A novel role for mutant mRNA degradation in triggering transcriptional adaptation to mutations

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

Vorgelegt beim Fachbereich 15
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

Von
Mohamed A. El-Brolosy
aus Kairo, Ägypten

Frankfurt 2020
(D30)

vom Fachbereich Biowissenschaften (FB15) der Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan: **Prof. Dr. Sven Klimpel**

Gutachter: **Prof. Dr. Didier Y. R. Stainier**

Prof. Dr. Erin M. Schuman

Datum der Disputation:

Reviewers

Prof. Dr. Didier Y. R. Stainier, Ph.D.
Department of Developmental Genetics
Max Planck Institute for Heart and Lung Research
Bad Nauheim, Germany

&

Prof. Dr. Erin M. Schuman, Ph.D.
Max Planck Institute for Brain Research
Frankfurt am Main, Germany

"The important thing is not to stop questioning.
Curiosity has its own reason for existing."
Albert Einstein

Table of Contents

Table of contents

1. Abstract

Robustness to mutations promotes organisms' well-being and fitness. The increasing number of mutants in various model organisms, and humans, showing no obvious phenotype (Bouche and Bouchez, 2001; Chen et al., 2016b; Giaever et al., 2002; Kok et al., 2015) has renewed interest into how organisms adapt to gene loss. In the presence of deleterious mutations, genetic compensation by transcriptional upregulation of related gene(s) (also known as transcriptional adaptation) has been reported in numerous systems (El-Brolosy and Stainier, 2017; Rossi et al., 2015; Tondeleir et al., 2012); however, the molecular mechanisms underlying this response remained unclear. To investigate this phenomenon, I develop and study multiple models of transcriptional adaptation in zebrafish and mouse cell lines. I first show that transcriptional adaptation is not caused by loss of protein function, indicating that the trigger lies upstream, and find that the response involves enhanced transcription of the related gene(s). Furthermore, I observe a correlation between levels of mutant mRNA degradation and upregulation of related genes. To investigate the role of mutant mRNA degradation in triggering the response, I generate mutant alleles that do not transcribe the mutated gene and find that they fail to induce a transcriptional response and display stronger phenotypes. Transcriptome analysis of alleles displaying mutant mRNA degradation revealed upregulation of a significant proportion of genes displaying sequence similarity with the mutated gene's mRNA, suggesting a model whereby mRNA degradation intermediates induce transcriptional adaptation via sequence similarity. Further mechanistic analyses suggested RNA-decay factors-dependent chromatin remodeling, and repression of antisense RNAs to be implicated in the response. These results identify a novel role for mutant mRNA degradation in buffering against mutations. Besides, they hold huge implications on understanding disease-causing mutations and shall help in designing mutations that lead to minimal transcriptional adaptation-induced compensation, facilitating studying gene function in model organisms.

2. List of Abbreviations

Abbreviation	Full name
mRNA	Messenger RNA
RBP	RNA binding proteins
DSB	Double strand break
diRNA	DSB-induced RNA
RNAa	RNA activation
dsRNA	Double-stranded RNA
PTC	Premature termination codon
NMD	Nonsense mediated decay
NGD	No-go decay
NSD	No-stop decay
ECJ	Exon-junction complex
IP	Immunoprecipitation
HRP	Horseradish peroxidase
PBS	Phosphated buffer saline
SSC	Saline-sodium citrate
TBE	Tris-borate-EDTA
PFA	Paraformaldehyde
BSA	Bovine serum albumin
EDTA	Ethylenediaminetetraacetic acid
SDS	Sodium dodecyl sulfate
DMSO	Dimethyl sulfoxide
PTU	1-Phenyl-2-thiourea
DAPI	4′,6-diamidino-2-phenylindole
NMDi14	Nonsense mediated decay inhibitor 14
TNF	Tumor necrosis factor

List of Abbreviations

dNTPs	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MEFs	Mouse embryonic fibroblasts
MKFs	Mouse kidney fibroblasts
mESCs	Mouse embryonic stem cells
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
gRNA	Guide RNA
NGS	Next-generation sequencing
ATAC-seq	Assay for Transposase-Accessible Chromatin by sequencing
WT	Wild type
$Puro_r$	Puromycin resistance gene
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRi	CRISPR-interference
siRNA	Small interfering RNA
miRNA	Micro RNA
Scr	Scrambled
PCR	Polymerase chain reaction
HRMA	High resolution melt analysis
qPCR	Quantitative PCR
mRNA	Messenger RNA
TSS	Transcription start site
PAM	Protospacer adjacent motif
N/A	Not applicable

List of Abbreviations

UV	Ultra-violet
hpf	Hours post fertilization
dpf	Days post fertilization
cDNA	Complementary DNA
HEK	Human embryonic kidney cells
lncRNA	Long non-coding RNA
LoF	Loss of function
PTV	Protein truncating variant

Table 1. List of abbreviations used in this dissertation

3. Introduction

Parts of this chapter have been published as a review article in the journal PLOS Genetics (El-Brolosy and Stainier, PLOS Genetics, 2017; 13(7):e1006780).

3.1. Genetic robustness

The development of a single-cell zygote to a complex organism with different cell types is a fascinating process that has been optimized over millions of years of evolution. To ensure similar developmental outcomes despite minor differences in genetic makeup or environmental conditions, organisms have evolved multiple buffering systems to ensure robustness; a process termed as canalization (Mather, 1953; Waddington, 1959). Our cells experience tens of thousands of DNA lesions per day (Lindahl and Barnes, 2000), and while most of such DNA lesions are repaired by the DNA repair machinery in our cells, some mis-repaired lesions can lead to mutations that can have cause devastating effects. Fortunately, however, not all mutations lead to disease owing to the robustness of our genome. A 2012 study analyzed 185 human genomes from the 1000 genomes project and suggested that each individual carries around 100 heterozygous and 20 homozygous mutations in protein-coding genes (MacArthur et al., 2012) and other later studies identified several loss-of-function mutations in healthy individuals (Narasimhan et al., 2016; Sulem et al., 2015), including in previously reported disease-associated genes (Chen et al., 2016b). Following the recent advancement in DNA targeting technologies such as zinc-finger nucleases (Durai et al., 2005), TALENs (Cermak et al., 2011; Doyle et al., 2013), CRISPR/Cas9 (Gagnon et al., 2014; Ran et al., 2013; Vejnar et al., 2016) and other technologies, scientists have gained the ability to better understand gene function by generation of mutant animals and studying gene function through analyzing the effect of gene loss on the animal. However, scientists have observed lack of an obvious phenotype in several engineered mutant animals, thereby hindering our understanding of gene function. For example, 80% of the yeast genome was reported to be not essential for growth (Giaever et al., 2002), and a study on *C. elegans* showed that 96% of induced mutations go unrecognized in the lab due to their minimal effect on the worm's fitness (Davies et al., 1999). In addition, lack of phenotypes was reported for many gene knock-outs in mice (White et al., 2013), Arabidopsis (Bouche and Bouchez, 2001), and zebrafish (Kok et al., 2015).

Dosage compensation in fruit flies was one of the first examples reported for genetic robustness. Male flies increase display an increase in transcription from their single X chromosome by twofold to achieve similar gene expression levels as their female counterparts having two X chromosomes (Mukherjee and Beermann, 1965; Muller, 1932). Females in mammals, on the contrary, through establishing a heterochromatic environment, inactivates one of their X chromosomes leading to similar gene expression levels as the male counterparts (Barr and Bertram, 1949; Heard and Disteche, 2006; Lyon, 1961). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

3.1.1. Rewiring of genetic networks

Several mechanisms were proposed to be underlying the phenomenon of genetic robustness. Rewiring of cellular networks, such as metabolic and genetic (transcriptional) ones, in response to loss of a gene, has been reported to confer robustness in several model organisms (Barabasi and Oltvai, 2004; Davidson and Levin, 2005). Mutations can disrupt tightly regulated networks leading to changes in the transcription of other genes within the network which can contribute to maintaining cellular fitness. For example, mutations in the ribosomal gene *Rpl22* in mice cause no significant defects in translation, due to the increased expression of its *Rpl22l1* (a paralogue), whose expression is inhibited by RPL22 under wild-type conditions (O'Leary et al., 2013). Furthermore, Beta-Catenin mutations in F9 teratocarcinoma cell lines inactivates the destruction complex and thereby increasing Gamma-Catenin stability (Fukunaga et al., 2005). In addition, upregulation of p52 (a NFKB protein) in *Nfkb1* mutant MEFs (Hoffmann et al., 2003), was reported to be due to increased processing of the precursor p100 to p52 (Basak et al., 2008).

Rewiring of genetic networks can also explain situations were acute loss of a protein leads to stronger phenotypes than chronic loss. For example, acute knockdown of the RAC1/CDC42 guanine nucleotide exchange factor *DOCK6* in HeLa cells leads to collapsing of the cytoskeleton and disruption of cell attachment and spread on glass surface. Knockout of *DOCK6*, or prolonged knockdown, displayed no overt phenotype. Prolonged loss of *DOCK6* was found to increase retention of the transcription factor MTRF-A in the cytoplasm, thereby reducing its ability to transcribe the ubiquitin-like modifier ISG15. Reduced ISG15 leads to decreased ISGylation of the GTPase-activating-like protein IQGAP1. IQGAP1 functions to stabilize RAC1

and CDC42 and reduced IQGAP1 ISGylation allows for stronger interaction with RAC/CDC42, leading to increased activity of the GTPases RAC1 and CDC42, thereby functionally compensating for the chronic loss of *DOCK6* (Cerikan et al., 2016). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

3.1.2. Acquiring adaptive mutations

Moreover, in response to a mutation, rapidly proliferating organisms such as yeast may accumulate adaptive mutations in other genes controlling the same affected activity or pathway, thereby ameliorating the potential effect of the first mutation (Chen et al., 2016a; Teng et al., 2013). For example, analyzing the yeast knockout collection has revealed that most of the knockout strains gain secondary activating or inactivating mutations in genes affecting nutrient responses or heat stress-induced cell death allowing normal growth of such knockout strains (Teng et al., 2013). Moreover, secondary acquired mutations in tumor suppressor homologs were also identified in several human tumors (Teng et al., 2013). Furthermore, a study in yeast revealed that some cells tend to be resilient to mutations in essential genes owing to acquiring secondary mutations. For example, mutations in the essential gene *ADE13*, which codes for the enzyme adenylosuccinate lyase (ADSL), mutations in which cause mental retardation and seizures in humans, can be rescued by different mutations upstream of ADSL in yeast and *C. elegans* proposing a loss-of-function therapeutic strategy for affected human individuals (Chen et al., 2016a). Similarly, knockout of ERD2 in yeast, which recycles protein-associated vesicles between the Golgi and endoplasmic reticulum (ER), prevents cells' growth due to accumulation of proteins in the Golgi (Hardwick et al., 1992; Townsley et al., 1994). However, some mutant cells are resilient owing to an acquired secondary mutation in the *ERV29* gene, which codes for a protein that transports proteins from the ER to the Golgi, thus slowing down protein transport to the Golgi and preventing its accumulation (Chen et al., 2016a).

3.1.3. Genetic redundancies

Another mode of robustness arises from redundant genes whereby loss of a given gene can be compensated by other redundant genes with overlapping functions (Cadigan et al., 1994; Cohen et al., 1987; González-Gaitán et al., 1994; Hoffmann, 1991; Santamaria et al., 2007; von Koch et al., 1997; Wang et al., 1996) (reviewed in

(Tautz, 1992)). For example, in yeast knockout of any of the paralogous genes STV*1* and *VPH1* (two alternative subunits of the V-ATPase complex that is involved in organelles' acidification), leads to cellular re-localization of the intact subunit to take over the function of the mutated gene (Manolson et al., 1994). Similarly, in mice loss of nidogen-1, a major component of basement membranes, leads to expression of nidogen-2, whose expression is normally restricted to endothelial basement membranes, in the basement membranes of skeletal muscle sarcomeres without an increase in its expression (Murshed et al., 2000). Redundancies due to changes in protein interactions were also reported for mouse voltage-gated Ca^{2+} channels subunits; loss of the β4 isoform leads to increased association of the remaining three β subunits with the α1 subunit without an increase in their transcription (Burgess et al., 1999). Furthermore, the maintained robustness of the yeast nuclear-pore complex upon loss of one of its subunits was reported to be due to the ability of redundant subunits to establish new protein-protein interactions (Diss et al., 2013). Similarly, mice lacking all interphase cyclin-dependent kinases (Cdk2, Cdk3, Cdk4, and Cdk6) developed normally to mid-gestation owing to the ability of Cdk1 to interact with their corresponding cyclins and promote cell cycle progression (Santamaria et al., 2007).

3.1.4. Genetic plasticity

Another form of robustness arises from biological plasticity. Nonsense mutations are one of the most common mutations used to study gene function. Such mutations lead to the introduction of a premature termination codon (PTC) in a coding exon of a given mRNA leading to non-sense mediated degradation of the mutated transcript or the production of truncated proteins that are unable to perform the function. Such mutation can be introduced either through point mutations or insertion-deletion mutations (indels) that disrupt the coding sequence and cause a frameshift that introduces a PTC. A number of studies have reported numerous plasticity mechanisms that render such kind of mutations less severe (Anderson et al., 2017; Jagannathan and Bradley, 2016; Lalonde et al., 2017; Prykhozhij et al., 2017; Smits et al., 2019; Tuladhar et al., 2019). According to these studies, frameshift and nonsense mutations can lead to alternative splicing and skipping of exons containing the PTC, thereby producing an in-frame mRNA that can code for a functional protein. They also reported re-initiation of translation downstream of the PTC leading to production of N-terminal truncated proteins that can be functional for some genes.

Moreover, PTC-readthrough was also reported as another mechanism through which cells evade premature translation termination or nonsense-mediated decay (Jagannathan and Bradley, 2016).

3.2. Genetic compensation and the discrepancies between knockout and knockdown phenotypes

The previously mentioned modes of genetic robustness were all a result of the loss of a specific protein's function. More recently, several studies reported a different mode of genetic robustness that is not triggered due to loss of the protein function, but rather by a then yet to be identified upstream trigger (Hall et al., 2013; Rossi et al., 2015; Sztal et al., 2018; Tondeleir et al., 2012; Zhu et al., 2017), a phenomenon that was termed as genetic compensation, or also referred to as transcriptional adaptation (Table 2). The recent advances in gene targeting technologies have revealed phenotypic discrepancies between knockout (mutant) and knockdown (via antisense methodologies) models in a range of model organisms such as Drosophila (Yamamoto et al., 2014), Arabidopsis (Braun et al., 2008; Chen et al., 2014; Gao et al., 2015), zebrafish (Kok et al., 2015; Law and Sargent, 2014; Sztal et al., 2018; Zhu et al., 2017), mouse (Daude et al., 2012; De Souza et al., 2006; McJunkin et al., 2011; Young et al., 2009), and human cultured cell lines (Evers et al., 2016; Karakas et al., 2007; Morgens et al., 2016). While toxicity or off-target effects of the knockdown methodologies, or genetic plasticity including the possibility of analyzing a hypomorphic mutant allele were provided as an explanation for some of the observed discrepancies (Baek et al., 2014; Olejniczak et al., 2010; Olejniczak et al., 2016; Robu et al., 2007) (reviewed in (Jackson and Linsley, 2010)), a pioneering study in zebrafish proposed genetic compensation responses through upregulation of related genes in knockout but not knockdown models as an underlying reason for the observed discrepancy (Rossi et al., 2015). While morpholino-mediated knockdown of *egfl7*, an endothelial extracellular-matrix (ECM) gene, leads to severe vascular developmental defects (Parker et al., 2004), *egfl7* mutants show no obvious vascular phenotype. Minimal or no vascular defects were observed upon injecting the *egfl7* MO into *egfl7* mutant embryos, suggesting that the phenotypic discrepancies are not due to MO off-target effects. Moreover, CRISPR interference (CRISPRi)-mediated knockdown of *egfl7* also led to vascular defects in the developing embryos. Through transcriptome and proteome analyses, the authors observed upregulation of another family of ECM genes, specifically Emilins, in *egfl7* mutants but not knockdowns.

Overexpression of *Emilin2* or *Emilin3* in *egfl7*-morpholino injected embryos partially rescued the circulation phenotype, suggesting that the upregulation of the *emilin* genes in *egfl7* mutants can explain the discrepancy between the mutant and knockdown phenotypes. The authors also observed upregulation of *vegfab* transcript levels in *vegfaa* mutants but not MO-injected embryos. Furthermore, they observed to increase in *vegfab* expression levels upon overexpression of a dominant-negative form of Vegfaa (which causes a phenotype similar to that of *vegfaa* mutants) in zebrafish embryos, suggesting that the trigger for such transcriptional adaptation responses is independent of the loss of protein function. Later, several other studies reported similar findings. For example, zebrafish *actc1b* mutants display mild muscular phenotypes, unlike the knockdowns, due to the upregulation of its paralogue *actc1a* (Sztal et al., 2018) and while MO-mediated knockdown of *nid1a* in zebrafish larvae leads to a short body length phenotype, *nid1a* mutants display normal body length due to the upregulation of other family members: *nid1b* and *nid2a* (Zhu et al., 2017).

Term	Definition
knockout	Genetic manipulations that aim to ablate gene function (Housden et al., 2017)
knockdown	Methodologies that reduce the amount of functional RNA or protein levels of specific genes through interfering with DNA, RNA, or proteins (Housden et al., 2017)
genetic compensation	Changes in transcript or protein levels that can lead to functional compensation for a loss of a given gene's function in response to a mutation (El-Brolosy and Stainier, 2017)
transcriptional adaptation	Changes in transcript levels as a consequence of a genetic mutation that is independent of the loss of protein function (El-Brolosy and Stainier, 2017)

Table 2. Glossary

Following a gene knockout, upregulation of related genes might be a direct consequence of the loss of protein function, for example through the loss of a negative feedback loop, which can thereby maintain genetic robustness as previously explained. Knockdown analyses may help differentiate whether an observed upregulation is due to loss of the protein function or transcriptional adaptation. For example, *RBL2* mutant human T lymphocytes display normal proliferation and immune function owing to the upregulation of RBL1, an upregulation which is also detected following knockdown of *RBL2* in human breast cancer cell lines (Jackson and Pereira-Smith, 2006; Mulligan et al., 1998). Similarly, knockout or knockdown of *HDAC1* leads to an upregulation of HDAC2 in several mouse and human cultured

cells and tissues, and vice versa (Hagelkruys et al., 2014; Jurkin et al., 2011; Lagger et al., 2002).

In contrast, lack of a compensatory upregulation response in knockdown and dominant-negative treated animals or cells, compared to their corresponding knockout models, implies that another form of compensation might be triggered independent of the loss of protein function, possibly due to the DNA lesion itself or presence of the mutant mRNA molecules. For example, while knockdown of any of the three Cyclin D genes was reported to reduce cellular proliferation in different cultured cell lines (Becker et al., 2010; Radulovich et al., 2010; Wang et al., 2012), knockout mice of any of the Cyclin D family members develop normally with minimal defects suggesting potential compensation by any of the other intact family members (Huard et al., 1999; Sicinska et al., 2003; Sicinski and Weinberg, 1997). Notably, double knockout mice, increase the expression of the intact Cyclin D gene and display minor phenotypes only in tissues that fail to induce the upregulation of the intact gene (Ciemerych et al., 2002). Similarly, *Cyclin D2* knockout mouse B-lymphocytes display no proliferation defects owing to the compensatory upregulation of *cyclin D3* (Lam et al., 2000). Furthermore, knockdown of TET1, an enzyme that catalyzes the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), in mESCs leads to a significant decrease in 5hmC levels and a loss of mESCs undifferentiated morphology; on the contrary, *Tet1* knockout embryonic stem cells display limited decrease in 5hmC levels and maintain their undifferentiated morphology (Dawlaty et al., 2011), suggesting a potential compensation by the paralogous enzyme, TET2, in knockout but not siRNA-treated cells (Freudenberg et al., 2012). Moreover, knockdown of the integrin co-activator gene *Fermt2* in mouse embryonic fibroblasts (MEFs) was reported to inhibit INTERLEUKIN 1 beta-mediated increase in focal adhesions (Rajshankar et al., 2012). However, *Fermt2* knockout cells are able to form focal adhesions owing to the de-novo expression of the paralogous gene *Fermt1* (Theodosiou et al., 2016). In addition, *Importinα5* knockdown in mESCs was reported to inhibit their differentiation to neuronal lineages (Yasuhara et al., 2007); however, *Importinα5* knockout mice exhibit normal brain development, potentially due to the increase in IMPORTINα4 expression levels (Shmidt et al., 2007). In another example, *Dystrophin* knockout mice were reported to not display a severe muscular dystrophy phenotype owing to the increased expression of several genes including the dystrophin-related protein UTROPHIN

(Deconinck et al., 1997; Law et al., 1994). Interestingly, however, *Dystrophin* knockdown mice do not display UTROPHIN upregulation (Ghahramani Seno et al., 2008). In addition, knockdown of *Tau* in cultured neural cells inhibits axonal elongation (Caceres and Kosik, 1990; Caceres et al., 1991). However, axonal elongation was not affected in cultured knockout *Tau* neurons, potentially due to the increased expression of microtubule-associated protein 1A (MAP1A) (Harada et al., 1994), an upregulation that was interestingly not detected following *Tau* knockdown in mouse cultured oligodendrocytes (Seiberlich et al., 2015).

