBS data.

## RNA expression analysis

For qRT-PCR, total RNA was isolated using Trizol (Invitrogen) and precipitated with isopropanol. RNA used for assaying expression of repeat sequences subsequently was treated with DNase using

TURBO DNA-free Kit (Ambion) prior to analyses. RNA was converted to cDNA using GoScript Reverse Transcriptase Kit (Promega) and Real Time PCR was performed using an Applied Biosystems 7500 PCR Machine. Analysis was performed using the  $2^{-\Delta\Delta C_t}$  method, with relative mRNA levels of all transcripts normalized to  $\beta$ -actin1 or 18S. All primer sequences are listed in **Supplementary file 2**.

For Northern blot analysis, total RNA was extracted with using Trizol (Invitrogen). 2  $\mu$ g of RNA was subjected to electrophoresis on 1% agarose gel and transferred to Amersham Hybond-N+ membrane (GE Healthcare). The membrane was probed with  $^{32}$ P-dCTP radiolabeled Sat1 DNA probe at 42°C. Hybridization signals were imaged and analyzed using a Typhoon phosphorimager (GE Life Sciences).

TAG-aided sense/antisense transcript detection (TASA-TD) strand-specific PCR was performed as described by (Henke et al., 2015). Oligos used are listed in **Supplemental file 3**.

## Transcriptome sequencing

After RiboGreen quantification and quality control by Agilent BioAnalyzer, 500 ng of total RNA underwent polyA selection and TruSeq library preparation according to instructions provided by Illumina (TruSeq Stranded mRNA LT Kit), with 8 cycles of PCR. Samples were barcoded and run on a HiSeq 2500 High Output in a 50 bp/50 bp paired end run, using the TruSeq SBS v4 Kit (Illumina). An average of 45.3 million paired reads was generated per sample. The percent of mRNA bases averaged 62.8%.

For single-mutant RNA-seq analysis presented in **Figure 3**, reads were mapped to the Zebrafish genome (danRer7) using the rnaStar aligner v2.5.0a (Dobin et al., 2013). We used the two-pass mapping method outlined in Engström et al. (2013). The first mapping pass used a list of known annotated junctions from Ensemble. Novel junctions found in the first pass were then added to the known junctions and a second mapping pass was done (on the second pass the RemoveNoncanonical flag was used). Expression counts (counts per million, cpm) were computed from the mapped reads using HTSeq v0.5.3 (Anders et al., 2015) and Ensemble D. rerio v79 gene annotations. Normalization and differential expression was performed using DESeq (Anders and Huber, 2010).

For RNA-seq analysis presented in **Figure 6**, raw RNA-seq FASTQ reads were trimmed for adapters and preprocessed to remove low-quality reads using Trimmomatic v0.33 (arguments: LEADING:3 TRAILING:3 MINLEN:36) (Bolger et al., 2014) prior to mapping to the *Danio rerio* GRCz10 reference genome assembly. Reads were mapped using TopHat v2.1.1 (Kim et al., 2013) supplied with a reference General Features File (GFF) to the *Danio rerio* GRCz10 reference genome assembly, and with the following arguments: -i 10 -l 5000 -library-type fr-firststrand. Gene expression was estimated using Cuffquant (a tool from Cufflinks v2.2.1), with following arguments -library-type fr-firststrand. Expression level were normalized in FPKM units by Cuffnorm (a tool from Cufflinks v2.2.1), with following arguments -library-type fr-firststrand.

## 5-aza-dC treatment

Zebrafish embryos were treated with 5-aza-dC (Sigma-Aldrich) to the final concentration of 25  $\mu$ M or 50  $\mu$ M within the first 2 hr post fertilization, when zebrafish are sensitive to 5-aza-dC treatments as described in Martin et al. (1999). At 24hpf, total RNA was collected for expression analysis. At 24hpf, genomic DNA was also collected and digested with methylation sensitive enzyme, HpaII, to test for global DNA hypomethylation.