Actb knockout mice display upregulation of many other *Actin* genes, such as *Actg1* and *Acta2* (Bunnell et al., 2011; Patrinostro et al., 2017; Tondeleir et al., 2012). Interestingly, restoring *Actb* expression in *Actb* knockout mouse embryonic fibroblasts does not lead to dampening *Actg1* upregulation levels, suggesting that this transcriptional adaptation response is triggered independently of loss of *Actb* protein function (Tondeleir et al., 2012). Furthermore, *Acta2* is upregulated in *Actg1* mutant mouse embryonic fibroblasts but not upon knockdown of *Actg1* (Patrinostro et al., 2017). Moreover, while knockdown of the centrosome protein AZI1 by siRNAs in mouse embryonic fibroblasts leads to a significant reduction in ciliogenesis, knockout *Azi1* MEFs display no ciliogenesis defects (Hall et al., 2013). Interestingly, *Azi1* knockout fibroblasts were resistant to the *Azi1* siRNA, suggesting that the knockdown phenotype is not due to siRNA off-target effects and that knockout MEFs display no phenotype due to induction of a genetic compensation response. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

Model organism	Gene	Mutant phenotype	Knockdown phenotype	Proposed compensating gene in mutants	References
Yeast	*Bem1*	No obvious phenotype	Cell polarity defects and reduced viability	N/A	(Jost and Weiner, 2015)
Zebrafish	*egfl7*	Minor or no vascular defects	Severe vascular defects	*emilin3a*	(Rossi et al., 2015)
	nid1a	No obvious	Short body length	*nid1a* and	(Zhu et al.,

		phenotype		*nid2a*	2017)
	actc1b	Mild muscular defects	Reduced skeletal muscle performance	*actc1a*	(Sztal et al., 2018)
Arabidopsis	*ABP1*	No obvious phenotype	Defects in leaf growth	N/A	(Braun et al., 2008; Chen et al., 2014; Gao et al., 2015)
Mouse	*Sprn*	*Sprn; Prnp* double knockout mice are viable.	Embryonic lethality upon knockdown of *Sprn* in *Prnp* knockout mice.	N/A	(Daude et al., 2012; Young et al., 2009)
	Azi1	No obvious defects in cultured embryonic fibroblasts	Reduced ciliogenesis in cultured embryonic fibroblasts	N/A	(Hall et al., 2013)
	Tet1	No obvious phenotype in cultured mESCs	Loss of mESCs undifferentiated morphology	*Tet2*	(Dawlaty et al., 2011; Freudenberg et al., 2012)
	Aqp4	No obvious defects in cultured astrocytes	F-Actin cytoskeleton rearrangement in cultured astrocytes	N/A	(Ma et al., 1997; Manley et al., 2000; Nicchia et al., 2005; Papadopoulos et al., 2004)
	Ppara	Mutant mice do not develop	Knockdown mice develop	N/A	(De Souza

		hypoglycemia or hypertriglyceridemia under normal feeding conditions	hypoglycemia and hypertriglyceridemia under normal feeding conditions		et al., 2006)
	Prkn	No mitochondrial defects following treatment with acetaminophen	Decreased mitophagy in the liver following treatment with acetaminophen	N/A	(Williams et al., 2015)
Human (Breast cancer cell lines)	*MELK*	No obvious defects	Reduced cellular proliferation	N/A	(Hebbard et al., 2010; Lin et al., 2017; Lin et al., 2007; Speers et al., 2016; Wang et al., 2014)

Table 3. Discrepancies between knockout and knockdown phenotypes.

Table adapted and reprinted with permission from (El-Brolosy and Stainier, 2017).

3.3. Mechanisms that potentially underlie transcriptional adaptation

As previously mentioned, transcriptional adaptation involves upregulation of related genes (hereafter referred to as adapting genes) in response to mutations in a manner independent of the loss of protein function. An upregulation response is often inferred as independent of the loss of the protein function if it is not induced by knockdown or dominant-negative approaches or if restoring the mutant gene's expression doesn't dampen the upregulation response (Rossi et al., 2015; Tondeleir et al., 2012). Ruling out loss of the protein function as the trigger, one can consider two potential triggers to underlie transcriptional adaptation: (1) the DNA lesion itself or (2) consequences of the presence of mutant mRNA molecules (Figure 1). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

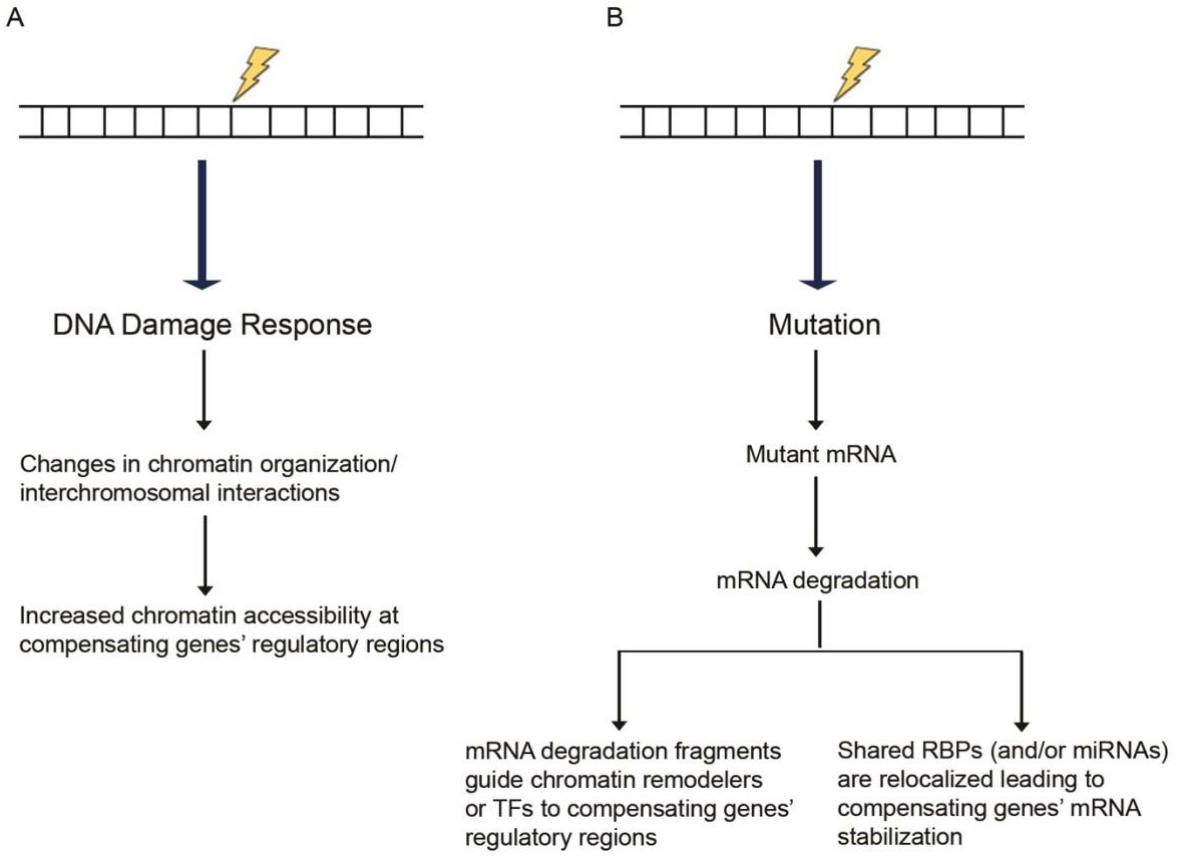


Figure1. Potential models of transcriptional adaptation.

(A) DNA lesions can induce a DNA damage response that leads to chromatin reorganization and modulates the accessibility of chromatin at the compensating/adapting genes' regulatory loci. **(B)** Mutations often lead to defective transcripts that are targeted to decay by the mRNA surveillance machinery. The resulting RNA decay intermediates (fragments) themselves may contribute to triggering the transcriptional adaptation response. Alternatively, as an indirect effect of the mutated gene's mRNA decay, stabilizing miRNAs or RBPs acting normally on both the mutated and the compensating/adapting genes' mRNAs, would be more available to stabilize the compensating genes' mRNAs. RBPs: RNA binding proteins; miRNAs: micro RNAs; TFs: transcription factors. Figure adapted and reprinted with permission from (El-Brolosy and Stainier, 2017).

3.3.1. DNA lesions as a potential trigger for transcriptional adaptation

Upon DNA damage, several chromatin remodelers and histone-modifying enzymes induce chromatin decondensation and reorganization ((Takahashi and Kaneko, 1985; Ziv et al., 2006), reviewed in (Downs et al., 2007)). It is thereby possible, that in response to mutations, global chromatin reorganization may lead to increased chromatin accessibility at the adapting genes' regulatory loci, leading to its increased

expression levels (Figure 1A). Dosage compensation in flies also involves chromatin reorganization, where a complex composed of the male-specific lethal (MSL) proteins and other proteins induce H4K16 acetylation on the X chromosome which leads to an increase in the accessibility of chromatin allowing for enhanced transcription (Stuckenholz et al., 1999). Furthermore, incomplete penetrance of the *skn-1* mutants gut phenotype in worms was attributed to the high variability in the expression levels of the compensating gene *end-1* (Bowerman et al., 1992; Raj et al., 2010). Notably, *end-1* variable expression was attributed to differences in the chromatin environment at *end-1* regulatory loci. Chromatin reorganization might also involve modification in DNA looping and nuclear architecture (Chambeyron and Bickmore, 2004). Several kinds of stress (such as temperature) may modulate inter- and intrachromosomal interactions ((Li et al., 2015), reviewed in (Wei et al., 2013). It is thereby possible that DNA damage-induced stress could lead to modifications in chromosomal interactions, such as intra-chromosomal interactions between the adapting genes' loci or even interchromosomal ones between the mutated gene and the adapting gene's regulatory loci, which can thereby modulate the adapting genes' expression.

DNA damage in HEK293T cells was reported to induce the expression of *GADD45A* leading to global DNA demethylation and thereby activation of previously-silenced methylated promoters (Barreto et al., 2007). It is thereby possible that such a response in response to mutations can lead to demethylation of the adapting genes' promoters and thereby contributing to transcriptional adaptation.

Small non-coding RNAs (ncRNAs) from regions spanning a DNA double-strand break (DSB), known as diRNAs (DSB-induced RNAs), were proposed to be essential for DNA repair possibly through recruiting proteins and chromatin remodelers to the DSB site (Francia et al., 2012; Wei et al., 2012) (reviewed in (d'Adda di Fagagna, 2014)). It is thereby possible that diRNAs may contribute to transcriptional adaptation through guiding positive chromatin remodelers or transcription factors through homology-mediated base-pairing to the adapting genes' regulatory loci and modulating their expression. This model of ncRNAs guiding chromatin remodelers is consistent with the role of the small-ncRNAs *roX1* and *roX2* in dosage compensation in *Drosophila* which guide the MSL complex assembly on the male X chromosome (Amrein and Axel, 1997; Franke and Baker, 1999; Meller et al., 1997) (reviewed in (Stuckenholz et al., 1999)).

To conclude, changes in the chromatin environment may play a role in transcriptional adaptation. It is likely that if such kind of responses are involved, they could be transmitted to the next generations through genomic imprinting via histone modification (Carr et al., 2007; Fournier et al., 2002; Yang et al., 2003) (reviewed in (McEwen and Ferguson-Smith, 2009)). Chromatin accessibility, chromosome capture, and epigenetic studies may help investigate such possibilities. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

3.3.2. Consequences of the presence of mutant mRNA molecules as a potential trigger for transcriptional adaptation

3.3.2.1. mutant mRNA degradation potential role in transcriptional adaptation

Mutations often lead to defective transcripts that are degraded by the mRNA surveillance machinery (see later chapter 3.4.) (Akimitsu, 2008; Harigaya and Parker, 2010; Isken and Maquat, 2007). A recent zebrafish study reported two different *mt2* mutant alleles (in the same exon) with different levels of phenotypic severity (Schuermann et al., 2015). Interestingly, the mutant allele displaying a milder phenotype exhibited higher levels of mutant mRNA degradation. Accordingly, MO-mediated knockdown of the NMD pathway in embryos with the milder phenotype led to a more severe phenotype. These data suggested that mutant mRNA degradation can trigger a genetic compensation response. One possibility would be that the mutant mRNA degradation intermediates may act as regulatory ncRNAs that can modulate the adapting genes' expression. Several recent studies have attempted to sequence RNA degradation intermediates and successfully identified a range of fragments that vary widely in size (Ibrahim et al., 2018; Ibrahim and Mourelatos, 2019; Kurosaki et al., 2018; Peach et al., 2015; Pelechano et al., 2016; Schmidt et al., 2015; Ueno et al., 2018; Valen et al., 2011). Moreover, several studies in the past decade have reported that mRNA degradation and gene expression are coupled processes ((Elkon et al., 2010; Haimovich et al., 2013; Hao and Baltimore, 2009; Rabani et al., 2011; Sun et al., 2012), reviewed in (Hartenian and Glaunsinger, 2019)). According to one model (Haimovich et al., 2013), following mutant mRNA degradation, decay factors can translocate back to the nucleus to bind near transcription start sites and promote transcription initiation and elongation rates, possibly through interacting with chromatin remodelers and histone modifiers as well (Berretta et al., 2008; Collins et al., 2007; Haimovich et al., 2013; Pinskaya et al.,

2009). It is thereby possible that decay intermediates can guide specific epigenetic modifiers to adapting genes' loci through homology-mediated base pairing in a fashion similar to that of long non-coding RNAs (reviewed in (Vance and Ponting, 2014)) (Figure 1B).

3.3.2.2. Antisense transcripts

Moreover, injection, or transfection, of short RNA fragments, ranging in size between 20 and 22 nucleotides, of a given mRNA was reported to increase the transcription levels of the corresponding gene (Ghanbarian et al., 2017; Wagner et al., 2008). Mechanistically, the authors provided evidence that such short RNA fragments act on antisense RNAs present at the corresponding genes locus, which in many cases act as negative regulators of gene expression (Faghihi and Wahlestedt, 2009; Modarresi et al., 2012). The resulting double-stranded RNA duplex formed between the fragment and the antisense RNA is then recognized by the RNAi machinery leading to reduced expression of antisense RNA and increased expression of the sense RNA. The authors further confirmed the requirement of the RNAi machinery as this response was dependent on Argonaute proteins. The human and mouse transcriptome includes several antisense transcripts are capable of forming pairs with complementary RNAs (Chen et al., 2004; Katayama et al., 2005; Kiyosawa et al., 2003; Yelin et al., 2003). Acting upon antisense transcripts is thereby one possibility through which RNA decay fragments can induce a transcriptional adaptation response.

3.3.2.3. RNA activation (RNAa)

The previous model is consistent with studies that identified transcriptional activation upon targeting of short double-stranded RNAs (dsRNA) to gene promoter regions or transcription start sites, a phenomenon termed as RNA activation (RNAa) (Hu et al., 2012; Janowski et al., 2007; Li et al., 2006; Portnoy et al., 2016; Schwartz et al., 2008; Turunen et al., 2009; Zhang et al., 2014). According to such model, the double-stranded RNA (dsRNA)-loaded AGO2 binds to the promoter region, or the nascent transcript, of a given gene recruiting a transcriptional activation complex which includes the helicase RHA and CTR9, a component of the PAF1 complex, that interacts with RNA polymerase II to enhance transcription initiation and elongation (Portnoy et al., 2016). The complex also recruits ubiquitin ligases that induce histone 2B monoubiquitination, a histone modification that increases transcription activation

through recruiting proteins that increases H3K4 di- and tri-methylation (Weake and Workman, 2008). Interestingly, other studies have reported that RNAa is induced through the action of the dsRNA on antisense RNAs present at the promoter regions of targeted genes (Chu et al., 2010; Matsui et al., 2010; Zhang et al., 2014). It is possible that RNA decay fragments can act in a similar fashion to that observed in RNAa, for example through binding to promoter region or the nascent transcript and other promoter-associated ncRNAs of the adapting gene.

3.3.2.4. Indirect effects of mutant mRNA degradation

In addition to their well-established gene silencing effects (Filipowicz et al., 2008), micro-RNAs (miRNAs), miRNAs were reported to be able to enhance gene expression through a number of different mechanisms. *miRNA-373*, for example, was reported to bind near *CDH1* and *CSDC2* promoter regions in human prostate cancer cell lines and increase their expression levels in a fashion similar to the previously explained RNAa (Place et al., 2008). Under starvation conditions, *miRNA10a* was observed to bind at the 5′ untranslated region of mRNAs coding for ribosomal protein and contribute to promoting their translation (Orom et al., 2008). miRNAs can have several target mRNAs (Jacobsen et al., 2013; Pasquinelli, 2012), and it is possible that following mutant mRNA degradation of a given gene, the miRNAs targeting the mutated gene will be more available to target other mRNAs or genes and can lead to their increased expression if they act in a similar fashion to the examples described above (Figure 1B).

RNA-binding proteins (RBPs) are key players in regulating gene expression (reviewed in (Glisovic et al., 2008)), one of their mechanisms is through regulating mRNA stability (Kuwano et al., 2009). RNA operons or RNA regulons is a term used to describe mRNAs coding for functionally related proteins that are co-regulated by similar RBPs (Gerber et al., 2004; Keene and Lager, 2005; Keene and Tenenbaum, 2002) (reviewed in (Keene, 2007)). If the mutant and the adapting genes are regulated by the same RNA binding proteins, it is possible that following mutant mRNA degradation (or mutation-induced changes in secondary structure of an mRNA), the RBPs co-regulating the mutant and adapting genes' mRNAs would become more available to stabilize the adapting genes' transcripts (Figure 1B). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

3.3.3. Other potential triggers for genetic compensation

Several other responses have been reported for different kinds of stress that leads to selective stabilization or increased expression of specific genes. For example, in response to heat shock, certain mRNAs were reported to be post-transcriptionally modified by N6-methylation of adenosines (m6A) or pseudouridylation which leads to stabilization of those messages and their increased translation (Schwartz et al., 2014; Zhou et al., 2015a) (reviewed in (Licht and Jantsch, 2016)). In yeast, starvation condition induces the phosphorylation of eIF2α which thereby which contributes in minimizing global translation events through reducing translation initiation rates by acting as a competitive inhibitor for the initiation factor eIF2B (Hinnebusch et al., 2007). This mechanism may allow, however, the increased translation of certain mRNAs under cellular stress conditions. For example, the yeast transcription factor gene *GCN4* has four upstream open reading frames (uORFs). Under normal growth conditions, the four uORFs are translated and minimal translation reinitiation happens at the GCN4 coding ORF. Under starvation conditions, the first upstream ORF of *GCN4* is efficiently translated; but due to the phosphorylation of eIF2α, translation reinitiation doesn't happen at the downstream uORFs and reinitiates only at the GCN4 coding ORF leading to increased GCN4 expression (Mueller and Hinnebusch, 1986). It is possible that specific gene mutations can induce a stress response that leads to activation of one of the aforementioned mechanisms to maintain robustness. However, to my knowledge, no examples currently exist on posttranscriptional modifications and uORF skipping as genetic compensation responses for mutations. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

3.4 The mRNA surveillance machinery

Mutations often lead to defective mRNAs that are directed for decay by the mRNA surveillance machinery (reviewed in (Akimitsu, 2008; Harigaya and Parker, 2010; Lykke-Andersen and Jensen, 2015)). The mRNA surveillance machinery is a very important conserved mechanism that prevents the cells from translating defective transcripts and thereby preventing the cell from producing toxic truncated proteins that can act in a dominant-negative or constitutively active fashion. It is also very important in regulating gene expression through controlling transcript levels

(Haimovich et al., 2013; Smith and Baker, 2015). Defects in the mRNA surveillance machinery are associated with multiple diseases (reviewed in (Wolin and Maquat, 2019). Three main modes of cytoplasmic mRNA surveillance exist: a) nonsense-mediated decay (NMD), b) no-go decay (NGD), and c) no-stop decay (NSD). Mutations that lead to a premature termination codon (PTC) are directed to NMD (Figure 2A), mutations that lead to stalling of the ribosome from translocation, for example by introducing changes in the secondary structure of the mRNA and formation of a stable loop, trigger NGD (Figure 2B), and mutations that eliminate stop codons from a transcript trigger NSD (Figure 2C).