## RNA synthesis and injections

Sat1 RNA and control RNAs were in vitro transcribed using Riboprobe in vitro transcription systems (Promega). Oligos to amplify the DNA template for in vitro transcription are included in **Supplementary file 3**. Sense and anti-sense transcripts were transcribed in vitro using the T3 and T7 RNA polymerases respectively. RNA was purified illustra MicroSpin G-50 Columns (GE Healthcare) and 50 ng of sense and antisense RNA was co-injected into zebrafish embryos at the 1 cell stage.

## Statistical analysis

The Student unpaired 2-tailed t-test was used for statistical analysis unless specified otherwise. Statistical analysis was performed using GraphPad PRISM software.

## Accession number

All ERRBS and RNA-Seq data reported in this paper have been deposited in GEO under the accession GSE116360.

## Acknowledgements

This research was supported by a grant from the National Institutes of Health (R01GM110092) to MGG. We thank the Goll laboratory for helpful discussions and critical reading of the manuscript and Kellee Siegfried-Harris (UMass, Boston) for advice on gonadal sections. We acknowledge the use of the Integrated Genomics Operation Core of MSKCC, funded by the NCI Cancer Center Support Grant (CCSG, P30 CA08748), Cycle for Survival, and the Marie-Josée and Henry R Kravis Center for Molecular Oncology. We also acknowledge the use of the Bioinformatics core of MSKCC for support with sequence analysis. ERRBS was performed in the Weill Cornell Medicine Epigenomics Core Facility.

---

## Additional information

### Funding

Funder	Grant reference number	Author
National Institutes of Health	R01GM110092	Mary Goll

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Srivarsha Rajshekar, Conceptualization, Resources, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Jun Yao, Paige K Arnold, Sara G Payne, Yinwen Zhang, Teresa V Bowman, Formal analysis, Investigation; Robert J Schmitz, Formal analysis, Investigation, Writing—review and editing; John R Edwards, Formal analysis, Visualization, Writing—original draft, Writing—review and editing; Mary Goll, Conceptualization, Resources, Supervision, Funding acquisition, Validation, Methodology, Project administration, Writing—review and editing

### Author ORCIDs

Srivarsha Rajshekar  <https://orcid.org/0000-0002-5224-5531>

Sara G Payne  <http://orcid.org/0000-0002-3572-9112>

Mary Goll  <http://orcid.org/0000-0001-5003-6958>

### Ethics

Animal experimentation: This study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All of the animals were handled according to approved institutional animal care and use committee (IACUC) protocols of Memorial Sloan Kettering Cancer Center (MSKCC- 10-08-014) and the University of Georgia (UGA A2017 05-021). The protocol was approved by the Committee on the Ethics of Animal Experiments of MSKCC and UGA.

### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.39658.032>

Author response <https://doi.org/10.7554/eLife.39658.033>

## Additional files

### Supplementary files

- Supplementary file 1. List of mutant alleles.  
DOI: <https://doi.org/10.7554/eLife.39658.024>
- Supplementary file 2. List of primers.  
DOI: <https://doi.org/10.7554/eLife.39658.025>
- Supplementary file 3. List of Oligos (5'—3').  
DOI: <https://doi.org/10.7554/eLife.39658.026>
- Supplementary file 4. List of DMRS.  
DOI: <https://doi.org/10.7554/eLife.39658.027>
- Transparent reporting form  
DOI: <https://doi.org/10.7554/eLife.39658.028>

### Data availability

Sequencing data have been deposited in GEO under accession code GSE116360. All data generated or analyzed during this study are included in the manuscript and supporting files.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Rajshekar S, Edwards JR, Goll MG	2018	Pericentromeric hypomethylation elicits an interferon response in an animal model of ICF syndrome	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116360">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116360</a>	NCBI Gene Expression Omnibus, GSE116360

## References

- Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* **11**:R106. DOI: <https://doi.org/10.1186/gb-2010-11-10-r106>, PMID: 20979621
- Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**:166–169. DOI: <https://doi.org/10.1093/bioinformatics/btu638>, PMID: 25260700
- Anderson RM, Bosch JA, Goll MG, Hesselson D, Dong PD, Shin D, Chi NC, Shin CH, Schlegel A, Halpern M, Stainier DY. 2009. Loss of Dnmt1 catalytic activity reveals multiple roles for DNA methylation during pancreas development and regeneration. *Developmental Biology* **334**:213–223. DOI: <https://doi.org/10.1016/j.ydbio.2009.07.017>, PMID: 19631206
- Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG, Tan W, Penheiter SG, Ma AC, Leung AY, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC. 2012. In vivo genome editing using a high-efficiency TALEN system. *Nature* **491**:114–118. DOI: <https://doi.org/10.1038/nature11537>, PMID: 23000899
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114–2120. DOI: <https://doi.org/10.1093/bioinformatics/btu170>, PMID: 24695404
- Boyle P, Clement K, Gu H, Smith ZD, Ziller M, Fostel JL, Holmes L, Meldrim J, Kelley F, Gnirke A, Meissner A. 2012. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biology* **13**:R92. DOI: <https://doi.org/10.1186/gb-2012-13-10-r92>, PMID: 23034176
- Burger A, Lindsay H, Felker A, Hess C, Anders C, Chiavacci E, Zaugg J, Weber LM, Catena R, Jinek M, Robinson MD, Mosimann C. 2016. Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**:2025–2037. DOI: <https://doi.org/10.1242/dev.134809>, PMID: 27130213
- Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF. 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* **39**:e82. DOI: <https://doi.org/10.1093/nar/gkr218>, PMID: 21493687
- Cheon H, Borden EC, Stark GR. 2014. Interferons and their stimulated genes in the tumor microenvironment. *Seminars in Oncology* **41**:156–173. DOI: <https://doi.org/10.1053/j.seminoncol.2014.02.002>, PMID: 24787290
- Chernyavskaya Y, Mudbhary R, Zhang C, Tokarz D, Jacob V, Gopinath S, Sun X, Wang S, Magnani E, Madakashira BP, Yoder JA, Hoshida Y, Sadler KC. 2017. Loss of DNA methylation in zebrafish embryos activates retrotransposons to trigger antiviral signaling. *Development* **144**:2925–2939. DOI: <https://doi.org/10.1242/dev.147629>, PMID: 28698226
- Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, Hein A, Rote NS, Cope LM, Snyder A, Makarov V, Budhu S, Buhu S, Slamon DJ, Wolchok JD, Pardoll DM, Beckmann MW, Zahnow CA, Merghoub T, Mergoub T, et al. 2015. Inhibiting dna methylation causes an interferon response in cancer via dsrna including endogenous retroviruses. *Cell* **162**:974–986. DOI: <https://doi.org/10.1016/j.cell.2015.07.011>, PMID: 26317466