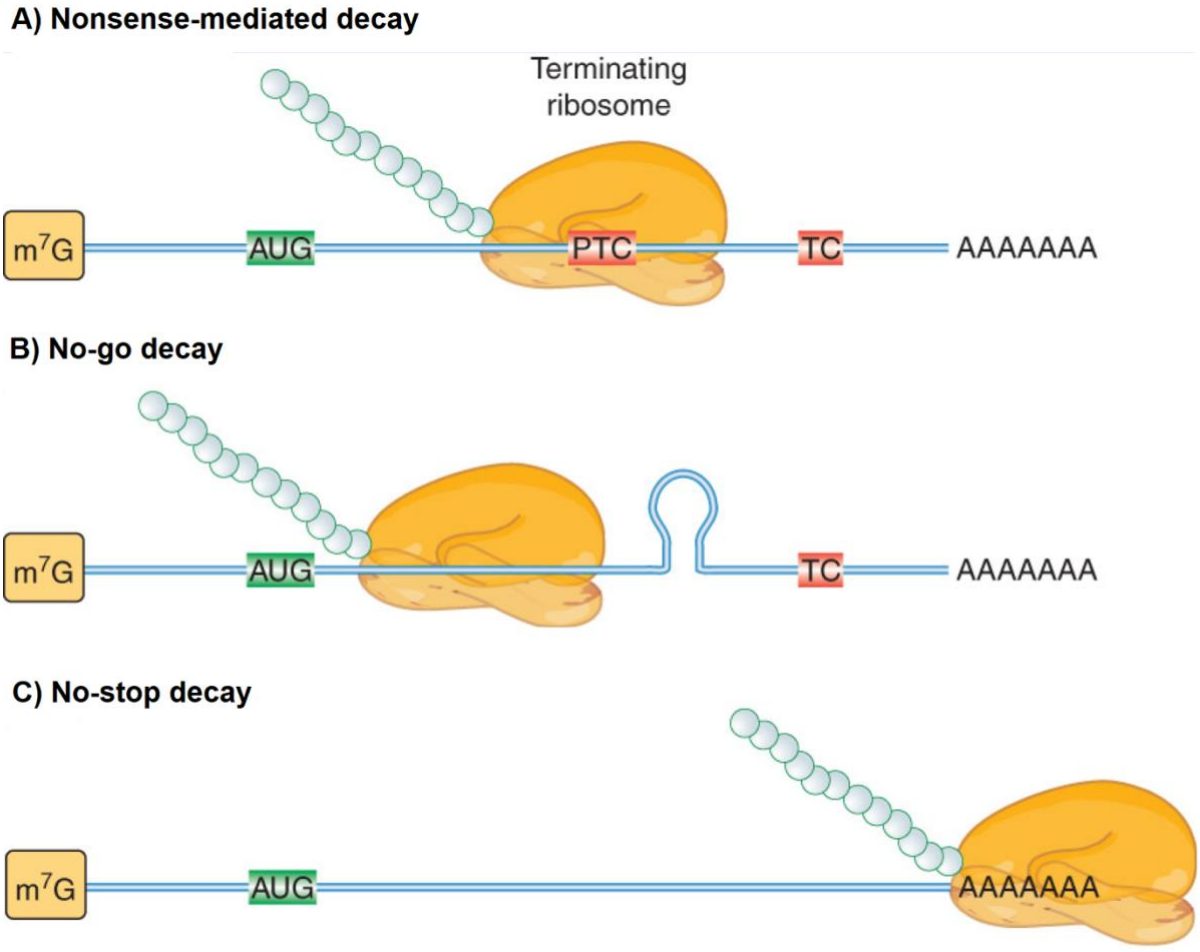


Figure 2. Cytoplasmic mRNA surveillance

(A) A translating ribosome stalls prematurely at a premature stop codon and thereby triggering nonsense-mediated decay. (B) A translating ribosome finds no stop codon to terminate at and thereby keeps on translating until it runs into the polyA of the mRNA leading it to stall and initiate no-stop decay. (C) A translating ribosome stalling at a very stable secondary structure (e.g., a stable loop) and is unable to translocate past it, leading to the initiation of no-go decay. Abbreviations, m7G: 7-methylguanylate cap; AUG: translation

start codon; PTC: Premature-termination codon; TC: Termination codon. Figure adapted from an RNA quality control lecture by Prof. Heikke Krebber, Uni Göttingen, Germany.

3.4.1. Nonsense-mediate decay

Nonsense mutations are mutations that lead to introduction of a premature termination codon (PTC) within the coding sequence of an mRNA (Figure 2A). As previously explained, such mutations are one very common in alleles used to study gene function. Following splicing, a multi-protein complex, known as the exon-junction complex (EJC) is deposited around 24 nucleotides upstream of exon-exon junctions (Le Hir et al., 2000; Le Hir et al., 2016). During the pioneer round of translation, the translocating ribosome normally removes proteins associated with 5'UTRs and the coding sequence of an mRNA, including the EJC, before terminating at a stop codon (Dostie and Dreyfuss, 2002; Sato and Maquat, 2009). More proteins become associated with the deposited EJC that regulate mRNA splicing, export, translation, and stability (reviewed in (Kurosaki et al., 2019; Tange et al., 2004)). The NMD factor UPF3B (also known as UPF3X) is one of the factors that are become associated with the EJC in the nucleus, and which then further recruits the other NMD component UPF2 in the cytoplasm (Chamieh et al., 2008; Gehring et al., 2003; Kim et al., 2001; Le Hir et al., 2001; Lykke-Andersen et al., 2001). According to one model, NMD is elicited if a PTC is present at least 50-55 nucleotides upstream of an exon-exon junction (Nagy and Maquat, 1998; Thermann et al., 1998). In that case, the ribosome is unable to remove the downstream EJC and translation termination becomes inefficient, possibly due to the EJC interfering in the interaction between the polyA binding proteins (PABP) and the release factor eRF3 (Kervestin et al., 2012; Silva et al., 2008). Normally, translation termination involves recruitment of the release factors eRF1 and eRF3 that promote the release of the peptide and the ribosome subunits. In the case of the presence of a PTC, and thereby inefficient translation termination, UPf1, and its kinase SMG1 are recruited to join the translation termination factors eRF1 and eRF3 forming the SURF complex (SMG1-UPF1-eRFs complex) (Ivanov et al., 2008). The SMG1 kinase is normally in a complex with SMG8 and SMG9 that inhibits SMG1's kinase activity (Deniaud et al., 2015; Yamashita et al., 2009). Once the SURF complex is assembled, however, it interacts with the downstream EJC-bound UPF2 leading to phosphorylation of UPF1 by SMG1 and initiation of mRNA degradation (Durand et al., 2016; Kashima et al., 2006; Kurosaki et al., 2014). UPF1 phosphorylation can lead to the recruitment of

the endonuclease SMG6 which cleaves the mRNA near the PTC (Eberle et al., 2009; Huntzinger et al., 2008). The resulting 5' cleavage product is then further degraded by the 3' to 5' exonucleases, mainly the exosome complex, but also the DIS3L2 exonuclease (Kurosaki et al., 2018; Schmid and Jensen, 2008). The 3' cleavage product is further degraded by the 5' to 3' exonuclease XRN1 (Nagarajan et al., 2013). Alternatively, phosphorylated UPF1 can recruit the SMG5-SMG7 heterodimer. SMG7 recruits the CCR4-NOT complex that deadenylates the mRNA allowing for efficient 3' to 5' degradation of the mRNA by the exosome complex and also recruits the decapping enzyme DCP2 that removes the 5'm7G cap allowing for efficient 5' to 3' degradation by XRN1. SMG5 itself also directly recruits the DCP2 (Loh et al., 2013; Unterholzner and Izaurralde, 2004; Yamashita et al., 2005).

Another mode of NMD has been proposed that is independent of the EJC. According to this model, PTCs that are at a distance of 1kb or more from the polyA tail can trigger NMD through what is commonly referred to as long 3'UTR model. This large distance between the PTC and the polyA tail prevents proper interaction between the polyA binding protein PABPC1 and the ribosome release factor eRF3, thereby preventing proper translation termination and recruits UPF1 to initiates NMD in a not very well understood mechanism (Amrani et al., 2004; Eberle et al., 2008; Silva et al., 2008). Such model fits for single-exon genes, which are more common in organisms like yeast (Celik et al., 2017; Malabat et al., 2015; Spingola et al., 1999). It also explains cases where NMD is triggered during the steady-state round of translation (after removal of the EJC during the pioneer round of translation) (Hoek et al., 2019). What triggers UPF1 phosphorylation in the absence of the EJC is still not very clear, however, reports have shown that UPF1 binds to mRNAs in a promiscuous fashion: it is present on most mRNAs and is normally removed by the translocating ribosome (Kurosaki et al., 2014; Lee et al., 2015). It is possible that a long 3'UTR with no ribosomes translocating through it would allow for increase UPF1 occupancy and higher chances of its phosphorylation in a stochastic manner.

3.4.2. No-go decay

Another mRNA surveillance machinery is triggered by slowing down or stalling of the translating ribosomes (Figure 2B). Such kind of stalls can be induced by very stable secondary structures (for example loops) that the ribosome can't translocate through, translation of rare codons with suboptimal levels of the corresponding tRNA or difficulties in forming a peptide bond between two particular amino acids (reviewed in

(Schuller and Green, 2018)). Mutations can often lead to changes in the mRNA secondary structure and indels can lead to frameshifts which may include rare codons, and thereby triggering NGD. Due to the slowdown in translation, the chance of two ribosomes colliding with each other increases. Such collision leads to the recruitment of mRNA decapping factors that removes the 5'm7G cap of the transcript allowing for efficient XRN1-dependent 5' to 3' degradation. Ribosome collision also exposes a 40S-40S interface that is recognized by the ubiquitin ligase ZNF598 (Hel2 in yeast) leading to ubiquitination of some of the 40S proteins which is required for endonucleolytic cleavage (Ikeuchi et al., 2019; Juszkiewicz et al., 2018). A recent yeast study identified the recruitment of a novel endonuclease called Cue2 (NEDD4 binding protein 2 (N4BP2) being the potential mammalian ortholog) to the collided ribosomes, through possibly recognizing the ubiquitinated 40S ribosomal proteins, leading to cleaves the transcript between collision sites producing 5' and 3' cleavage products. The ribosome rescue factors PELO (Dom34 in yeast) and HBS1 then release the ribosomes them from the 5' and 3' cleavage products allowing for efficient degradation of the fragments by the different exonucleases (D'Orazio et al., 2019; Tsuboi et al., 2012), reviewed in (Wolin and Maquat, 2019)).

3.4.3. No-stop decay

A third model of mRNA surveillance involves mRNAs lacking a stop codon and thereby proper release of the translating ribosome (Figure 2C). Lack of a stop codon may lead to the ribosome translating into the polyA signal which leads to ribosome stalling due to the electrostatic interactions between the polyA-encoded positively charged poly-lysine peptide and the negatively charged peptide tunnel in the ribosome according to some studies (Ito-Harashima et al., 2007; Lu and Deutsch, 2008) and other more complex mechanisms according to other studies (Chandrasekaran et al., 2019; Juszkiewicz and Hegde, 2017). Stalling of the ribosome at the 3' end of the transcript initiates no stop decay in a manner dependent on the exosome associated protein Ski7 in yeast (HBS1L in mammals) (Frischmeyer et al., 2002). According to one study, the C-terminus of Ski7 has a GTPase domain that is closely related to that of the ribosome release factors eRF1 and eRF3. The authors thereby proposed that Ski7 binds in the A-site of ribosomes stalled on no-stop mRNAs and recruits the 3' to 5' exonuclease, the exosome complex, to initiate mRNA degradation in a manner independent of deadenylation by the CCR4-NOT complex ((van Hoof et al., 2002), reviewed in (Klauer and van Hoof,

2012)). Other studies reported similar requirements for NSD as that of NGD. Release of the stalled ribosomes by PELO and HBS1 and endonucleolytic cleavage was reported to also influence NSD (Tsuboi et al., 2012). A recent preprint also proposed Cue2 to be the endonuclease involved in NSD (Glover et al., 2019), and a previous study suggested endonucleolytic roles for the exosome complex involved in NSD (Schaeffer and van Hoof, 2011).

3.5. Importance of understanding the molecular mechanisms underlying transcriptional adaptation

Several recent studies have identified healthy individuals with homozygous loss-of-function mutations in several genes (MacArthur et al., 2012; Narasimhan et al., 2016; Sulem et al., 2015). A study has even recently identified healthy resilient individuals who harbor mutations in 8 different disease-associated genes, without manifesting any of the symptoms associated with the disease (Chen et al., 2016b). Functional analysis of the reported alleles remains to be characterized, but it is likely genetic compensation may contribute to the resilience of such individuals. Investigating the mechanisms underlying genetic compensation and transcriptional adaptation may help us understand why certain mutations cause disease and others not and may lead to the development of more new therapies, ones that can promote an individual's robustness to a mutation rather than correcting its effect.

4. Aims of the study

The aim of this Ph.D. thesis is to identify the molecular mechanisms underlying the transcriptional adaptation, a phenomenon that underlies genetic compensation. As mentioned earlier, genetic compensation is a form of genetic robustness to mutations that involves upregulation of related genes (such as paralogs). Previous studies reported that such as response is triggered independent of the loss of protein function (Rossi et al., 2015; Tondeleir et al., 2012), suggesting the presence of an unknown trigger, and thereby a potential novel gene expression regulation mechanism. Moreover, the increasing number of generated mutant animals showing no obvious phenotype (Bouche and Bouchez, 2001; Chen et al., 2016b; Giaever et al., 2002; Kok et al., 2015) has hindered our understanding of gene function. With genetic compensation being one of the main underlying reasons (El-Brolosy and Stainier, 2017; Rossi et al., 2015), identifying the underlying molecular mechanisms may identify guidelines in better designing mutant alleles with minimal compensation responses, and thereby facilitating studying gene function.

Thereby, I have the following two specific aims:

Aim 1: Identify transcriptional adaptation's molecular trigger.
I will first aim to confirm that loss of protein function is not the trigger underlying transcriptional adaptation. Next, I will investigate two other possibilities that might underly the response 1) the DNA lesion, and 2) the mutant mRNA molecules.

Aim 2: Provide guidelines for the generation of mutant alleles with minimal transcriptional adaptation-derived compensation.
Following identification of the trigger underlying the response, I will aim at providing guidelines to circumvent such trigger when designing mutant alleles. These guidelines shall help in better designing alleles with minimal genetic compensation responses and thereby enable unmasking hidden phenotypes to better understand gene function.

Aim 3: Investigate the molecular mechanisms underlying transcriptional adaptation.

After identifying the trigger, I will aim at understanding the molecular mechanisms through which it induces the response.

5. Materials and Methods

5.1. Materials

5.1.1. Antibiotics

Antibiotics	Working concentration
Ampicillin	100 µg/ml
Kanamycin	50 µg/ml
Streptomycin	100 µg/ml
Penicillin	100 U/ml
G418	0.5 mg/ml or 2 mg/ml
Puromycin	0.5 µg/ml

Table 4. Antibiotics used and their working concentrations.

5.1.2. Antibodies

Antibody	Dilution/Working concentration	Application	Supplier
FERMT2	1:1000	Western Blotting	Millipore
RELA	1:1000	Western Blotting	Cell Signaling Technology
ACTB	1:1000	Western Blotting	Cell Signaling Technology
anti-mouse IgG-HRP	1:10000	Western Blotting	Thermo Fisher Scientific
anti-rabbit IgG-HRP	1:10000	Western Blotting	Thermo Fisher Scientific
Alexa Fluor 568 Phalloidin	1:50 (for zebrafish larvae) and 1:1000 (for mESCs)	Immunostaining	Thermo Fisher Scientific
WDR5	4µg/IP	Chromatin immunoprecipitation	Cell Signaling Technology
H3K4me3	4µg/IP	Chromatin immunoprecipitation	Cell Signaling Technology
rabbit IgG	4µg/IP	Chromatin immunoprecipitation	Thermo Fisher Scientific

Table 5. Antibodies used, their working concentrations, applications they were used for and supplier's name.

5.1.3. Bacterial strains

Bacterial strain	Application
DH5α	Competent cells for transformation

Table 6. Bacterial strain used and its application.

5.1.4. Buffers and solutions

Buffer/Solution	Composition
Egg water	3g Instant Ocean 0.75g Calcium Sulfate 10 liters dH_2O
PBS	8g NaCl 0.2g KCl 1.44g Na2HPO4 0.24g KH2PO4 pH 7.4 1 liter dH_2O
20X SSC	175.3g NaCl 88.2g Sodium Citrate pH 7 1 liter dH_2O
10X TBE	121g Tris 62g Boric Acid 7.4g EDTA 1 liter dH_2O
4% PFA	Add 8 g of PFA to 140 ml of PBS then heat the solution to 60ºC until PFA gets dissolved. Adjust pH to 7 after cooling then make up volume to 200 ml with dH_2O. Filter solution.
RIPA buffer	150 mM NaCl 1.0% IGEPAL CA-630 0.5% sodium deoxycholate 0.1% SDS 50 mM Tris pH 8.0
Blocking buffer	5% BSA or 5% non-fat milk
PBST	PBS 0.1% Tween 20

PBT	PBS 0.1% Triton X-100
PBDT	1% DMSO, 1% BSA and 0.5% Triton-X in PBS
TE buffer	10 mM Tris pH8.0 1mM EDTA
ChIP low salt buffer	0.1% SDS 1% Triton X-100 2mM EDTA 20mM Tris-HCl pH 8.1 150 mM NaCl
ChIP high salt buffer	0.1% SDS 1% Triton X-100 2mM EDTA 20mM Tris-HCl pH 8.1 500 mM NaCl
ChIP LiCl buffer	10mM Tris-HCl pH 8.0 250mM LiCl 1% NP-40 1% deoxycholic acid 1mM EDTA
ChIP elution buffer	1% SDS 100mM NaHCO3
Zebrafish embryo and larvae lysis buffer	50 mM NaOH

Table 7. List of the buffers and solutions used and their composition.

5.1.5. Chemicals and reagents

Chemical	Supplier
Mineral Oil	Sigma
SOC media	Thermo Fisher Scientific
BSA	Sigma
Chloroform	Merck
DMSO	Sigma
DNA ladder (1kb and 100bp)	Thermo Fisher Scientific
Gel Loading Dye	Thermo Fisher Scientific
Ethanol	Roth

Glycerol	Sigma
Trizol	Ambion
Low Melting Agarose	Sigma
Paraformaldehyde	Sigma
Pronase	Roche
Methylene blue	Sigma
SYBR Safe	Thermo Fisher Scientific
Tris	Sigma
Triton-X	Sigma
Tween-20	Sigma
Precision Plus Protein Standard	Bio Rad
Isopropanol	Roth
Phenylmethylsulfonyl fluoride	Applichem
EDTA-free complete protease inhibitor cocktail	Roche
Tricaine	Pharmaq
PBS	Sigma
CutSmart buffer	NEB
Nuclease-free water	Ambion
Agarose	Peqlab
Milk	Sigma
PTU	Sigma
Phenol-chloroform-isoamyl alcohol	Thermo Fisher Scientific
Formamide (deioinized)	Ambion
Phenol Red	Sigma
DAPI	Thermo Fisher Scientific
Penicillin-Streptomycin	Thermo Fisher Scientific
Trypsin-EDTA	Thermo Fisher Scientific
NMDi14	Merck Millipore
Cycloheximide	Sigma
4-thiouridine	Sigma

Biotin-HPDP	Sigma
Actinomycin D	Sigma
Mouse TNFα	Sigma
MTT	Thermo Fisher Scientific
Dako fluorescent mounting medium	Aglient
Clarity Western ECL Substrate	Bio-Rad
10X Tris/Glycine/SDS buffer	Bio-Rad
Lipofectamine RNAiMax	Thermo Fisher Scientific
Lipofectamine MessengerMax	Thermo Fisher Scientific
FuGENE 6	Promega
FuGENE HD	Promega
LB agar	Roth
LB medium	Roth
DMEM high glucose, pyruvate	Thermo Fisher Scientific
PluriQ-ES-DMEM	Thermo Fisher Scientific
Bovine Calf Serum	HyClone
ESCs-qualified fetal bovine serum	Millipore
2i	Sigma
ESGRO	Chemicon international
dNTP mix (10 mM each)	Thermo Fisher Scientific
DTT	Thermo Fisher Scientific

Table 8. List of chemicals and reagents used along with their suppliers.

5.1.6. Kits

Kit	Supplier
RNA Clean and Concentrator kit	Zymo Research
T7 mMessage mMACHINE kit	Ambion
SP6 mMessage mMACHINE kit	Ambion
Maxima First Strand cDNA Synthesis kit with dsDNase	Thermo Fisher Scientific
GeneJET PCR purification kit	Thermo Fisher Scientific
GeneJET gel extraction kit	Thermo Fisher Scientific

pGEM-T-easy vector kit	Promega
GeneJET plasmid miniprep kit	Thermo Fisher Scientific
Cell line nucleofector kit R	Lonza
SMARTer Stranded Total RNA Sample Prep Kit–HI Mammalian	Clonetech
TruChIP chromatin shearing reagent kit	Covaris
NucleoSpin Gel and PCR Clean-up kit	Macherey-Nagel
µMacs Streptavidin kit	Miltenyi
miRNeasy Micro kit	Qiagen
RNeasy MinElute clean up kit	Qiagen
Quick-DNA 96 kit	Zymo Research
Nextera DNA Sample Preparation Kit	Illumina
MinElute PCR Purification Kit	Qiagen
MARTer Stranded Total RNA Sample Prep Kit–HI Mammalian	Clonetech
Cold fusion cloning kit	SBI biosystems

Table 9. List of kits used along with their suppliers.

5.1.7. Growth media

Growth medium	Composition
***E. coli* SOC medium**	Tryptone 2% Yeast extract 0.5% NaCl 0.05% KCl 0.0186% Dissolve in dH_2O Adjust pH to 7, then add: MgCl2 10 mM D-glucose 20 mM Autoclave
MEFs culture medium	DMEM high glucose, pyruvate 10% Bovine Calf Serum 1% Penicillin/Streptomycin
mESCs culture medium	PluriQ-ES-DMEM 15% ESCs-qualified fetal bovine serum 2 mM glutamine 1% non-essential amino acids

	0.1 mM β-mercaptoethanol 1,000 U/ml ESGRO (Lif) 2i (3 µM CHIR99021 and 1 µM PD0325901) 1% Penicillin/Streptomycin

Table 10. List of growth media used along with their composition.

5.1.8. Enzymes

Enzyme	Supplier
RQ1 RNase-free DNase	Promega
T4 DNA ligase	Takara
KAPA 2G fast DNA polymerase	Kapa Biosystem
Proteinase K	Roche
T7 RNA polymerase	Promega
Sp6 RNA polymerase	Promega
RNasin ribonuclease inhibitor	Promega
SYBR green PCR master mix	Thermo Fisher Scientific
SeqAmp DNA polymerase	Takara
T4 DNA polymerase	NEB
NotI-Hf, XbaI, BamHI, BbsI-HF, XhoI and other restriction enzymes	NEB
Tn5 Transposase	Illumina
DNase-Free DNase Set	Qiagen

Table 11. List of enzymes used and their respective suppliers.

5.1.9. Centrifuges

Centrifuge	Supplier
Centrifuge 5417 R (200 µl tubes)	Eppendorf
Centrifuge 5418 (1.5-2ml tubes)	Eppendorf
Centrifuge 5415D (1.5-2ml tubes)	Eppendorf
Centrifuge 5810 R (15-50 ml tubes and 96 well plates)	Eppendorf

Table 12. List of centrifuges used and their respective suppliers.

5.1.10. Microscopes

Microscope	Supplier
LSM 700 confocal microscope	Zeiss
Spinning disc CSU-X1 confocal microscope	Zeiss
LSM880 Axio Examiner confocal microscope	Zeiss
Stemi 2000 stereomicroscope	Zeiss
SMZ18 stereomicroscope	Nikon
SMZ25 stereomicroscope	Nikon

Table 13. List of microscopes used and their respective suppliers.

5.1.11. Laboratory equipment

Equipment	Supplier
Nanodrop 2000 c	Thermo Fisher Scientific
ChemiDoc MP	Bio-Rad
Trans blot turbo transfer system	Bio-Rad
Criterion SDS PAGE Chamber	Bio-Rad
CFX connect real time PCR detection system	Bio-Rad
Bioruptor sonicator	Diagenode
FLUOstar Omega spectrophotometer	BMGH Labtech
PCR mastercycler Pro	Eppendorf
Eco Real-time PCR system with HRMA	Illumina
Electrophoresis power supply	Bio-Rad
Gel Doc EZ System	Bio-Rad
Injection micromanipulator	World precision instruments
Microscale	Novex
Weighing balance	Sartorius
Heating blocks	VWR
Microwaves	Bosch

Bacterial incubator shaker	Infors HAT
Bacterial incubator	Heraeus
Nucleofector 2b device	Lonza
Zebrafish aqua culture system	Techniplast
Zebrafish breeding tanks	Techniplast
Zebrafish embryo and larvae incubator	Binder
Cell culture CO_2 incubators	Thermo Fisher Scientific
Cell culture laminar flow hoods	Thermo Fisher Scientific
DynaMag-2 Magnet	Thermo Fisher Scientific
BD FACS Aria III sorter	BD Biosciences
LUNA-II automated cell counter	Logos Biosystems
MOXI Z Mini Automated Cell Counter Kit	Orflo
NextSeq500 platform	Illumina
Next Advance Bullet Blender Homogenizer	Scientific Instrument Services

Table 14. List of equipment used and their respective suppliers.

5.1.12. Laboratory supplies

Laboratory supply	Supplier
Latex gloves	Roth
Nitrile gloves	VWR
Pipetboy	Integra
Pipettes (2 µl, 20 µl, 100 µl, 200 µl, 1000 µl)	Gilson
Pipette tips	Greiner bio-one
Pipette filter tips	Greiner bio-one
Centrifuge tubes (1.5 ml, 2 ml)	Sarstedt
PCR tubes (200 µl, 500 µl)	Sarstedt
Falcons tubes (15 ml, 50 ml)	Greiner bio-one
Bacterial culture tubes	Sarstedt
Beakers	VWR
Petri dishes (90 mm, 60 mm, 35 mm)	Greiner bio-one

Glass bottles (100 ml, 250 ml, 500 ml, 1000 ml, 2000 ml)	Duran
Conical flasks (100 ml, 500 ml)	VWR
Laboratory film	Parafilm
Forceps	Dumont
CELLSTAR cell culture flasks (T25, T75, T125)	Greiner bio-one
CELLSTAR cell culture multi-well plates (96, 48, 24, 12, 6 well plates)	Greiner bio-one
CELLSTAR cell culture dishes (35 mm, 60 mm, 100 mm, 145 mm)	Greiner bio-one
Magna ChIP Protein A+G Magnetic Beads	Millipore
4–20% Mini-PROTEAN TGX Precast Protein Gels	Bio-Rad
Trans-Blot Turbo Mini Nitrocellulose Transfer Pack	Bio-Rad
LabChip Gx Touch 24	Perkin Elmer
Bradford protein assay	BioRad
Primers	Sigma

Table 15. List of miscellaneous supplies used and their respective suppliers.

5.1.13. Software and databases

Software/Database	Purpose	Source	Citation (if available)	Online link (if web-based tool)
ApE	Nucleotide sequence management	University of Utah		
Adobe Photoshop & Illustrator	Figure formatting, illustrations	Adobe		
Primer3	Primer design	University of Massachusetts medical school		http://biotools.umassmed.edu/bioapps/primer3_www.cgi
BLASTn	Identifying sequence alignments and similarities	NIH	(Altschul et al., 1990)	https://blast.ncbi.nlm.nih.gov
Ensembl	Genomic sequences analysis and BLAST	Ensembl		http://www.ensembl.org/
IGV	Genomic sequence and next generation sequencing data analysis	Broad Institute	(Robinson et al., 2011)	

UCSC genome browser	Genomic sequence and next generation sequencing data analysis	University of California Santa Cruz		https://genome.ucsc.edu/
Trimmomatic v.0.33	Trimming NGS reads		(Bolger et al., 2014)	
Reaper v.13-100	Trimming NGS reads		(Davis et al., 2013)	
FastQC	Assessing sample quality for NGS	Babraham bioinformatics		http://www.bioinformatics.babraham.ac.uk/projects/fastqc
STAR 2.4.2a	Mapping NGS reads to mouse genome		(Dobin et al., 2013)	
Picard 1.136	Deduplicating NGS reads	Broad Institute		http://broadinstitute.github.io/picard/
MACS2 peak caller v.2.1.0	Identifying peaks from ATAC-seq	https://github.com/taoliu/MACS		
featureCounts 1.6.0	Counting RNA-seq reads that align to genes	Subread package	(Liao et al., 2014)	
bamCoverage	Converting binary alignment map (BAM) files to bigWig format	Deep Tools	(Ramirez et al., 2014)	
DESeq2	Normalizing raw counts for unified peaks from NGS and identifying differentially expressed genes		(Anders and Huber, 2010; Love et al., 2014)	
R	Bioinformatic analyses	R Foundation for Statistical Computing		
genomecov	Normalizing BigWig files for IGC	bedtools	(Quinlan and Hall, 2010)	
Kablammo	Visualizing BLASTn alignments		(Wintersinger and Wasmuth, 2015)	http://kablammo.wasmuthlab.org/
MUSCLE	Multiple sequences alignment	University of California, Berkeley	(Edgar, 2004)	
KOBAS 2.0	Gene set enrichment analysis	Peking University	(Xie et al., 2011)	
ZFIN	Gene expression, nomenclature	ZFIN		https://zfin.org/

Prism	Data analyses and visualization	Graphpad		
Imaris	Image processing	BitPlane		
Image J	Image processing	NIH	(Schindelin et al., 2012)	
Zen (Blue & Black)	Image visualization and processing	Zeiss		
CHOPCHOP	gRNA design for CRISPR/Cas9 mutagenesis	University of Bergen	(Labun et al., 2016)	http://chopchop.cbu.uib.no/
Microsoft office (Word, Excel, PowerPoint)	Writing, data analysis and figure formatting	Microsoft		

Table 16. List of databases and software used along with the purpose they were used for, their sources, citations and online links, if available.

5.1.14. Zebrafish food

Food	Developmental stage
SDS100	5 dpf – 12 dpf
SDS200	1-2 months
SDS300	2-3 months
SDS400	>3 months
Brine Shrimp	>1 month

Table 17. List of zebrafish food used for different developmental stages

5.1.15. Zebrafish lines

Line/allele number	Description	Source
Tüb/AB	Wild Type	
hif1ab^{bns90}	*hif1ab*$^{\Delta 8}$ mutant	(Gerri et al., 2017)
vegfaa^{bns1}	*vegfaa*$^{\Delta 10}$ mutant	(Rossi et al., 2016)
egfl7^{s980}	*egfl7*$^{\Delta 3}$ mutant	(Rossi et al., 2015)
egfl7^{s981}	*egfl7*$^{\Delta 4}$ mutant	(Rossi et al., 2015)
hbegfa^{sa18135}	*hbegfa*^{ENU} mutant	Sanger institute zebrafish mutation project (Kettleborough et al., 2013); http://www.sanger.ac.uk/resources/zebrafish/zmp/
vcla^{sa14599}	*vcla*^{ENU} mutant	Sanger institute zebrafish mutation project

		(Kettleborough et al., 2013); http://www.sanger.ac.uk/resources/zebrafish/zmp/
Tg(*fli1a*:*eGFP*)$_{y1}$	Vascular-specific reporter line	(Lawson and Weinstein, 2002)
TgBAC(*etsrp*:*eGFP*)$_{ci1}$	Vascular-specific reporter line	(Proulx et al., 2010)
alcama$_{bns201}$	*alcama*$_{\Delta 10}$ mutant	Dr. Jenny Pestel (MPI for Heart and Lung research); (El-Brolosy et al., 2019)
upf1$_{ya3319}$	*upf1*$_{\Delta 3ins1}$ mutant	Dr. Carter Takacs and Prof. Antonio Giraldez (Yale University, New Haven, CT, USA); (El-Brolosy et al., 2019)
hbegfa$_{bns189}$	*hbegfa*$_{\Delta 7}$ mutant	This study; (El-Brolosy et al., 2019)
hbegfa$_{bns203}$	*hbegfa*$_{\Delta 3}$ mutant	
hbegfa$_{bns243}$	*hbegfa*$_{full\ locus\ del.}$ mutant	
vcla$_{bns241}$	*vcla*$_{\Delta 13}$ mutant	
vcla$_{bns300}$	*vcla*$_{exon22\_ins1}$ mutant	
vegfaa$_{bns301}$	*vegfaa*$_{5'UTR\Delta 10}$ mutant	
vegfaa$_{bns242}$	*vegfaa*$_{promoter\text{-}less}$ mutant	
egfl7$_{bns303}$	*egfl7*$_{5'UTR\Delta 3}$ mutant	
egfl7$_{bns302}$	*egfl7*$_{full\ locus\ del}$ mutant.	
alcama$_{bns244}$	*alcama*$_{promoter\text{-}less}$ mutant	

Table 18. List of zebrafish lines used in this thesis and their sources.

5.1.16. Mouse cell lines

Line	Source
WT MKFs	(Theodosiou et al., 2016)
Fermt2 K.O. MKFs	
WT MEFs	(Gapuzan et al., 2005; Gapuzan et al., 2002)
Rela K.O. MEFs	

Rela$_{promoter\text{-}less}$ MEFs	This study; (El-Brolosy et al., 2019)
Actg1 K.O. (*Actg1*$_{NSD}$) MEFs	
Actg1$_{full\text{-}locus\ del.}$ MEFs	
WT mESCs	J. Kim (MPI for Heart and Lung Research, Bad Nauheim, Germany)
Actb K.O. mESCs	Dr. Andrea Rossi and Dr. Zacharias Kontarakis (MPI for Heart and Lung research); (El-Brolosy et al., 2019)
Actb heterozygous mESCs	
Actb$_{full\text{-}locus\ del.}$ mESCs	

Table 19. List of mouse cell lines used in this thesis and their sources

5.1.17. Plasmids

Plasmid	Purpose	Bacterial resistance	Source	Citation
pT3TS-nlsCas9nls	Cas9 expression	Ampicillin	Addgene (#46757)	(Jao et al., 2013)
pGEM-T	Vector for sequence cloning	Ampicillin	Promega	
pCS2+	Vector for cDNA cloning under a T7 promoter	Ampicillin		
pCDNA3.1	Mammalian expression vector	Ampicillin	Thermo Fisher Scientific	
XRN1-resistant sequence plasmid	Plasmid containing the XRN1-resistant sequence	Ampicillin		(Boehm et al., 2016)
PX458	Co-expression of Cas9-2A-GFP and gRNA in mouse cell lines	Ampicillin	Addgene (#48138)	(Ran et al., 2013)
PX459	Co-expression of Cas9-2A-Puro$_r$ and gRNA in	Ampicillin	Addgene (#62988))	(Ran et al., 2013)

	mouse cell lines			
pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-GFP	CRISPRi	Ampicillin	Addgene (#71237)	(Thakore et al., 2015)

Table 20. List of plasmids used in this thesis, the purpose they were used for, their bacterial resistance gene and source.

5.1.18. Oligonucleotides

5.1.18.1. siRNAs

siRNA target	**Provider**	Sequence (5' to 3') or company code if pool of siRNAs	Used in combination?	Epigenetic role, if applicable
Control (Scr)	Sigma	SIC001		
UPF1	Sigma	GUUCCAUCCUCAUUGACGA[dT][dT]		
EXOSC4	Sigma	GCCUGUUCUUCUCCCGAGU[dT][dT]		
ERF1	Santa Cruz	sc-37872		
SMG6	Santa Cruz	sc-61570		
XRN1	Santa Cruz	sc-61812		
SETD7	Sigma	EMU064601		Deposits the permissive H3K4me1 histone mark (Nishioka et al., 2002; Wang et al., 2001)
KDM6a	Santa Cruz	sc-76882	Yes (KDM6)	Removes the repressive H3K27me3 histone mark (Agger et al., 2007; De Santa et al., 2007; Lan et al., 2007)
KDM6b	Sigma	EMU206511		
KDM4a	Sigma	EMU055901	Yes (KDM4)	Removes the repressive H3K9me3 histone mark (Cloos et al., 2006; Fodor et al., 2006; Whetstine et al., 2006)
KDM4b	Sigma	EMU014571		
KDM4c	Sigma	EMU053281		
WDR5	Sigma	EMU055581		Part of the COMPASS complex that deposits the permissive H3K4me3 histone mark

				(Ruthenburg et al., 2006; Shilatifard, 2012; Wysocka et al., 2005)
CREBBP	Sigma	EMU207651	Yes (P300/CREBBP)	Deposits the permissive H3K27ac histone mark (Jin et al., 2011; Ogryzko et al., 1996)
P300	Sigma	EMU078861		
TET1	Sigma	EMU089111	Yes (TET)	DNA demethylation (Wu and Zhang, 2017)
TET2	Sigma	EMU147291		
TET3	Sigma	EMU207421		

Table 21. List of siRNAs used in this thesis, their providers and sequence or company code.