- Crowl JT**, Gray EE, Pestal K, Volkman HE, Stetson DB. 2017. Intracellular nucleic acid detection in autoimmunity. *Annual Review of Immunology* **35**:313–336. DOI: <https://doi.org/10.1146/annurev-immunol-051116-052331>, PMID: 28142323
- de Greef JC**, Wang J, Balog J, den Dunnen JT, Frants RR, Straasheijm KR, Aytekin C, van der Burg M, Duprez L, Ferster A, Gennery AR, Gimelli G, Reisli I, Schuetz C, Schulz A, Smeets D, Sznajder Y, Wijmenga C, van Eggermond MC, van Ostaïjen-Ten Dam MM, et al. 2011. Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *The American Journal of Human Genetics* **88**:796–804. DOI: <https://doi.org/10.1016/j.ajhg.2011.04.018>, PMID: 21596365
- Dobin A**, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21. DOI: <https://doi.org/10.1093/bioinformatics/bts635>, PMID: 23104886
- Doyle EL**, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ. 2012. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Research* **40**:W117–W122. DOI: <https://doi.org/10.1093/nar/gks608>, PMID: 22693217
- Ehrlich M**. 2003. The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clinical Immunology* **109**:17–28. DOI: [https://doi.org/10.1016/S1521-6616\(03\)00201-8](https://doi.org/10.1016/S1521-6616(03)00201-8), PMID: 14585272
- Ehrlich M**, Sanchez C, Shao C, Nishiyama R, Kehrl J, Kuick R, Kubota T, Hanash SM. 2008. ICF, an immunodeficiency syndrome: DNA methyltransferase 3B involvement, chromosome anomalies, and gene dysregulation. *Autoimmunity* **41**:253–271. DOI: <https://doi.org/10.1080/08916930802024202>, PMID: 18432406
- Engström PG**, Steijger T, Sipos B, Grant GR, Kahles A, Rättsch G, Goldman N, Hubbard TJ, Harrow J, Guigó R, Bertone P. RGASP Consortium. 2013. Systematic evaluation of spliced alignment programs for RNA-seq data. *Nature Methods* **10**:1185–1191. DOI: <https://doi.org/10.1038/nmeth.2722>, PMID: 24185836
- Enukashvily NI**, Donev R, Waisertreiger IS, Podgornaya OI. 2007. Human chromosome 1 satellite 3 DNA is decondensed, demethylated and transcribed in senescent cells and in A431 epithelial carcinoma cells. *Cytogenetic and Genome Research* **118**:42–54. DOI: <https://doi.org/10.1159/000106440>, PMID: 17901699
- Fanelli M**, Caprodossi S, Ricci-Vitiani L, Porcellini A, Tomassoni-Ardori F, Amatori S, Andreoni F, Magnani M, De Maria R, Santoni A, Minucci S, Pelicci PG. 2008. Loss of pericentromeric DNA methylation pattern in human glioblastoma is associated with altered DNA methyltransferases expression and involves the stem cell compartment. *Oncogene* **27**:358–365. DOI: <https://doi.org/10.1038/sj.onc.1210642>, PMID: 17653095
- Gagnon JA**, Valen E, Thyme SB, Huang P, Akhmetova L, Ahkmetova L, Pauli A, Montague TG, Zimmerman S, Richter C, Schier AF. 2014. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLOS ONE* **9**:e98186. DOI: <https://doi.org/10.1371/journal.pone.0098186>, PMID: 24873830
- Garrett-Bakelman FE**, Sheridan CK, Kacmarczyk TJ, Ishii J, Betel D, Alonso A, Mason CE, Figueroa ME, Melnick AM. 2015. Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. *Journal of Visualized Experiments*:e52246. DOI: <https://doi.org/10.3791/52246>, PMID: 25742437
- Geiman TM**, Tessarollo L, Anver MR, Kopp JB, Ward JM, Muegge K. 2001. Lsh, a SNF2 family member, is required for normal murine development. *Biochimica Et Biophysica Acta (BBA) - General Subjects* **1526**:211–220. DOI: [https://doi.org/10.1016/S0304-4165\(01\)00129-5](https://doi.org/10.1016/S0304-4165(01)00129-5)
- Goll MG**, Bestor TH. 2005. Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry* **74**:481–514. DOI: <https://doi.org/10.1146/annurev.biochem.74.010904.153721>, PMID: 15952895
- Grant CE**, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**:1017–1018. DOI: <https://doi.org/10.1093/bioinformatics/btr064>, PMID: 21330290
- Härtlova A**, Erttmann SF, Raffi FA, Schmalz AM, Resch U, Anugula S, Lienenklaus S, Nilsson LM, Kröger A, Nilsson JA, Ek T, Weiss S, Gekara NO. 2015. DNA damage primes the type I interferon system via the cytosolic DNA sensor STING to promote anti-microbial innate immunity. *Immunity* **42**:332–343. DOI: <https://doi.org/10.1016/j.immuni.2015.01.012>, PMID: 25692705
- Henke C**, Strissel PL, Schubert MT, Mitchell M, Stolt CC, Faschingbauer F, Beckmann MW, Strick R. 2015. Selective expression of sense and antisense transcripts of the sushi-ichi-related retrotransposon-derived family during mouse placentogenesis. *Retrovirology* **12**:9. DOI: <https://doi.org/10.1186/s12977-015-0138-8>, PMID: 2588968
- Jao LE**, Wente SR, Chen W. 2013. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *PNAS* **110**:13904–13909. DOI: <https://doi.org/10.1073/pnas.1308335110>, PMID: 23918387
- Jenness C**, Giunta S, Müller MM, Kimura H, Muir TW, Funabiki H. 2018. HELLS and CDCA7 comprise a bipartite nucleosome remodeling complex defective in ICF syndrome. *PNAS* **115**:E876–E885. DOI: <https://doi.org/10.1073/pnas.1717509115>, PMID: 29339483
- Kim D**, Perteau G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**:R36. DOI: <https://doi.org/10.1186/gb-2013-14-4-r36>, PMID: 23618408
- Labun K**, Montague TG, Gagnon JA, Thyme SB, Valen E. 2016. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Research* **44**:W272–W276. DOI: <https://doi.org/10.1093/nar/gkw398>, PMID: 27185894
- Lei H**, Oh SP, Okano M, Jüttermann R, Goss KA, Jaenisch R, Li E. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**:3195–3205. PMID: 8898232
- Leonova KI**, Brodsky L, Lipchick B, Pal M, Novototskaya L, Chenchik AA, Sen GC, Komarova EA, Gudkov AV. 2013. p53 cooperates with DNA methylation and a suicidal interferon response to maintain epigenetic silencing