5.1.18.2. Genotyping primers

Mutant line	Primer sequence (5' to 3') (Top: forward; Bottom: reverse)	Method of genotyping
hbegfa$_{bns189}$ zebrafish	TTTTGAACGCGGAGAAAC	HRMA
	TAGACAGGGGTTTCTTTTCTTG	
vcla$_{bns241}$ zebrafish	GTGCAGGGAAAGAGAGACGA	HRMA
	GACACCTGATCCGTCATCTG	
hif1ab$_{bns90}$ zebrafish	CTCATCTGTGAGCCCATTCC	HRMA
	GCTGAGGAAGGTCTTGCTGT	
egfl7$_{s981}$ zebrafish	GAGTCGTCGCGTGTCTTACA	HRMA
	GGCACATGGTGATGTAGGGTT	
vegfaa$_{bns1}$ zebrafish	CGAGAGCTGCTGGTAGACATC	HRMA
	GGATGTACGTGTGCTCGATCT	
alcama$_{bns201}$ zebrafish	CTGCCTGATTTCTCCCAGTT	HRMA
	AGCTCGAAGGAAAAGCTGTG	
hbegfa$_{sa18135}$ zebrafish	GGGAAAAGGCAAGAAAAGAAA	PCR
	CCGTGGATGCAAAAGTCC	
vcla$_{sa14599}$ zebrafish	GCCGTACAGACAGGAGCTG	HRMA
	CATCATCAGCTGCTCCACCT	
Actb K.O. mESCs	GCCTTCTTTTGTGTCTTGATAGTTC	HRMA

	CACGATGGAGGGGAATACAG	
Actb heterozygous mESCs	GGGAATGGGTCAGAAGGACT	HRMA
	AATGGGGTACTTCAGGGTCA	
Rela$_{promoter\text{-}less}$ MEFs	GGCCTCGAACTAGGAGACCT	Out-out PCR
	TGTCCAGCCATAGGGTTTGT	
Actb$_{full\text{-}locus\ del.}$ mESCs	GCACAGCTTCTTTGCAGCTC	Out-out PCR
	GTCAAAGAAAGGGTGTAAAACG	
hbegfa$_{bns243}$ zebrafish	GCAGGTAACCATACCAGGGATA	Out-out PCR
	AGACAAGGCTTCATCAGTACTAAAA	
vegfaa$_{bns242}$ zebrafish	ACACAGCAGTTTGCGAGAAA	Out-out PCR
	CAAATAAACAACCAAGTTCAT	
alcama$_{bns244}$ zebrafish	TGGTGTCAGTTTCTTCCTCAGA	Out-out PCR
	CAAACGATCCAGTTTTCACG	
upf1$_{ya3319}$ zebrafish	GGGGAAAACCAGCCAACT	HRMA
	GTGTCTTCCTCGTCCTCCTC	
egfl7$_{bns303}$ zebrafish	GCCCAAAAGCATCTCATTTC	HRMA
	CATTTCTTATAAACTCTTCTTCAGTCT	
vegfaa$_{bns301}$ zebrafish	GCCAAAACAGTCACGGAAAT	HRMA
	GAGCAAAGGCTTGCTGTAAA	
vcla$_{bns300}$ zebrafish	GTGCACAACGCCCAGAAC	HRMA
	AGCGGCTTCTGCTTCTCTC	
egfl7$_{bns302}$ zebrafish	AAATGCTGGCGATGGAGAGG	Out-out PCR
	TGCTTACAGGCTCCGATTCT	
Actg1$_{NSD}$ MEFs	AGGAGATCACAGCCCTAGCA	Out-out PCR
	CCTGATCCTGTCACCTCCAC	
Actg1$_{full\text{-}locus\ del.}$ MEFs	GGCTTACACTGCGCTTCTTG	Out-out PCR
	CCTGATCCTGTCACCTCCAC	

Table 22. List of primers used to genotype zebrafish and mouse mutant lines and method of genotyping.

5.1.18.3. qPCR primers

Organism	Targeted gene	Primer sequence (5' to 3') (Top: forward; Bottom: reverse)
Zebrafish	*hbegfa* mRNA	TCTTTACCATCGTGGCTGTG
		CTCAGCGCCTCCAATAAATC
	hbegfb mRNA	GCTAAACCTGCCAAGAGTGG
		TTCCTTTGCCCTTCCTTTTT
	vcla mRNA	TCATCTTCTGGAGGCAATTA
		AAAACATCAGACACTTCCTG
	vclb mRNA	GGTCAAATCAGCCAGAGACC
		CATGGTCTCGAAGTGCTCAA
	hif1ab mRNA	CCTCTGGATCAAAACCCAAG
		TCAAGAGGTCATCTGGCTCA
	epas1a mRNA	AATTTCAATGTTCCACCACC
		TCCTAAACTCATGCCTTTCT
	epas1b mRNA	CAGCAAGACATTTCTGAGTC
		TGTAACCCTTTCATCACAGT
	egfl7 mRNA	GACGATTCTGCCAAATAGAT
		CGTATTCACACACTTCTGAG
	emilin3a mRNA	CAAACTTCCACAACAGTAGA
		GTTGTCATCATGTTCTTGGT
	emilin2a mRNA	CACCAGCCTGACTGTGAGC
		TTCAGTCACGGTCTTGTATGCT
	emilin3b mRNA	CAAGTGCATATGGGGACAGA
		TGACCGATTTTGTACTGTGGTT
	vegfaa mRNA	CGAGAGCTGCTGGTAGACATC
		GGATGTACGTGTGCTCGATCT
	vegfab mRNA	GGTGCTGCAATGATGAAATG

		TGTCACCCTGATGACGAAGA
	alcama mRNA	TGTACGGTGAGACCATCGAA
		GGAGCCGTCATCTTTCACAT
	alcamb mRNA	CATGAATCAAGTGTGTGGTT
		ATTTCAACAGTGTCTCCGTA
	rpl13 mRNA	TAAGGACGGAGTGAACAACCA
		CTTACGTCTGCGGATCTTTCTG
	hbegfb premRNA	ACCCTCTGAATGGCTTCTCA
		CTGTAACGCTGCTCCTCCTT
	emilin3a premRNA	AGCTGGAATGGAAGTGCTGT
		CTTTTGTGCAGGCATTGGT
	hbegfa premRNA	TTGCATGGTTTTCATTGCAT
		TGCACCTCTCACAGCCACTA
	egfl7 premRNA	CTACATCACCATGTGCCAAAA
		GATGGTGCATGGTACGGTTT
	alcama premRNA	ACCATCAGCCCCTGTAATCA
		CGGAAGCTCAGGGTCTTAAA
	vegfaa uncapped specific	AAACTGTGGACGGCTTTTTC
		GCCAAGCTATTTAGGTGACACTATAG
	hif1ab uncapped specific	CCTGCGAGTGTTTTATGACCT
		GCCAAGCTATTTAGGTGACACTATAG
	actb1 mRNA	GTATGCAGAAGGAAATCACC
		TCTTGATCTTCATGGTGGAA
	vclb antisense RNA-1 *ZFLNCT11314*	TTTCTATGAAAGCCCGTTCC
		TCTGACTTTATAACACGTAATTCTGGA
	vclb antisense RNA-2 *ZFLNCT11315*	TTTCTATGAAAGCCCGTTCC
		AACTCGTAATTCTGACTTTATAACTCG
	hbegfb antisense RNA-1	TGAATTGCTTCTTGCCACAC

	ZFLNCT18094	CCCCCTCAAAAGTATTGACTTG
Mouse	*Fermt2* mRNA	AAGTTTTCAAGCCGAAGAAG
		ATCCGGAGTAACTTCACATC
	Fermt1 mRNA	AAATACAGGACTTTGCAACC
		TAGGGATGTCAGTTATGTCC
	Rela mRNA	GCAGAAAGAGGACATTGAG
		GTGCACATCAGCTTGC
	Rel mRNA	AAAGACAACTCTGCTTTTCC
		CTTCCTTCTCCAATTGAACC
	Actb mRNA	CTGTATTCCCCTCCATCGTG
		CTCGTCACCCACATAGGAGTC
	Actg1 mRNA	GGAAGAAGAAATCGCCGCAC
		cctcgtcacccacgtatgag
	Actg2 mRNA	CTTCTACAATGAGCTTCGAG
		ACATGATCTGGGTCATCTTC
	Gapdh mRNA	ACCCAGAAGACTGTGGATGG
		ACACATTGGGGGTAGGAACA
	Rn18s	CGGACAGGATTGACAGATTG
		CAAATCGCTCCACCAACTAA
	Fermt1 premRNA	TTTGGAAGTGACCCTGGAAG
		ACATTTTCTCTCCGCTCCAA
	Rel premRNA	TGGAAAAGATTGCAGAGATGG
		TGCTGAAGGTTTCTGTCACTG
	Fermt2 premRNA	GACCGAGTCTTCAAGGCTGT
		ATGCCAAGTACCAGCCACAC
	Rela premRNA	TATTCCTGGCGAGAGAAGCA
		CGTTCCACCACATCTGTGTC
	Upf1 mRNA	ATTTGGTTAAGAGACATGCG

		GAGCTCAATAGCAATCTCATC
	Exosc4 mRNA	ATCGACATCTACGTGCAG
		CATGCGCATACAAAGTCC
	Smg6 mRNA	CAGTTGGCTCTACTGGCAGTGT
		CTTGGCAGTCAGGATAGGGTTG
	Cdk9 mRNA	TAAAGCCAAGCACCGTCAG
		GATTTCCCTCAAGGCTGTGAT
	Sox9 mRNA	AGACCCTTCGTGGAGGAG
		TCGGTTTTGGGAGTGGTG
	Fermt1 ChIP TSS	ACTGGACCCGCTGTACCTT
		ggagagctcacCTGCTATGG
	Rel ChIP TSS	CCGTGTAGAGACCTCGATCC
		AATCTCGTCCTCTTGCTGCT
	Actg2 ChIP TSS	CTTACCAGAGGCCAGCATGT
		TTGGCTATGGCCTAAACACC
	Fermt1 ChIP non-promoter region	TTTGGAAGTGACCCTGGAAG
		ACATTTTCTCTCCGCTCCAA
	Rel ChIP non-promoter region	TGGAAAAGATTGCAGAGATGG
		TGCTGAAGGTTTCTGTCACTG
	Actg2 ChIP non-promoter region	GGACTTTCTCCCCTCCAGAC
		GGGGCTTTGTGAGGATGTTA
	Ubap1l premRNA	AAGACTGGTGTCTGGGCTGT
		CCAGAGTATGCTGGCAGTGA
	Fmnl2 premRNA	TGCATGGGCTTAATGTGTGT
		AAGTGTGTAAGGGGCGTGAC
	Cdk12 premRNA	CAAGAATCTTCCTGCCTTGC
		GAGGACCGATGTTTGCATCT
	Actr1a premRNA	GGACTTTTGCCTGAGAGTGC

		ATCCTTCTGGGGGTTTATGG
	Erf1 mRNA	AGTGCTGCCGATAGGAACG
		GCCACTCGTGAAATCTGGTCTT
	Xrn1 mRNA	TACCTCGATTTGAGAGATACC
		AAGAATTTTCCTGACCCTTC
	Tet1 mRNA	CTACCCCTTACATGAAAAACAG
		CACAAAAATCCATGCAACAG
	Tet2 mRNA	CCGATGCATACAATAATCAGG
		GATTGTCTTCTCTATTGAGGG
	Tet3 mRNA	AGGATCGGTATGGAGAAAAG
		CAGGATCAAGATAACAATCACG
	Setd7 mRNA	TAGCAGTTGGACCTAATACTG
		ACATCAATGACTGTCTCCTC
	p300 mRNA	CAATAGAGCGGAATACTATCAC
		CAGTATTCATAGGAACTGGAC
	Crebbp mRNA	CCAATCCACTGATGAATGATG
		GCTTGAACGAGTTTATGGAC
	Kdm4a mRNA	CAAAGTCTTGGTACTCTGTTC
		AAAGTGATCATAAACTCGCC
	Kdm4b mRNA	AAGTACTGGAAGAACCTGAC
		CATGCCAAAGTACAAGTAGG
	Kdm4c mRNA	TTAAGAAAGCCTCCCAAGAG
		CAAATTCCTCATCTTCTTCAGG
	Kdm6a mRNA	AATTCAGTTTCACATTGCCC
		TGATGCATCCAACCTAATTG
	Kdm6b mRNA	TCAAGATGATCAAGTTCTGC
		CCATTCTCACTTGTAACGAAC
	Wdr5 mRNA	CATTGATGATGACAATCCTCC

		GTACTTCTCATTCTTGTGGC
Human	*BDNF* mRNA	GAACTCCCAGTGCCGAACTA
		CTTATGAATCGCCAGCCAAT
	BDNF antisense RNA	AGTGGCTAATCTTACAACAGCACAA
		CTCAGTAGTCAAGTGCCTTTGGA

Table 23. List of sequences of qPCR primers used in this thesis.

5.1.18.4. gRNAs

Organism	gRNA name	Sequence (5' to 3')	Mutant generated/ purpose	For RNA-less alleles, distance of PAM from ATG or Stop codon or TSS
Zebrafish	*hbegfa* exon 3 gRNA	GGCTCAAAGAGGAAGGGGCT	*hbegfa*$_{bns189}$	N/A
	vcla exon 8 gRNA	CCTGGGAACAGCCAAAACCC	*vcla*$_{bns241}$	
	upf1 gRNA 1	GGGGAAAACCAGCCAACTTC	*upf1*$_{ya3319}$	
	upf1 gRNA 2	GGCCGAGCTGAACTTTGAGG		
	vcla exon 22 gRNA	CATGCAGTCTGTGAAGGAGA	*vcla*$_{bns300}$	
	egfl7 5'UTR gRNA	GTCGGATCACGGGGAGCATC	*egfl7*$_{bns303}$	
	vegfaa 5'UTR gRNA	GTCAAACCAAGGATTGCCAC	*vegfaa*$_{bns301}$	
	hbegfa 5' full locus del. gRNA	CTGTAACGTTAGCAGACACT	*hbegfa*$_{bns243}$	113- from ATG
	hbegfa 3' full locus del. gRNA	TGGATGGCGAGGATGTAG		707+ from Stop
	egfl7 5' full locus del. gRNA	TGGACCGAATGGCCTC	*egfl7*$_{bns302}$	15183- from ATG
	egfl7 3' full locus del.	CGAACAGGTCGGCTTTCTGG		1637- from Stop

	gRNA			
	vegfaa promoter gRNA	TGCAATGGGAATTATGCGT	*vegfaa*$_{bns242}$	652- from TSS
	vegfaa 5'UTR_promoterless gRNA	CTCACCGCAACACTCCAC		572- from TSS
	alcama promoter gRNA	TGTCTGCCTAATGACAAAGT	*alcama*$_{bns244}$	1523- from TSS
	alcama intron1 gRNA	AGTCCGTGATCGGGGGAC		1047+ from TSS
Mouse	*Rela* promoter gRNA	GATGGGTTCAAACTATGTAG	*Rela*$_{promoter\text{-}less}$ MEFs	1144- from TSS
	Rela exon 3 gRNA	TCATCGAACAGCCGAAGCAA		1121+ from TSS
	Actb exon 3 gRNA	GGACTCCTATGTGGGTGACG	*Actb* heterozygous mESCs	N/A
	Actb exon 2 gRNA	CAGCGATATCGTCATCCA	*Actb* K.O.	
	Actb exon 2 gRNA	CAGCGATATCGTCATCCA	*Actb*$_{full\text{-}locus\ del.}$ mESCs	1+ from ATG
	Actb last exon gRNA	GCACCGCAAGTGCTTCTAGG		1+ from Stop
	Actg1 exon 1 gRNA	ATCTGCGCAGGAAGAAGCCC	*Actg1*$_{full\text{-}locus\ del.}$ MEFs	24- from ATG
	Actg1 3' gRNA	TGGTAAGACTGGTTATCCAA		1116+ from Stop
	Actg1 intron 5 gRNA	GTCCTAAGGCCAGCTCAGGC	*Actg1* K.O. (*Actg1*$_{NSD}$) MEFs	N/A
	Actg1 3' gRNA	TGGTAAGACTGGTTATCCAA		
	Fermt2 CRISPRi gRNA 1	GACAAACCCCACCGAGTCCG	*Fermt2* CRISPRi	
	Fermt2 CRISPRi gRNA 2	TTAAACGGGTGTCTTCCTCA		

	Fermt2 CRISPRi gRNA 3	AAACTCGAGTCAAATGCCAG		

Table 24. List of sequences of gRNAs used in this thesis.

5.1.18.5. Common sequencing primers

Primer name	Sequence (5' to 3')
SP6 F	ATTTAGGTGACACTATAG
T3 F	GCAATTAACCCTCACTAAAGG
T7 F	TAATACGACTCACTATAGGG
SV40 R	TGGGCGAAGAACTCCAGCATGAGAT
M13 F	GTAAAACGACGGCCAGT

Table 25. List of common sequencing primers used in this thesis.

5.2. Methods

5.2.1. Zebrafish husbandry

All zebrafish husbandry was performed under standardized conditions according to institutional (Max Planck Society) and national welfare and ethical guidelines approved by the committee for animal experiments ethics at the Regierungspräsidium Darmstadt, Germany.

5.2.2. Zebrafish maintenance

Zebrafish adults (Danio rerio, strain: Tüb/AB) were maintained at a water temperature of 28ºC in fish rooms with 29ºC temperature at a light-dark cycle of 14 hours of light and 10 hours of dark in a fish aqua culture system (Tecniplast) according to guidelines from (Westerfield, 2000). The aqua culture system is composed of tanks connected to freshwater source and a recycling system in racks. The sponge-like material of the recycling system's biological filters allows growth of aerobic denitrifying bacteria as Nitrobacter and Nitrosomonas. In addition, water is sterilized by UV light before going to the reservoir of fresh water. Embryos were kept in egg water at a 28ºC incubator in 90 mm Petri-dishes. For imaging purposes, PTU was added in the water to 24 hpf embryos to avoid pigment formation. To dechorionate embryos, 1mg/ml of pronase was added for 10 minutes.

5.2.3. Zebrafish mating

Male and female fish were bred in special mating tanks with egg permeable insets separated by a transparent divider around late afternoon/evening. The divider was removed on the next morning and around 20 minutes later, giving time for the females to be chased by the males and lay eggs, the fertilized eggs at the bottom of the mating tank were collected by filtering the water through a sieve. The eggs were then kept in a 90 mm petri-dish with 50 ml of egg water. Later in the afternoon the dead eggs were discarded after checking the eggs under a stereomicroscope and divided into 50 eggs/petri-dish.

5.2.4. Microinjections

To prepare injection plates, agarose was mixed with egg water to make a 2% solution that was then heated with a microwave to dissolve the agarose. The solution was then kept for a short period of time at room temperature, allowing the

agarose solution to solidify a bit. 25 ml of the agarose solution was then poured in 90 mm Petri-dishes and a special mold with lanes was then gently placed on top of the solution until it solidified. After removal of the mold, the injection plates, now with lanes, were stored at 4ºC and always warmed to 28ºC before injections.

To prepare injection needles, a glass capillary was placed in a needle puller where the midpoint of the capillary was melted with a platinum heating filament followed by pulling both ends of the capillary apart. Two needles with thin tips were then obtained for injections. The needles were usually loaded with around 3 µl of the injection mix using special thin and long tips before being on a micromanipulator. The needle tip was broken using forceps to allow, along with pressure conditions like pulse duration, a small droplet of 1 nl in size (measured using a microscale) to come out the needle. The fertilized eggs were then placed in the lanes of the injection plate and injected at the 1-cell stage. gRNAs, *Cas9* mRNA and other mRNAs were injected into the yolk along with phenol-red for visualization of the injection mix.

5.2.5. Cell culture

mESCs were maintained in PluriQ-ES-DMEM, composed of high-glucose DMEM supplemented with 2 mM glutamine, 1% non-essential amino acids, 100 U/ml penicillin, 100 µg/ml streptomycin, 15% ES cell-qualified fetal bovine serum (Millipore), 0.1 mM β-mercaptoethanol and 1,000 U/ml ESGRO (LIF) and 2i (3 µM CHIR99021 and 1 µM PD0325901). mESCs were grown on 0.1% gelatin-coated plates and split every other day. MEFs and MKFs were cultured in DMEM supplemented with 100 U/ml penicillin and 100 µg/ml streptomycin 10% bovine calf serum. All cells were grown at 37 °C, 5% CO2, 9% humidity, and experiments were performed on cells under 20 passages. To split cells, the medium was aspirated followed by a washing step of cell monolayer with PBS before incubating with 2.5% Trypsin-EDTA to detach the cells from the plate. 5 minutes later, serum-containing media was added to stop trypsin activity and cells were centrifuged down and split at 1:10. All cell lines tested negative for mycoplasma contamination.

5.2.6. PCR amplification

PCR reactions were assembled with high-fidelity or non-high-fidelity polymerases depending on experiment purposes as per the manufacturer’s protocol to amplify

coding sequences from cDNA or genomic regions from genomic DNA (gDNA). The common thermal cycler steps for the PCR reaction are as follows:

Step	Temperature	Time	Step description
1	95ºC	1 to 2 minutes, depending on polymerase used	Initial denaturation
2	95ºC	10 to 15 second	Denaturation
	57 to 65ºC, depending on primer melting temperature	15 or 30 seconds per kbp depending on polymerase used	Annealing
	68 or 72ºC depending on polymerase used	10 to 20 seconds	Extension
3	72ºC	10 minutes	Final extension
4	4 ºC	Indefinite time	hold

Table 26. Standard thermal cycler conditions for PCR.

5.2.7. Agarose gel electrophoresis

The PCR reaction products were mixed with gel loading dye and loaded on a 1% agarose gel where SYBR safe was added to label the DNA along with the appropriate DNA ladder (100 bp or 1kbp ladder). The reaction products were run on the electrophoresis gel for around 30-45 minutes at 160V before analyzing the gel with a blue light transilluminator and taking a picture of the gel using a UV light in a gel imager system.