- of repeats and noncoding RNAs. *PNAS* **110**:E89–E98. DOI: <https://doi.org/10.1073/pnas.1216922110>, PMID: 23236145
- Li C, Lan Y, Schwartz-Orbach L, Korol E, Tahiliani M, Evans T, Goll MG. 2015. Overlapping requirements for tet2 and tet3 in normal development and hematopoietic stem cell emergence. *Cell Reports* **12**:1133–1143. DOI: <https://doi.org/10.1016/j.celrep.2015.07.025>, PMID: 26257178
- Mankan AK, Schmidt T, Chauhan D, Goldeck M, Höning K, Gaidt M, Kubarenko AV, Andreeva L, Hopfner KP, Hornung V. 2014. Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. *The EMBO Journal* **33**:2937–2946. DOI: <https://doi.org/10.15252/embj.201488726>, PMID: 25425575
- Martin CC, Laforest L, Akimenko MA, Ekker M. 1999. A role for DNA methylation in gastrulation and somite patterning. *Developmental Biology* **206**:189–205. DOI: <https://doi.org/10.1006/dbio.1998.9105>, PMID: 9986732
- Nakagawa T, Kanai Y, Ushijima S, Kitamura T, Kakizoe T, Hirohashi S. 2005. DNA hypomethylation on pericentromeric satellite regions significantly correlates with loss of heterozygosity on chromosome 9 in urothelial carcinomas. *The Journal of Urology* **173**:243–246. DOI: <https://doi.org/10.1097/01.ju.0000141577.98902.49>, PMID: 15592089
- Narayan A, Ji W, Zhang XY, Marrogi A, Graff JR, Baylin SB, Ehrlich M. 1998. Hypomethylation of pericentromeric DNA in breast adenocarcinomas. *International Journal of Cancer* **77**:833–838. DOI: [https://doi.org/10.1002/\(SICI\)1097-0215\(19980911\)77:6<833::AID-IJC6>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-0215(19980911)77:6<833::AID-IJC6>3.0.CO;2-V), PMID: 9714050
- Parichy DM, Elizondo MR, Mills MG, Gordon TN, Engeszer RE. 2009. Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Developmental Dynamics* **238**:2975–3015. DOI: <https://doi.org/10.1002/dvdy.22113>, PMID: 19891001
- Park Y, Wu H. 2016. Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics* **32**:1446–1453. DOI: <https://doi.org/10.1093/bioinformatics/btw026>, PMID: 26819470
- Phillips RB, Reed KM. 2000. Localization of repetitive DNAs to zebrafish (*Danio rerio*) chromosomes by fluorescence in situ hybridization (FISH). *Chromosome Research* **8**:27–35. DOI: <https://doi.org/10.1023/A:1009271017998>, PMID: 10730586
- Qu GZ, Grundy PE, Narayan A, Ehrlich M. 1999. Frequent hypomethylation in Wilms tumors of pericentromeric DNA in chromosomes 1 and 16. *Cancer Genetics and Cytogenetics* **109**:34–39. DOI: [https://doi.org/10.1016/S0165-4608\(98\)00143-5](https://doi.org/10.1016/S0165-4608(98)00143-5), PMID: 9973957
- Rai K, Nadauld LD, Chidester S, Manos EJ, James SR, Karpf AR, Cairns BR, Jones DA. 2006. Zebra fish Dnmt1 and Suv39h1 regulate organ-specific terminal differentiation during development. *Molecular and Cellular Biology* **26**:7077–7085. DOI: <https://doi.org/10.1128/MCB.00312-06>, PMID: 16980612