5.2.8. Restriction digestion

DNA samples (such as plasmids) were mixed with 1 µl of the restriction enzyme of interest and 5 µl of the appropriate 10X buffer in a 50 µl reaction and incubated at 37ºC for 1 hour before running on an agarose gel.

5.2.9. DNA ligation

Ligation DNA templates and plasmids digested with similar restriction enzymes (thereby generating matching 5' and 3' overhangs sticky ends) was performed by the T4 DNA ligase as per the manufacturer's instructions. The DNA template, linearized plasmid, and T4 DNA ligase were incubated for 1 hour at room temperature before being transformed into DH5 alpha competent cells.

5.2.10. TA cloning

To sequence CRISPR/Cas9 induced mutations, PCR products were cloned into pGEM-T-easy vector through TA cloning. 3 µl of the PCR product was mixed with 1 ul of pGEM-T-easy vector, 1 ul T4 DNA ligase, and 5 ul of 2X rapid ligation buffer and incubated for at least 1 hour at room temperature before being transformed into DH5 alpha competent cells.

5.2.11. Cold fusion cloning

pCS2+ or pCDNA3.1 plasmids (destination vectors) were linearized using the restriction enzymes BamHI and XbaI. Forward and reverse primers used for PCR amplification of the desired insert were designed to have a 15 base pair homology sequence with the linearized vector ends, and the BamH1 or Xba1 restriction sequence, respectively, at their 5' ends. Following PCR amplification of the insert 1 µl (50-100 ng) of the PCR product was mixed with 1 µl of the linearized vector and 0.5 µl of the 5X cold fusion master mix. The mixture was then incubated 5 minutes at room temperature followed by 10 minutes on ice before being transformed into competent cells.

5.2.12. DNA purification from enzymatic reactions

PCR products or products of restriction enzymes digestion were purified using a GeneJET PCR purification kit as per the manufacturer's protocol. If agarose gel electrophoresis was required, the band of interest was cut out of the gel and purified using the GeneJet gel extraction kit as per the manufacturer's protocol.

5.2.13. Transformation into DH5 alpha competent cells.

Competent E. Coli cells were thawed on ice for a few minutes. 5 to 10 µl of the plasmid (from ligation and cloning reaction) were mixed with the competent cells and incubated on ice for 20 minutes. A heat shock was then applied by placing the

Eppendorf tube containing the mixture at 42ºC for 1 minute, followed by 2 minutes incubation on ice. 200 µl of SOC medium was then added to the tube then plated on the appropriate LB agar plate with the appropriate antibiotic for selection using glass beads and incubated at a 37ºC overnight.

5.2.14. Plasmid isolation from bacterial cells

Following transformation, a single bacterial colony was inoculated with 6 ml of LB medium with the appropriate antibiotic in a culture tube and incubated overnight at 37ºC in a bacterial shaker. Next morning, the bacterial culture was centrifuged at 4000 rotations per minute for 10 minutes before discarding the supernatant. The bacterial pellet was then used to isolate the plasmid using a GeneJet plasmid MiniPrep kit are per the manufacturer's instructions.

5.2.15. In vitro transcription

Zebrafish embryonic cDNA was used to PCR amplify cDNAs encoding *alcama*, *hif1ab*, *egfl7,* and *vegfaa* full-length mRNAs before being cloned into a pCS2+ vector by cold fusion cloning as described above. For overexpression experiments, the generated plasmids were sequenced to verify sequences. In vitro transcription using the mMESSAGE mMACHINE SP6 kit was performed as per the manufacturer's instruction after linearizing the respective plasmid using NotI. Briefly, 100 to 500 ng of the linearized plasmid was mixed with 10 µl 2x NTP/Cap mix, 2 µl 10x reaction buffer, and 2 µl of recombinant SP6 enzyme. The mixture was then incubated at 37ºC for 3 to 4 hours. To stop the reaction, 1 µl of Turbo DNAse was added to the mixture and incubated for 15 minutes at 37ºC. RNA was next purified by an RNA Clean and Concentrator kit as per the manufacturer's instructions. For overexpression experiments, 10 to 100 pg of mRNA was injected into one-cell stage embryos from heterozygous incrosses. Embryos were collected in TRIzol at 22-30 hpf for qPCR analysis.

For uncapped RNA transcription, cDNA was used to amplify zebrafish *hif1ab* and *vegfaa* and mouse Actb, whereas mouse *Cdk9* and *Sox9* were amplified from genomic DNA as it was not possible to amplify their sequences from cDNA owing to their low expression levels (only a single exon was amplified for *Cdk9*, while the full *Sox9* genomic locus was amplified). An adaptor sequence of 5′-GCCAAGCTATTTAGGTGACACTATAG-3′ was added to all reverse primers used for

amplification of cDNAs for uncapped experiments. Later, this adaptor sequence was used to detect the injected/transfected uncapped transcripts by qPCR. The amplified sequences were cloned into pCS2+ as described above and the obtained construct was linearized by NotI for in vitro transcription of uncapped RNAs. Alternatively, instead of cloning into a pCS2+ vector, a T7 sequence was added upstream of the forward primer used to amplify the cDNA and the PCR product was directly used for in vitro transcription of uncapped RNAs. To make XRN1-resistant uncapped transcripts, the XRN1-resistance sequence (Boehm et al., 2016) was cloned into pCS2+ upstream of the *hif1ab* and *vegfaa* coding sequences. In vitro transcription of uncapped transcripts was performed using SP6 RNA polymerase (Promega) (or T7 in case of the uncapped transcripts corresponding to the non-coding strand) through preparing a reaction mix that does not contain the Ribo-m_7G Cap Analog as per the manufacturer's protocol. 50 pg of RNA was injected into zebrafish embryos at the one-cell stage or 1 µg was transfected into MEFs or mESCs. Zebrafish embryos were collected in TRIzol at 6 hpf while cells were collected 6-24 hours post-transfection.

5.2.16. Transfection into mouse cell lines

FERMT2 and RELA overexpression plasmids were generated by cloning *Fermt2* and *Rela* cDNAs in pCDNA3.1 using cold fusion. 3.3 µg of each plasmid was transfected into the respective knockout cell line in 6 well plates using FuGENE 6 as per the manufacturer's protocol. 48 hours post-transfection, the cells were split into T75 flask and transfected cells were selected for using 0.5 µg/ml G418 for Fermt2 K.O. cells or 2 µg/ml for *Rela* K.O. cells. One week later, the selected cells were lysed in RIPA buffer to isolate proteins for western blot analysis.
Lipofectamine Messenger Max was used to transfect uncapped *Actb*, *Cdk9*, *Sox9,* and *eGFP* RNAs to MEFs and mESCs, while lipofectamine RNAiMax was used to transfect siRNAs.

5.2.17. Measurement of nucleic acid concentrations

DNA and RNA concentrations were measured using a Nanodrop spectrophotometer using 1 µl of the sample. Absorptions at 230, 260, and 280 nm were recorded and the concentration was calculated by the program as per the Lambert-Beer law. The quality of DNA and RNA was evaluated based on the 260/280 ratio (>1.8 for DNA and >2.0 for RNA) and 260/230 ratio (around 2.0 to 2.2 for both DNA and RNA).

5.2.18. DNA sequencing

0.5 to 1 µg of the DNA samples (such as plasmids or PCR products) was mixed 3 µM sequencing primer and sent to sequencing by SeqLab (Göttingen) using the Sanger chain termination method.

5.2.19. CRISPR/Cas9 mutagenesis

5.2.19.1. gRNA design

The online tool CHOPCHOP (http://chopchop.cbu.uib.no/) was used to design gRNAs (Table 22). To generate RNA-less alleles, two gRNAs flanking either the promoter region or the full gene locus were designed to generate promoter-less and full-locus deletion alleles, respectively. gRNAs aiming at deleting promoter regions were designed at least 500 bp upstream and downstream from the transcription start site (TSS) of the targeted gene.

5.2.19.2. Generation of zebrafish mutant lines

Zebrafish mutants were generated as previously described by the CRISPR/Cas9 system (Gagnon et al., 2014). Briefly, to synthesize gRNAs, 1 µl of a 100 µM oligo containing the T7 promoter and the gene-specific gRNA sequences (with the following sequence: TAATACGACTCACTATAggXXXXXXXXXXXXXXXXXXXXGTTTTAGAGCTAGAAATAGCAAG, where the multiple Xs represent the gene-specific gRNA sequence) and another 1 µl of a 100 µM constant oligo containing the gRNA scaffold sequence (with the following sequence: TAATACGACTCACTATAggagaaggtgaaggacactgGTTTTAGAGCTAGAAATAGCAAG) were mixed in a 10 µl solution. The solution was then placed in a thermal cycler to allow the oligos to anneal before filling out the overhangs using a T4 DNA polymerase as per the protocol shown below:

Step	Temperature	Time
1	95ºC	5 minutes
2	95ºC to 85ºC	A decrease by 2ºC every second
3	85ºC to 25ºC	A decrease by 0.1ºC every second

Pause		
To fill in the annealed oligos add the following reagents 2.5 µl dNTPs (10 µM) 2 µl 10X NEB buffer 2.1 0.2 µl 100X BSA 0.5 µl T4 DNA polymerase 4.8 µl dH_2O		
4	12ºC	20 minutes
5	4ºC	Indefinite (hold)

Table 27. Protocol for annealing and filling in the gRNA synthesis template.

The product was then purified using the GeneJet PCR purification kit and eluted in 30 µl dH_2O. The sample was then run on a gel to confirm appropriate size before using 6 µl to synthesize the gRNA using the T7 mMessage mMACHINE kit.

1-cell stage zebrafish embryos were then co-injected with 50 pg of gRNA and 100 pg of *Cas9* mRNA. To generate RNA-less alleles, two gRNAs were co-injected with the *Cas9* mRNA. Injected embryos were then raised to adulthood.

To identify founders (i.e., F0 fish that passed the mutation through germline), individual F0 fish were outcrossed to wild-type fish and the embryos were genotyped by HRMA or PCR. Once founders were identified, they were outcrossed with wild-type lines and embryos were raised to obtain F1 fish which are fully heterozygous mutants. Experiments were performed on embryos and larvae from obtained from F2 or later heterozygous parents.

5.2.19.3. Generation of mouse mutant cell lines

gRNAs targeting *Rela* and *Actg1* were cloned into the PX458 vector as previously described (Ran et al., 2013). Briefly, a forwarded and a reverse oligo of the gRNA sequence were designed with a 5'TGGG overhang sequence added to the forward oligo and a 3'AAAC added to the reverse oligo. The oligos were then annealed according to steps 1-3 of Table 25 then ligated to a BbsI-linearized PX458 vector. The final plasmid expressed both the gRNA targeting the gene of interest, Cas9, and eGFP as a selection marker (hereafter referred to as nuclease plasmid). 5 µg of the nuclease plasmid(s) was then electroporated into MEFs through nucleofection (Lonza) according to the manufacturer's protocol. Cells expressing eGFP were then

subjected to single-cell sorting into 96-well plates, two days following transfection using a FACSAria III sorter. Three weeks later, growing clones were split into two 48-well plates, where one plate was used to propagate the clones and the other to isolate genomic DNA from the clones for genotyping purposes by PCR and sequencing. Clones of the desired genotype were then propagated for further experiments.

5.2.20. Genomic DNA and RNA isolation and purification

For genotyping adult zebrafish, fish were anesthetized then a small part of the caudal fin was clipped by scissors and placed in 0.2 ml tubes. DNA was then extracted using the NaOH/Tris method. In brief, samples were boiled at 95°C for 10 minutes then Tris with pH 8 was added at 1/10 volume. The extracted DNA was then used for high-resolution melt analysis (HRMA) or PCR-based genotyping.

To genotype zebrafish from heterozygous fish incross for qPCR experiments, DNA and RNA were co-isolated from at least 24 embryos or larvae using TRIzol followed by phenol-chloroform extraction. Samples were homogenized in TRIzol using a Next Advance Bullet Blender homogenizer and chloroform was added followed by vortexing and centrifugation to allow phase separation. The top RNA-containing aqueous phase was isolated and stored at −80 °C, while the bottom DNA-containing organic phase was subjected to ethanol purification to purify the DNA. In brief, 100% ethanol was added to the organic phase and centrifuged following vortexing. The ethanol-organic phase mixture was then aspirated and the DNA pellet was washed with 80% ethanol before being aspirated and leaving the pellet to dry shortly at room temperature. The purified DNA pellet was then dissolved in water and genotyping was performed using HRMA or PCR. RNA from genotyped wild-type (+/+) and homozygous mutant (−/−) embryos or larvae were then pooled separately and purified using the Zymo RNA clean and concentrator kit as per the manufacturer's protocol. For each experiment, this process was performed on larvae or embryos from three different crosses at least.

For mouse cell lines, DNA was extracted using the Quick-DNA 96 kit, as per the manufacturer's protocol, or with the NaOH/Tris method following trypsinization and centrifugation. The isolated DNA was then used for PCR-based genotyping. RNA was isolated using TRIzol followed by phenol-chloroform extraction. Following phase separation, RNA was purified from the top aqueous phase using the Zymo RNA clean and concentrator kit.

5.2.21. Genotyping

5.2.21.1. PCR followed by HRMA

PCR followed by HRMA was used to genotype zebrafish mutants with small insertions or deletions. Primer pairs for HRMA were to amplify an amplicon <100 bp in size. A primer mix solution was made by adding 1 µl of each forward and reverse primer (100 µm) to 140 µl of dH_20. The following mixture was then assembled for PCR followed by HRMA: 5 µl SYBR green, 4.5 µl primer mix, and 0.5 µl of the gDNA containing solution. Following PCR, the machine performed HRM analysis to separate the PCR products based on their melting temperature, thereby allowing the identification of different genotypes. The PCR and HRMA steps are shown below:

Step	Temperature	Time	Description
1	95ºC	7 minutes	Polymerase activation
2	95ºC	10 seconds	PCR cycle (Denaturation and annealing step), 35 cycles
	60ºC	15 seconds	
3	95ºC	15 seconds	HRMA
	55ºC to 95ºC	Gradual increase in temperature every 1 second	
	95ºC	15 seconds	

Table 28. Thermocycler conditions used PCR followed by HRMA.

5.2.21.2. PCR followed by gel electrophoresis

PCR followed by gel electrophoresis was used to genotype zebrafish and mouse cell line mutants with large deletions such as RNA-less alleles. A primer pair designed outside of the deletion region was used to differentiate wild-type, heterozygous and homozygous mutant fish (out-out PCR). Following PCR and gel electrophoresis, the wild types will display a single band large in size, the homozygous mutants will display a single band smaller in size while heterozygous animals will display both bands. For very large deletions (>4 kbp), since an out-out PCR would not be able to

amplify an undeleted wild-type allele, a primer pair binding inside of the deletion region was used to differentiate between heterozygous and homozygous mutants. Wild types were identified as those showing no band on a gel following an out-out PCR, heterozygous mutants being those that show a band both with out-out and inside PCRs and homozygous mutants being those showing no band with an inside PCR. To amplify repetitive regions such as promoter and intronic regions, the high fidelity SeqAmp polymerase was used.

5.2.22. In vitro generation of synthetic uncapped transcripts containing different sequences of *hif1ab* mRNA

120 nt oligonucleotides containing *hif1ab* cDNA sequences sharing sequence similarity to *epas1a* genomic locus were ordered from Sigma and ligated together. The oligos were designed that the odd numbered ones had the sense strand sequence while the even numbered ones had the antisense strand sequence. Each oligo had 5' and 3' 20 nucleotide (nt) overhangs with the previous and following oligo. Oligos were annealed and filled out as indicated in Table 25. The first oligo included a 5' T7 promoter sequence that was important for invitro transcription of the uncapped RNAs using the Promega T7 RNA polymerase. Similar sequences were determined using BLASTn analysis that was highly sensitive (an *E* value of up to 1,000,000 and word size of 7). The same approach was used to generate uncapped RNAs of *hif1ab* cDNA sequences not sharing sequence similarity with *epas1a* genomic locus.

5.2.23. cDNA synthesis

Following RNA isolation by TRIzol and purification, 0.5 to 1 µg of RNA was used for cDNA synthesis (in some cases where the analyzed transcript was expected to have very low expression levels (like anti-sense RNAs) 5 µg of RNA was used). The Maxima first strand cDNA synthesis with dsDNase kit was used to synthesize the cDNA as per the manufacturer's protocol. Briefly, the appropriate amount of RNA was mixed with 1µl of dsDNase and 1 µl 10x DNase buffer in a 10 µl reaction and incubated at a 37ºC for 2 minutes to digest any DNA contamination. 2 µl of Maxima enzyme mix (that included the M-MuLV reverse transcriptase and an RNase inhibitor), 4 µl of a 5X reaction mix (that included reaction buffer, dNTPs, random hexamer primers and oligo dT) and 4 µl of dH_2O were added to the reaction mix before being incubated in a thermal cycler with the following program:

Step	Temperature	Time
1	25ºC	10 minutes
2	50ºC	30 minutes
3	85ºC	5 minutes
4	4ºC	Indefinite

Table 29. Thermocycler conditions for cDNA synthesis.

The cDNA was then directly used for qPCR analysis or stored at -20ºC for future use.

5.2.24. Gene expression analysis by quantitative PCR

A CFX Connect Real-Time System (Biorad) was used for qPCR analysis. Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) was used to design qPCR primers that amplify an amplicon of <200 bp in size. The PCR reaction was assembled similar to the reaction mix used for PCR followed by HRMA analysis (section **5.2.21.1**) except that cDNA was used instead of gDNA. During each PCR cycler, the fluorescent-dye in the SYBR green polymerase intercalated with the amplified dsDNA, allowing a real-time quantification. Thermal cycler conditions resembled that used for PCR followed by HRMA analysis (Table 26).
Three biological replicates were used to interpret results and reactions were performed in at least technical duplicates. For zebrafish experiments, qPCR was performed at the embryonic or larval stage where the respective studied mutated gene displayed its highest expression levels in wild types.
Primers to detect mRNAs were designed mostly around exon-exon junction, while those for pre-mRNA were designed around intron-exon boundaries. Several primer pairs, detecting different regions of the gene's cDNA, were used to assess transcript levels in promoter-less alleles generated. Only mutant alleles showing transcription less than 10% with any of the primer pairs were further used as promoter-less alleles. Allele-specific primers were designed to amplify just the wild-type allele but not the mutant one. To determine the injected capped and uncapped transcript levels, the adaptor sequence at the 3′ end of the injected transcripts was used to design a universal reverse primer that was used along with forward primers designed to be in close proximity; distinguishing between endogenous and injected transcripts was

thereby facilitated in that way. *rpl13* and *gapdh*, besides *Actb, Gapdh* and *Rn18s* were used as house-keeping genes to normalize zebrafish and mouse analyses, respectively.

The $2^{-\Delta\Delta Ct}$ method was used to calculate fold changes. In brief the house keeping gene's Ct values was deducted from the of the analyzed gene and to have the ΔCt value for each sample. ΔΔCt was then calculated by subtracting the average of ΔCt values obtained from control experiments from that of the experimental sample. Fold changes were then calculated as $2^{-\Delta\Delta Ct}$. Ct values observed in this study were between 21 and 27 for any mRNA expression analysis or between 27 and 29 for pre-mRNA expression analysis and between 12 and 23 for housekeeping genes except for *Rn18s* where the values were between 6 and 8.

Data from zebrafish *upf1* double mutants were analyzed in a way where the adapting genes' expression levels double mutants are shown relative to their expression in *upf1* single mutants. Mutated genes' expression levels in *upf1* double mutants were, on the other hand, shown relative to their expression in *hbegfa*, *vcla* or *vegfaa* single mutants. For siRNA experiments, adapting genes' expression levels in knockout cells were shown relative to their expression levels in wild-type transfected with the same siRNA. On the other hand, mutant gene expression levels in knockout cells transfected with a particular siRNA are shown relative to their expression levels in knockout cells transfected with scrambled (Scr) siRNA. Similarly, for NMD inhibition by NMDi14 in zebrafish or cycloheximide in cell lines, adapting genes' expression levels in mutant fish or knockout cells were shown relative to their levels in wild-type fish or cells treated with the same drug. Mutant gene expression levels in mutant fish or knockout cells treated with a particular drug are shown relative to their expression in mutant larvae that were untreated, or DMSO-treated knockout cells.

5.2.25. RNA interference

To silence genes of interest (Table 19), MEFs and mESCs were transfected with siRNA duplexes targeting the genes of interest using the Lipofectamine RNAiMAX. Briefly, 200,00 cells were seeded in fully supplemented medium a day prior to the transfection day in 6-well plates. The following day, the siRNA(s) of interest was mixed with the transfection reagent and incubated for 20 minutes before applying the mixture to the cells to obtain a final concentration of 10 nM for the siRNA in the cell culture medium (except when knocking down XRN1 in mESCs, where a lower concentration of siRNA was used (2.5 nM) was used as higher concentrations led to

disruption of housekeeping genes transcript levels as well). In most cases cells were collected in TRIzol for qPCR analysis 2 days post transfection except for XRN1, SMG6 and ERF1 knockdowns where cells were collected 1 day post transfection. In most cases, the siRNAs led to knockdown efficiencies of 70–90% (except for XRN1 knockdown in mESCs where the 2.5 nM duplex led to just 20% knockdown of *Xrn1* transcript levels). A scrambled (Scr) siRNA (Sigma, SIC00), that do not target any of the mouse RNAs, was used as negative control.

5.2.26. CRISPR interference (CRISPRi)

To knockdown *Fermt2* transcription, 3 gRNAs targeting the promoter and transcription start sites of *Fermt2* were designed and cloned into a plasmid encoding a catalytically dead Cas9 fused to the Krüppel-associated box (KRAB) repressor and eGFP (Addgene, 71237) as described previously (Thakore et al., 2015). The three plasmids were transfected by nucleofection (Lonza) to *Fermt2* K.O. cells were and forty-eight hours later, cells positive for eGFP were sorted using a FACSAria III cell sorter into TRIzol for qPCR analysis.

5.2.27. Drug treatments to inhibit RNA decay

NMDi14 (Martin et al., 2014) was used to inhibit NMD in wild-type and *hbegfa* mutant zebrafish. 72 hpf wild-type and mutant larvae were raised in egg water with 10 µM NMDi14, or DMSO as control, for three days later before they were collected at the 6 dpf stage in TRIzol for RNA extraction. The data was analyzed at the 6 dpf stage, rather than at 72 hpf (as were most of the analyses on *hbegfa* mutants in the study) because the drug was only effective in NMD inhibition when fish were treated for 3 days (in this case between 3 and 6 dpf) and earlier treatment with the drug was not possible due to toxicity (no gross morphological alterations were observed in larvae treated with the drug between 3 and 6 dpf).
Cycloheximide (Sigma) was used to inhibit RNA degradation through blocking mRNA translation. Wild-type and *Rela* K.O. MEFs were treated with 200 µg/ml cycloheximide, or DMSO as a control, for 5 hours before being collected in TRIzol for qPCR analysis.

5.2.28. mRNA half-lives quantification

Transcription inhibition by actinomycinD (Sigma) was used to asses mRNA half-lives. In brief, 10 µg/ml actinomycin D was used to added to cultured wild-type, *Fermt2*

K.O., *Actg1* K.O. and *Rela* K.O. cells which were then collected in TRIzol for qPCR analysis at 0, 1, 2, 4 and 8 hours post treatment. Expression levels of *Fermt2*, *Actg1* and *Rela* mRNA levels were assessed by qPCR analysis at each time point relative to their expression levels at the 0-hour time point. Half-lives were then quantified from fitted nonlinear exponential decay curves. *Rn18s* was used for normalization purposes (as housekeeping gene), as its expression level was not affected across the treatment time course.

5.2.29. RNA metabolic labeling

RNA metabolic labelling was performed as described previously (Radle et al., 2013; Sun and Chen, 2018). Briefly, wild-type, *Fermt2* K.O., *Actg1* K.O. and *Rela* K.O. cells were treated with 200 μM 4-thiouridine (4sU) for 1 h to label newly transcribed RNAs, before extracting RNA using phenol–chloroform extraction. 80 μg of the isolated RNA was then incubated with biotin-HPDP (Thermo) to biotinylate the4sU-labelled, newly transcribed transcripts. The μMacs Streptavidin Kit (Miltenyi) was then used to pull down the biotinylated transcripts. The RNA containing solution was incubated with streptavidin-coated magnetic beads for 15 minutes at room temperature, thereby enabling biotinylated transcripts to be specifically isolated when the magnetic beads where separated from the solution using a magnetic stand. Multiple washing steps were performed before eluting the biotinylated RNAs using 100 mM DTT. The pulled down RNA was then purified using RNeasy MinElute clean up kit (Qiagen) and at least 100 ng of the purified RNA was then used for cDNA synthesis and qPCR analysis. This experiment was performed only once.

5.2.30. Cytotoxicity analysis

Wild-type, *Rela* K.O. and *Rela*$_{promoter\text{-}less}$ cells were plated in 96-well plates (around 7,000 cells per well) and incubated with culture media supplemented with 25 ng/ml of mouse TNF. 24 hours later, cells incubated for five hours with media containing 3 mg/ml MTT, or DMSO as control, after a washing step with PBS. Viable cells are able to metabolize MTT to purple-colored Formazan crystals which can be used to assess cell viability after dissolving it in a 50% DMSO:50% ethanol solution. Optical density (O.D.) of the formazan solution in each well was measured using a FLUOstar Omega spectrophotometer at 572 nm to assess cell viability. The following formula: (O.D. DMSO − O.D. TNF)/O.D. DMSO, was then used to assess percent cytotoxicity. This experiment was performed only once.

5.2.31. mESCs staining

4% paraformaldehyde was used to fix cells for 15 min at room temperature. Cells were then permeabilized for 10 minutes using 0.3% Triton X-100 in PBS before incubation with phalloidin Alexa-568 (1:1000; to stain actin filaments) in 3% BSA at room temperature for one hour. After three washes with PBT of 15 minutes each, samples were incubated DAPI (1:5,000) for 5 minutes at room temperature before being mounted for imaging with Dako fluorescent mounting medium. Following imaging, Actin filaments protrusion length was measured using ImageJ.

5.2.32. Zebrafish larvae immunostaining

100 hpf larvae from heterozygous *alcama*$^{promoter\text{-}les}$ and *alcama*$^{\Delta 8}$ incrosses were fixed overnight in 4% paraformaldehyde. Larvae were then washed trice (each for 15 minutes) with 0.1% PBST, before being permeabilized for 1 hours using 3 µg/ml proteinase K. Following permeabilization, larvae were washed with PBDT trice (each for 15 minutes) then incubated for two hours at room temperature with phalloidin Alexa-568 (Invitrogen) to label actin filaments. After 6 washing steps (of 15 minutes each) with 0.1% PBST, larvae were mounted for imaging the heart using the confocal LSM880 microscope. Distance from the apex of the ventricle to the junction of the ventricle with the bulbus arteriosus was measured using Zen Black to assess the ventricle length (long axis of the ventricle).

5.2.33. Confocal microscopy

Imaging using confocal laser scanning microscopy such as Zeiss LSM700, LSM880 or a Zeiss spinning disc microscopes was performed to assess phenotypes of different mutant zebrafish and mouse cell lines. Live imaging of the trunk and brain vasculature in wild-type and *egfl7*$^{full\text{-}locus\ del.}$, *vegfaa*$^{\Delta 10}$ and *vegfaa*$^{promoter\text{-}less}$ mutant zebrafish embryos was performed using the LSM700 confocal microscope after embryos were anesthetized using a low dose of tricaine (0.04%) and placed in glass bottom Petri dish after being immobilized using 1.2% low-melting agarose. Plan-Apochromat 10×/0.45 and LCI Plan-Neofluar 25×/0.8 objectives were used. The LSM700 was also used to image fixed and immunostained mESCs using an LD C-Apochromat 63×/1.15 W Corr M27 objective. LSM880 was used to stain fixed and immunostained zebrafish larval heart using a W Plan-Apochromat 20×/1.0 objective.

Time-lapse imaging of the trunk region was performed a Zeiss Spinning disc CSU-X1 high-speed camera confocal microscope to measure blood flow velocity in 78 hpf wild-type, maternal zygotic hbegfa$^{\Delta 7-/-}$ and maternal zygotic hbegfa$^{\text{full-locus del.}-/-}$ zebrafish larvae. Short videos were made and the time required by erythrocytes to move 200 µm in the dorsal aorta at the level of the 5th and 6th somites was quantified to measure the blood flow velocity in the different genotypes, using Zen Blue (as described previously in (Kwon et al., 2016)).

5.2.34. Image processing and analysis

Images and videos obtained were processed and analyzed using software such as Zen, Imaris and ImageJ.

5.2.35. Western blotting

Western blot analyses were performed using precast gradient gels according to standard protocols. In brief, cells were lysed in modified RIPA buffer to which protease inhibitors (cOmplete ULTRA Mini, Roche) and phenylmethylsulfonyl fluoride (PMSF) were added. Protein concentration of the cell lysates was measured using the Bradford protein assay and 35 µg were separated on precast TGX gradient gels using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). The Transblot Turbo Transfer System was then used to transfer proteins electrophoretically to polyvinylidine fluoride membranes (Bio-Rad). Membranes for FERMT2 or RELA were then blocked using 5% non-fat milk or 5% BSA, respectively, for one hour before being probed overnight at 4°C with primary antibodies (Probing for ACTB was also included as a loading control). On the next day, the membranes were washed 3 times (each for 15 minutes) with PBST the peroxidase-conjugated secondary antibodies were applied to the membranes for one hour at room temperature. The membranes were again washed 3 times with PBST then bands were visualized by chemiluminescence after membranes were incubated with enhanced chemiluminescence (ECL) substrate (Clarity Western ECL Substrate, Bio-Rad) and imaged by a ChemiDoc MP system. This experiment was performed only once.

5.2.36. Chromatin immunoprecipitation (ChIP)

Approximately 30 million cells were fixed and nuclei were isolated for chromatin shearing using the truChIP Chromatin Shearing Reagent kit (Covaris) as per the

manufacturer's protocol. Briefly, cells were fixed with 11.1% formaldehyde for 10 minutes at room temperature before being quenching the fixative with the kit's quenching buffer for 5 minutes. Cells were incubated with the kit's cell lysis buffer to isolate nuclei that was then washed multiple times before being resuspended with shearing buffer and transferred to thin walled 500 µl Eppendorf tubes (150 µl resuspended nuclei in each tube) in preparation for chromatin shearing. Chromatin shearing was performed using Bioruptor (Diagenode) with 4-5 cycles of 15 seconds ON and 15 seconds OFF for 15 minutes each cycle, generating fragments of 200–400 bp in size. Immunoprecipitation was then performed as previously described (Blecher-Gonen et al., 2013). Briefly, sheared chromatin was incubated with antibodies targeting the proteins of interest (Table 3), in addition to IgG as control, overnight at 4°C on a rotor with slow rotation. The next day, pre-blocked Protein A/G-coated magnetic beads were added to each tube and incubated for 1 hour overnight at 4°C to allow the antibodies constant region to attach to the magnetic beads. Next, the beads were separated from the solution using a magnetic stand and washed twice with each of the following buffers in order: 1) ChIP low salt buffer 2) ChIP high salt buffer 3) ChIP high salt buffer 4) TE buffer. To elute the immunoprecipitated chromatin, the beads where incubated with ChIP elution buffer at 65°C for 15 minutes. The tubes were placed on the magnetic stands and the supernatant was isolated before another elution step was performed. Reverse cross-linking was then performed by incubating the samples at 65°C overnight. The following day, the samples were incubated for 1 hour with RNase A (Thermo) at 37°C followed by incubation with 100 µg proteinase K for 3 hours at 50°C. The immunoprecipitated DNA was then purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel), as per the manufacturer's instruction for sodium dodecyl sulfate-containing samples. H3K4me3 and WDR5 ChIP-qPCR experiments were performed using primers designed at the transcription start sites of the respective genes. Primers within the gene body (exonic or intronic regions) were also used as controls. Data were analyzed as fold enrichment over IgG. ChIP–qPCR experiments on *Rela* knockout cells following UPF1/EXOSC4 or XRN1 knockdown were performed once.

5.2.37. ATAC-seq material extraction and library preparation

After cells were trypsinized and washed with PBS, they were counted with MOXI Z Mini Automated Cell Counter Kit (Orflo) and the ATAC library was prepared from

50,000 cells using the Tn5 transposase (Illumina). Briefly, the cell pellet was resuspended in 50 µl PBS and 25 µl tagmentation DNA (TD) buffer, 2.5 µl Tn5, 0.5 µl 10% NP-40 and 22 µl water and incubated for half an hour at 37°C with intermittent snap-mixing. Next, the samples were incubated at 50 °C for half an hour with 500 mM EDTA pH 8.0 for recovery of the digested DNA fragments. 100 µl of 50 mM MgCl2 was then added to neutralize EDTA before purifying the DNA fragments using the MinElute PCR purification kit. Amplification and indexing of the library were performed as described previously (Buenrostro et al., 2013). Equimolar ratios of the libraries were then mixed and sequenced on a NextSeq500 platform using v2 chemistry. ATAC-seq data were deposited on the Gene Expression Omnibus under the accession code GSE107075.

5.2.38. ATAC-seq analysis

FastQC (http:// www.bioinformatics.babraham.ac.uk/projects/fastqc) was used to assess sample quality. Trimmomatic v.0.3350 was utilized to trim reads after a quality drop below a mean of Q20 in a five nucleotides window. Only reads >30 nucleotides were cleared for further analysis. 27 million reads per sample were randomly selected for further analysis, in order to normalize all samples to a similar sequencing depth. Reads were mapped against the mm10 (GRCm38) Ensembl mouse genome version with STAR 2.4.2a (Dobin et al., 2013) using only unique alignments to eliminate reads with unclear placing. To avoid PCR artefacts leading to multiple copies of the same original fragment, reads were additionally deduplicated using Picard 1.136 (http://broadinstitute.github.io/picard/). To identify peaks, the MACS2 peak caller v.2.1.0 was utilized. The minimum q value was fixed to −1.5 and the false discovery rate was changed to 0.01. To determine thresholds for significant peaks, the data were inspected manually on the Integrated Genomics Viewer (IGV; (Robinson et al., 2011)). Peaks overlapping known mis-assemblies, satellite repeats and other blacklisted region from ENCODE were eliminated. To compare peaks between different samples, the final lists of peaks that were significant were overlapped and unified to represent identical regions. Counts per unified peak per sample were computed with bigWigAverageOverBed (UCSC Genome Browser Utilities, http://hgdownload.cse.ucsc.edu/downloads.html) following conversion of binary alignment map (BAM) files to bigWig format with deep Tools bamCoverage (Ramirez et al., 2014). For normalization, raw counts for unified peaks were submitted to DESeq2 (Anders and Huber, 2010). Spearman correlations were

produced using R to identify degree of reproducibility between samples. To allow a normalized display of samples on IGV, the raw BAM files were normalized for noise level (number of reads inside peaks) and sequencing depth (number of mapped deduplicated reads per sample). Two factors were computed and applied, using bedtools genomecov (Quinlan and Hall, 2010), to the original BAM files, resulting in normalized bigWig files that can be visualized on IGV.

5.2.39. RNA sequencing (RNA-seq)

The miRNeasy micro Kit (Qiagen) was used to isolate RNA. On-column DNase digestion (DNase-Free DNase Set, Qiagen) was performed on samples to remove any contaminating DNA. RNA and library integrity were verified on the LabChip Gx Touch 24 (Perkin Elmer). 1 µg of the total RNA was utilized as input for the SMARTer Stranded Total RNA Sample Prep Kit–HI Mammalian. RNA sequencing was performed on a NextSeq500 instrument using v2 chemistry, generating an average of 25–30 million reads per library, with a 1 × 75 bp single-end setup. RNA-seq data were deposited on the Gene Expression Omnibus under the accession code GSE114212.

5.2.40. RNA sequencing analysis

FastQC was used to assess the resulting raw reads were quality, duplication rates and adaptor content. To trim reads after a quality drop below a mean of Q20 in a 10 nucleotides window, Reaper v.13-100 (Davis et al., 2013) was used. For subsequent analysis, only reads of at least 15 nucleotides were cleared. STAR 2.5.3a was used to align trimmed and filtered reads were against the mm10 (GRCm38) Ensembl mouse genome version using with the parameters '−outFilterMismatchNover- Lmax 0.1−alignIntronMax 200000' (Dobin et al., 2013). featureCounts 1.6.0 from the Subread package (Liao et al., 2014) was used to count the numbers of reads that aligned to genes. To admit a particular read, it had to map at least partially inside an exon; these reads were aggregated per gene. Reads aligning to multiple regions or overlapping with several genes were eliminated. DESeq2 v.1.14.1 (Love et al., 2014) was used to identify differentially expressed genes. $P < 0.05$ (Wald test) was used to classify a gene as significantly differentially expressed, without assigning specific maximum or minimum threshold for fold change as transcriptional adaptation might not necessarily cause strong transcript upregulation levels. The Ensembl annotation

was enriched with UniProt data (release 24 March 2017) based on Ensembl gene identifiers.

5.2.41. Gene set enrichment analysis

Genes that were commonly highly upregulated in *Fermt2*, *Actg1* and *Actb* K.O. cells compared to their respective wild-type controls (log2(expression level in K.O. cells/gene expression level in WT cells) >0.585, $P < 0.05$ (Wald test)) were used for gene-set enrichment analyses using KOBAS (Xie et al., 2011).

5.2.42. Sequence similarity and subsampling analysis

To identify genes exhibiting sequence similarity to *Fermt2*, *Actg1* and *Actb*, the longest respective transcript was chosen (ENSMUST00000071555, ENSMUST00000045905 and ENSMUST00000100497, respectively) and compared to the whole genome using BLASTn (Altschul et al., 1990). Genes were identified to be sharing sequence similarity to the respective gene's mRNA when a partial match was identified within the gene body or promoter region (that was identified as 2 kb upstream of the transcription start site). To identify the optimal degree of similarity, multiple alignment parameters were assessed. Alignment length, bit score and *E* value were queried to determine the optimal values and dynamic range using subsampling analysis (as described below). Bit score is a combined assessment of alignment quality and length. *E* value denotes the possibility that a given match resulted by chance, upon considering the whole target database (the genome in this case). For each K.O. cell line, a subsampling approach was used to calculate a ranked *P* value for the significance of the percentage of upregulated genes in K.O. cells compared to WT (*P* value < 0.05) in subsamples of a specific size (equivalent to the number of protein-coding genes sharing sequence similarity). Briefly, this algorithm was repeated ten-thousand times: 1) select X random protein-coding genes, 2) identify the percentage of significantly upregulated genes in this subsample. The resultant list was filtered for subsamples with equal or higher than expected number/percentage of upregulated genes according to a preceding comparison (e.g., for *Fermt2*, 18 protein-coding genes display sequence similarity to its mRNA (=subsample size), 9 of which were also upregulated (=expectation)). The number of subsamples displaying at least as many upregulated genes as the expectation depict the comparison rank. The ranked *P* value was computed by dividing the rank by the overall number of iterations (= 10,000).

Optimal thresholds varied for the different K.O. cell line models, ranging between 1) an alignment length between 20 and 180 nucleotides; 2) a bitscore between 40 and 200; and 3) a maximum *E* value between 10 and 6.73×10^{-50}. The following maximum *E* values from the optimal range were selected for follow-up similarity analysis: 5.1 for *Fermt2* and *Actg1*, and 2×10^{-48} for *Actb*. A stricter *E* value for *Actb* was necessary due of its repetitive 3′UTR, that lead to 'noisy' misleading matches. Such *E* value thresholds translate to local nucleotide sequence alignments ranging between 24 and 1,901 nucleotides in length, with 75% to 96% identity.

5.2.43. Sequence alignments of *hif1ab* and *epas1a*, besides *actb1* and *Actb*

Kablammo (Wintersinger and Wasmuth, 2015) was used to visualize BLASTn alignments of *hif1ab* longest transcript (ENSDART00000018500) with *epas1a* gene body and promoter (2 kb upstream of the transcription start site) using word size 7 and an *E* value of 25. MUSCLE (Edgar, 2004) was used to align the synthetic *hif1ab* transcript to the original source transcript (ENSDART00000018500), while BLASTn (Zhang et al., 2000) was used to show its alignment to *epas1a* gene body and promoter. The uncapped transcript composed only of the *hif1ab* sequences that are share sequence similarity to *epas1a* was 1,277 nucleotides in length while that composed only of *hif1ab* sequences not sharing sequence similarity to *epas1a* was 1,929 nucleotides in length. The alignment between the zebrafish *actb1* transcript ENSDART00000054987 (query) coding sequence to that of the mouse *Actb* transcript ENSMUST00000100497 (subject) was assessed using MUSCLE.