- Roulois D, Loo Yau H, Singhania R, Wang Y, Danesh A, Shen SY, Han H, Liang G, Jones PA, Pugh TJ, O'Brien C, De Carvalho DD. 2015. Dna-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell* **162**:961–973. DOI: <https://doi.org/10.1016/j.cell.2015.07.056>, PMID: 26317465
- Schneider WM, Chevillotte MD, Rice CM. 2014. Interferon-stimulated genes: a complex web of host defenses. *Annual Review of Immunology* **32**:513–545. DOI: <https://doi.org/10.1146/annurev-immunol-032713-120231>, PMID: 24555472
- Smith ZD, Meissner A. 2013. DNA methylation: roles in mammalian development. *Nature Reviews Genetics* **14**:204–220. DOI: <https://doi.org/10.1038/nrg3354>, PMID: 23400093
- Stancheva I, Meehan RR. 2000. Transient depletion of xDnmt1 leads to premature gene activation in *Xenopus* embryos. *Genes & Development* **14**:313–327. PMID: 10673503
- Suzuki T, Fujii M, Ayusawa D. 2002. Demethylation of classical satellite 2 and 3 DNA with chromosomal instability in senescent human fibroblasts. *Experimental Gerontology* **37**:1005–1014. DOI: [https://doi.org/10.1016/S0531-5565\(02\)00061-X](https://doi.org/10.1016/S0531-5565(02)00061-X), PMID: 12213551
- Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics* **9**:465–476. DOI: <https://doi.org/10.1038/nrg2341>, PMID: 18463664
- Tanne A, Muniz LR, Puzio-Kuter A, Leonova KI, Gudkov AV, Ting DT, Monasson R, Cocco S, Levine AJ, Bhardwaj N, Greenbaum BD. 2015. Distinguishing the immunostimulatory properties of noncoding RNAs expressed in cancer cells. *PNAS* **112**:15154–15159. DOI: <https://doi.org/10.1073/pnas.1517584112>, PMID: 26575629
- Tao Y, Xi S, Shan J, Maunakea A, Che A, Briones V, Lee EY, Geiman T, Huang J, Stephens R, Leighty RM, Zhao K, Muegge K. 2011. Lsh, chromatin remodeling family member, modulates genome-wide cytosine methylation patterns at nonrepeat sequences. *PNAS* **108**:5626–5631. DOI: <https://doi.org/10.1073/pnas.1017000108>, PMID: 21427231
- Thijssen PE, Ito Y, Grillo G, Wang J, Velasco G, Nitta H, Unoki M, Yoshihara M, Suyama M, Sun Y, Lemmers RJ, de Greef JC, Gennery A, Picco P, Kloeckener-Gruissem B, Güngör T, Reisli I, Picard C, Kebaili K, Roquelaure B, et al. 2015. Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome. *Nature Communications* **6**:7870. DOI: <https://doi.org/10.1038/ncomms8870>, PMID: 26216346
- Thompson JJ, Kaur R, Sosa CP, Lee JH, Kashiwagi K, Zhou D, Robertson KD. 2018. ZBTB24 is a transcriptional regulator that coordinates with DNMT3B to control DNA methylation. *Nucleic Acids Research*:10034–10051. DOI: <https://doi.org/10.1093/nar/gky682>, PMID: 30085123
- Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S, Rivera MN, Bardeesy N, Maheswaran S, Haber DA. 2011. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331**:593–596. DOI: <https://doi.org/10.1126/science.1200801>, PMID: 21233348
- Trede NS, Langenau DM, Traver D, Look AT, Zon LI. 2004. The use of zebrafish to understand immunity. *Immunity* **20**:367–379. DOI: [https://doi.org/10.1016/S1074-7613\(04\)00084-6](https://doi.org/10.1016/S1074-7613(04)00084-6), PMID: 15084267

- Tsuda H**, Takarabe T, Kanai Y, Fukutomi T, Hirohashi S. 2002. Correlation of DNA hypomethylation at pericentromeric heterochromatin regions of chromosomes 16 and 1 with histological features and chromosomal abnormalities of human breast carcinomas. *The American Journal of Pathology* **161**:859–866. DOI: [https://doi.org/10.1016/S0002-9440\(10\)64246-0](https://doi.org/10.1016/S0002-9440(10)64246-0), PMID: 12213714
- Tuck-Muller CM**, Narayan A, Tsien F, Smeets DF, Sawyer J, Fiala ES, Sohn OS, Ehrlich M. 2000. DNA hypomethylation and unusual chromosome instability in cell lines from ICF syndrome patients. *Cytogenetic and Genome Research* **89**:121–128. DOI: <https://doi.org/10.1159/000015590>, PMID: 10894953
- Ueda Y**, Okano M, Williams C, Chen T, Georgopoulos K, Li E. 2006. Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. *Development* **133**:1183–1192. DOI: <https://doi.org/10.1242/dev.02293>, PMID: 16501171
- van der Vaart M**, van Soest JJ, Spaink HP, Meijer AH. 2013. Functional analysis of a zebrafish myd88 mutant identifies key transcriptional components of the innate immune system. *Disease Models & Mechanisms* **6**:841–854. DOI: <https://doi.org/10.1242/dmm.010843>, PMID: 23471913
- Velasco G**, Grillo G, Touleimat N, Ferry L, Ivkovic I, Ribierre F, Deleuze JF, Chantalat S, Picard C, Francastel C. 2018. Comparative methylome analysis of ICF patients identifies heterochromatin loci that require ZBTB24, CDCA7 and HELLS for their methylated state. *Human Molecular Genetics* **27**:2409–2424. DOI: <https://doi.org/10.1093/hmg/ddy130>, PMID: 29659838
- von Bernuth H**, Ravindran E, Du H, Fröhler S, Strehl K, Krämer N, Issa-Jahns L, Amulic B, Ninnemann O, Xiao MS, Eirich K, Kölsch U, Hauptmann K, John R, Schindler D, Wahn V, Chen W, Kaindl AM. 2014. Combined immunodeficiency develops with age in Immunodeficiency-centromeric instability-facial anomalies syndrome 2 (ICF2). *Orphanet Journal of Rare Diseases* **9**:116. DOI: <https://doi.org/10.1186/s13023-014-0116-6>, PMID: 25330735
- Weisenberger DJ**, Campan M, Long TI, Kim M, Woods C, Fiala E, Ehrlich M, Laird PW. 2005. Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Research* **33**:6823–6836. DOI: <https://doi.org/10.1093/nar/gki987>, PMID: 16326863
- Wu H**, Thijssen PE, de Klerk E, Vonk KK, Wang J, den Hamer B, Aytekin C, van der Maarel SM, Daxinger L. 2016. Converging disease genes in ICF syndrome: ZBTB24 controls expression of CDCA7 in mammals. *Human Molecular Genetics* **25**:4041–4051. DOI: <https://doi.org/10.1093/hmg/ddw243>, PMID: 27466202
- Xi Y**, Li W. 2009. BSMAP: whole genome bisulfite sequence MAPPING program. *BMC Bioinformatics* **10**:232. DOI: <https://doi.org/10.1186/1471-2105-10-232>, PMID: 19635165
- Xu GL**, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, Viegas-Péquignot E. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**:187–191. DOI: <https://doi.org/10.1038/46052>, PMID: 10647011