5.2.44. Statistics and reproducibility analyses

No statistical methods were used to predetermine sample size and experiments were not randomized. Statistical analysis was performed using an unpaired, two-tailed Student's *t*-test. A *P* value less than 0.05 was considered statistically significant. Experiments and results analyses were not done blindly. All experiments were performed at least twice unless otherwise indicated.

6. Results

Parts of this chapter have been published as an article in the journal Nature (El-Brolosy et al., Nature. 2019; 568, 193-197).

The author's contribution was described as follows in the paper: “M.A.E.-B. designed and performed most of the experiments, analyzed the data and wrote the manuscript; Z.K. and A.R. designed and performed mESC experiments and some imaging and edited the manuscript; C.K. performed bioinformatics analyses; S.G. performed ATAC-seq and RNA-seq; N.F. generated some zebrafish mutants and performed some qPCR experiments; K.K. performed some qPCR experiments; G.L.M.B. performed some imaging; C.M.T. generated the *upf1* mutant, under the supervision of A.J.G.; S.-L.L., R.F. and C.G. provided unpublished mutants; and D.Y.R.S. helped to design the experiments and analyze data, supervised the work and wrote the manuscript. All authors commented on the manuscript.”

6.1. Transcriptional adaptation models in zebrafish and cultured mouse cells.

Gene duplication and whole-genome duplication events were one of the main drivers of evolution, as they lead to new genetic material that can be subsequently modified by the process of natural selection (De Grassi et al., 2008; Ohno, 1970). Besides, their role in evolution, they play a central role in maintaining genetic robustness. Gene duplication events lead to the formation of duplicated genes that can have overlapping functions (known as paralogs; or ohnologues if they arise from whole genome-duplication events). Paralogous genes have a higher potential to functionally compensate for each other’s loss (Diss et al., 2014; Gu et al., 2003; Hsiao and Vitkup, 2008; Su et al., 2014). Zebrafish is an ideal model for my analysis as it had undergone an additional event of whole-genome duplication around 300 million years ago (Glasauer and Neuhauss, 2014). I, thereby, decided to start my investigation on the molecular mechanisms underlying genetic compensation by analyzing the expression levels of paralogous genes besides other related genes (hereafter referred to as adapting genes) in several zebrafish and mouse mutants that harbor a PTC or have their last exon deleted. To this end, I made use of the previously published PTC-containing *vegfaa*, *egfl7* (Rossi et al., 2015), *hif1ab* (Gerri et al., 2017) zebrafish mutants, besides *hbegfa*, *alcama* and *vcla* mutants that were previously generated in the lab (Figure 3). In addition, I analyzed the previously

published last exon-deleted *Fermt2* knockout mouse kidney fibroblasts (MKFs) (Theodosiou et al., 2016), PTC-containing *Rela* knockout mouse embryonic fibroblasts (MEFs) ((Gapuzan et al., 2002); clone 1 in (Gapuzan et al., 2005)), besides PTC-containing *Actb* knockout mouse embryonic stem cells (mESCs) that were generated by colleagues in the lab, in addition to last exon-deleted *Actg1* knockout MEFs that I generated (Figure 3).

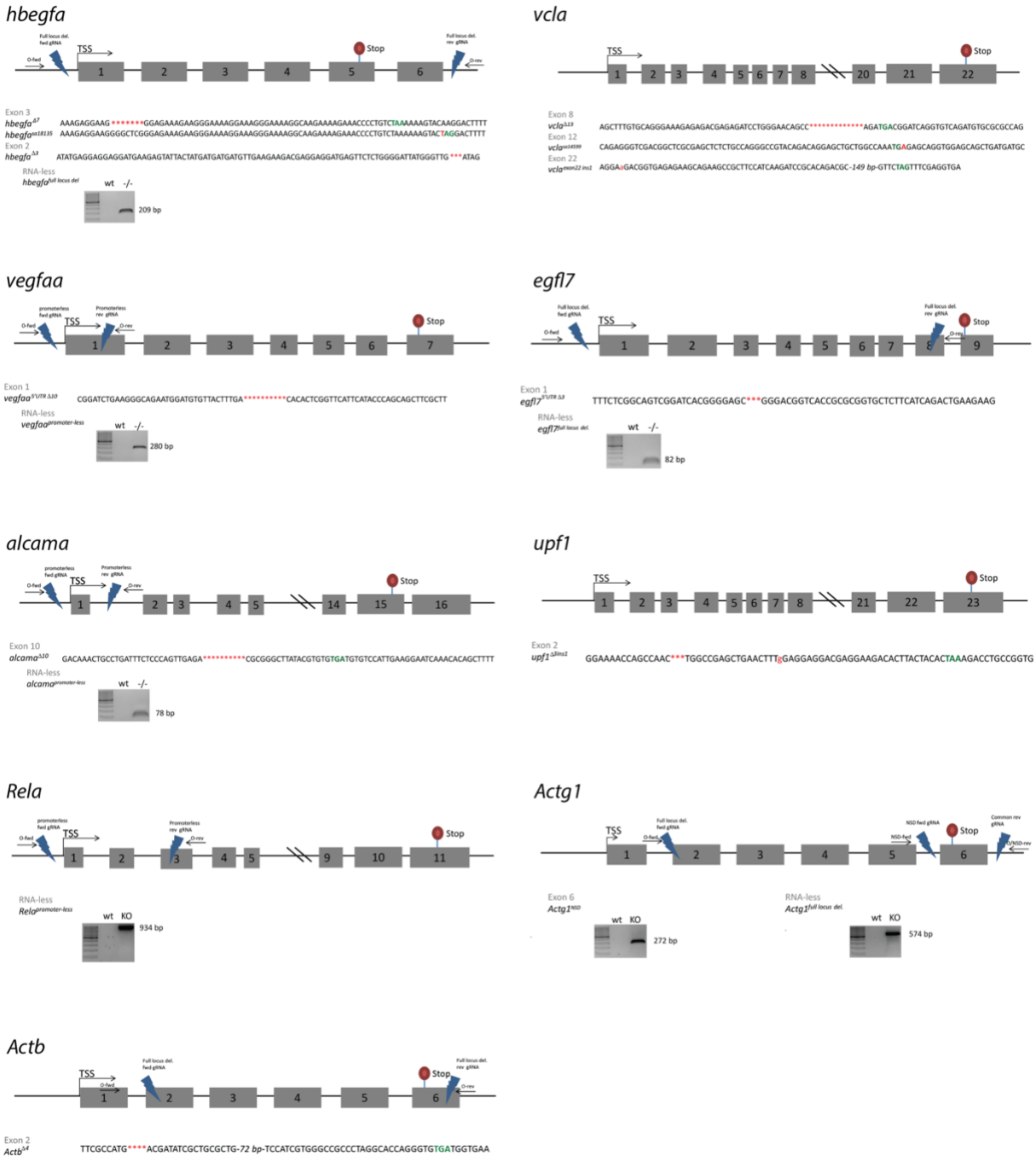


Figure 3. Mutant alleles generated for this thesis study.

Partial sequences of the generated mutant alleles for this thesis study, and gel snapshots providing evidence for the generated deletions in the different deletion alleles. Red letters: the generated mutation; green letters: PTCs; blue thunderbolts: gRNAs; arrows: primers used

for genotyping deletion alleles. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

hbegfa, *vcla*, *hif1ab*, *vegfaa*, *egfl7,* and *alcama* zebrafish mutants displayed an upregulation of either a paralogue or a related gene's mRNA expression levels, specifically *hbegfb*, *vclb*, *epas1a* and *epas1b*, *vegfab*, *emilin3a,* and *alcamb*, respectively (Figure 4A). Interestingly, injection of wild-type (wt) mRNAs of *hif1ab*, *vegfaa*, *egfl7* and *alcama* to the respective mutants did not dampen the upregulation levels of the adapting genes suggesting that the upregulation responses were not triggered due to loss of protein function but rather due to a transcriptional adaptation response (Figure 4B). Moreover, I observed upregulation of the adapting genes in *vcla*, *hif1ab,* and *egfl7* heterozygous animals, although to a lower extent than the corresponding homozygous mutants (Figure 4C). Furthermore, I observed upregulation of the wt allele in *hbegfa*, *hif1ab*, *vegfaa,* and *alcama* heterozygous fish (Figure 4D), suggesting that transcriptional adaptation modulates the mutated genes' expression levels as well.

Furthermore, *Fermt2* knockout MKFs, *Rela* and *Actg1* knockout MEFs, and *Actb* knockout mESCs (hereafter referred to as the knockout (K.O.) alleles) upregulated the mRNA expression levels of *Fermt1*, *Rel*, *Actg2,* and *Actg1*, respectively (Fig. 5A). Moreover, transfection of wild-type *Fermt2* and *Rela* in the *Fermt2* and *Rela* knockout cells did not lead to the reduction in the upregulation levels of *Fermt1* and *Rel*, suggesting that these responses were triggered upstream of the loss of the protein function as well (Figure 5B-D). In addition, *Actb* heterozygous mESCs also upregulated *Actg1* (Figure 5E). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

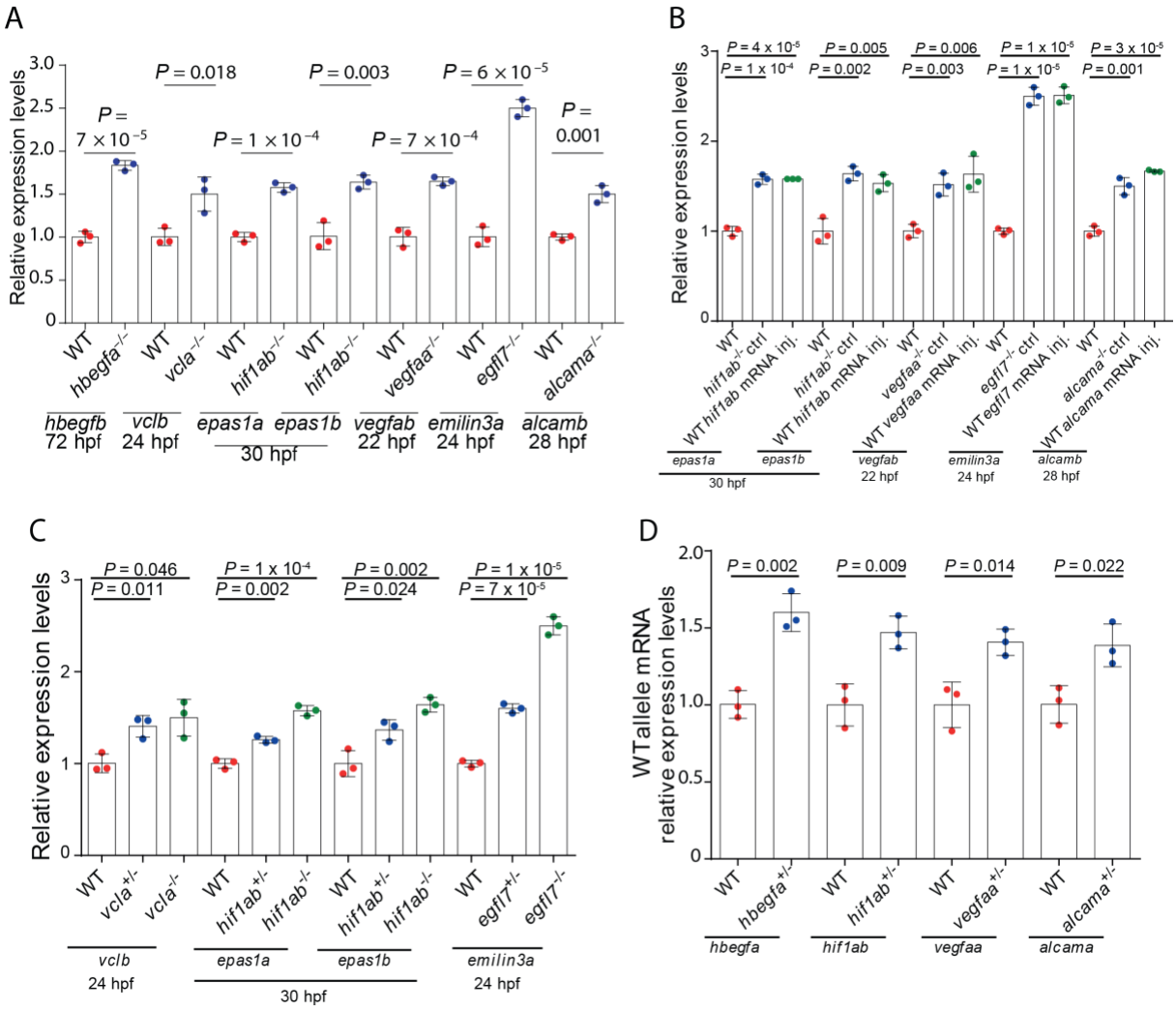


Figure 4. Zebrafish transcriptional adaptation models.

A) qPCR analysis of the mRNA expression levels of *hbegfb*, *vclb*, *epas1a* and *epas1b*, *vegfab*, *emilin3a*, and *alcamb* in *hbegfa*, *vcla*, *hif1ab*, *vegfaa*, *egfl7* and *alcama* wild-type and homozygous zebrafish mutants (−/−). B) qPCR analysis of the mRNA expression levels of *epas1a* and *epas1b*, *vegfab*, *emilin3a* and *alcamb* in *hif1ab*, *vegfaa*, *egfl7* and *alcama* wild-type fish and mutants injected (inj.) with *eGFP* mRNA as control (ctrl) or wild-type *hif1ab*, *vegfaa*, *egfl7* or *alcama* mRNA. C) qPCR analysis of the mRNA expression levels of *vclb*, *epas1a*, and *epas1b*, and *emilin3a* in *vcla*, *hif1ab*, and *egfl7* wild-type, homozygous and heterozygous (+/-) zebrafish mutants. D) qPCR analysis of the wild-type allele mRNA expression levels of *hbegfa*, *hif1ab*, *vegfaa*, and *alcama* in *hbegfa*, *hif1ab*, *vegfaa*, and *alcama* wild-type and heterozygous zebrafish animals using primers specific for the wild-type allele. n=3 independent biological samples. Control or wild-type expression levels were set at 1 for the analyses. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

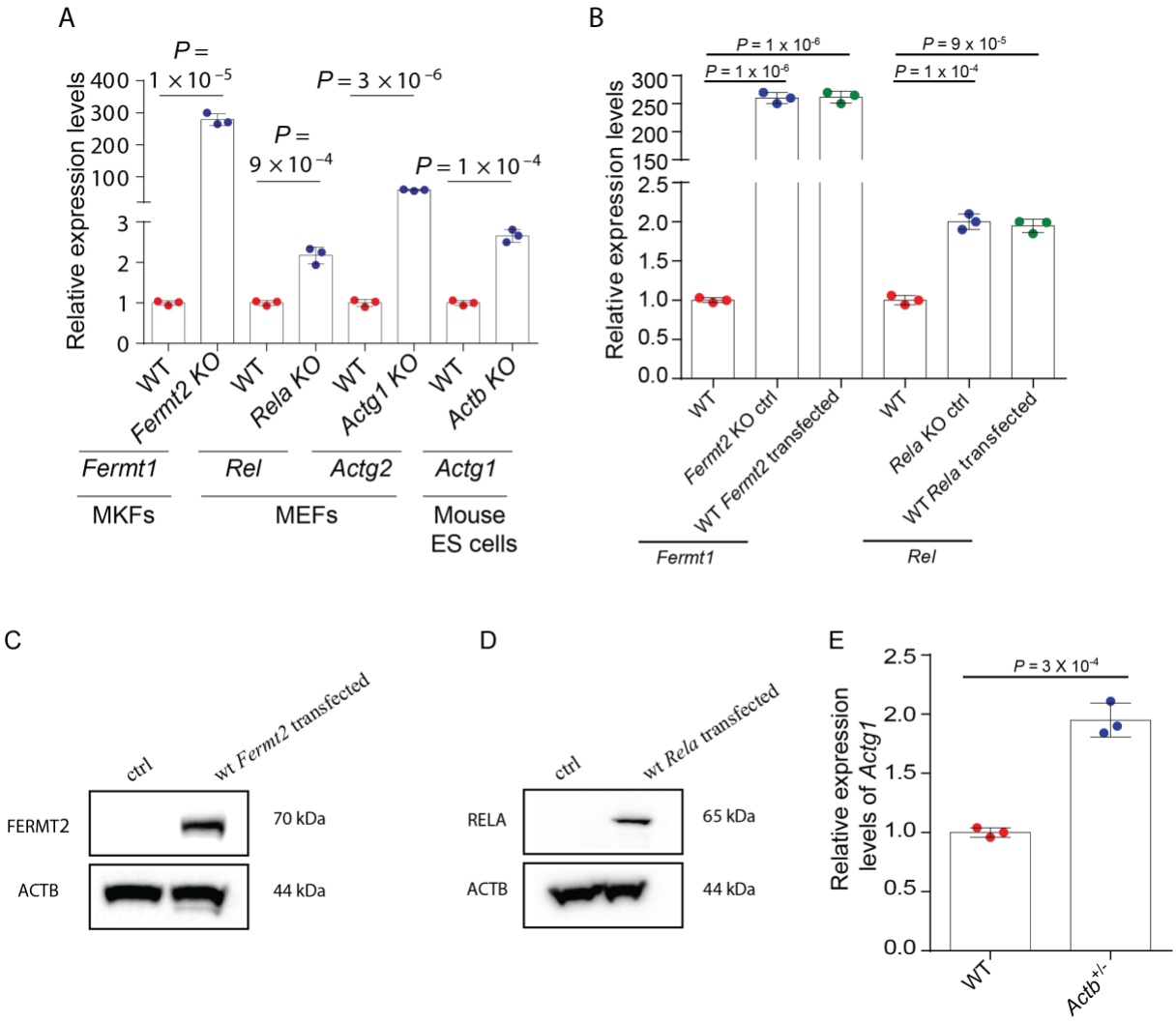


Figure 5. Mouse transcriptional adaptation models.

A) qPCR analysis of the mRNA expression levels of *Fermt1*, *Rel*, *Actg2* and *Actg1* in *Fermt2*, *Rela*, *Actg1* and *Actb* wild-type and knockout (KO) cell lines. B) qPCR analysis of the mRNA expression levels of *Fermt1* and *Rel* in *Fermt2* and *Rela* wild-type cells and knockout cells transfected with empty vectors as control (ctrl) or plasmids encoding wild-type FERMT2 or RELA (transfected). C) Western blot analysis of FERMT2 and ACTB (as a loading control) levels in *Fermt2* KO cells transfected with empty vectors as control (ctrl) or wild-type FERMT2-encoding plasmids. D) Western blot analysis of RELA and ACTB (as a loading control) levels in *Rela* KO cells transfected with empty vectors as control (ctrl) or wild-type RELA-encoding plasmids. E) qPCR analysis of the mRNA expression levels of *Actg1* in *Actb* wild-type and heterozygous mESCs. n=3 independent biological samples. Control or wild-type expression levels were set at 1 for the analyses. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

6.2. Increased transcription of adapting genes underlies transcriptional adaptation.

To identify if the increased mRNA expression levels of the adapting genes was due to increased transcription or increased stability of the mRNA, I analyzed pre-mRNA levels of *hbegfb and emilin3a* in *hbegfa* and *egfl7* zebrafish mutants and found that they were upregulated to almost the same levels as the respective mRNA (Figure 6A). *Fermt1* and *Rel* pre-mRNA levels were similarly upregulated in *Fermt2* and *Rela* knockout mouse cells (Figure 6B). In addition, ATAC-seq analysis revealed increased chromatin opening at the *Fermt1* transcription start site (TSSs) in *Fermt2* knockout cells (Figure 6C). Altogether, these data indicated that transcriptional adaptation involves increased transcription of the adapting genes rather than increasing mRNA stability. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

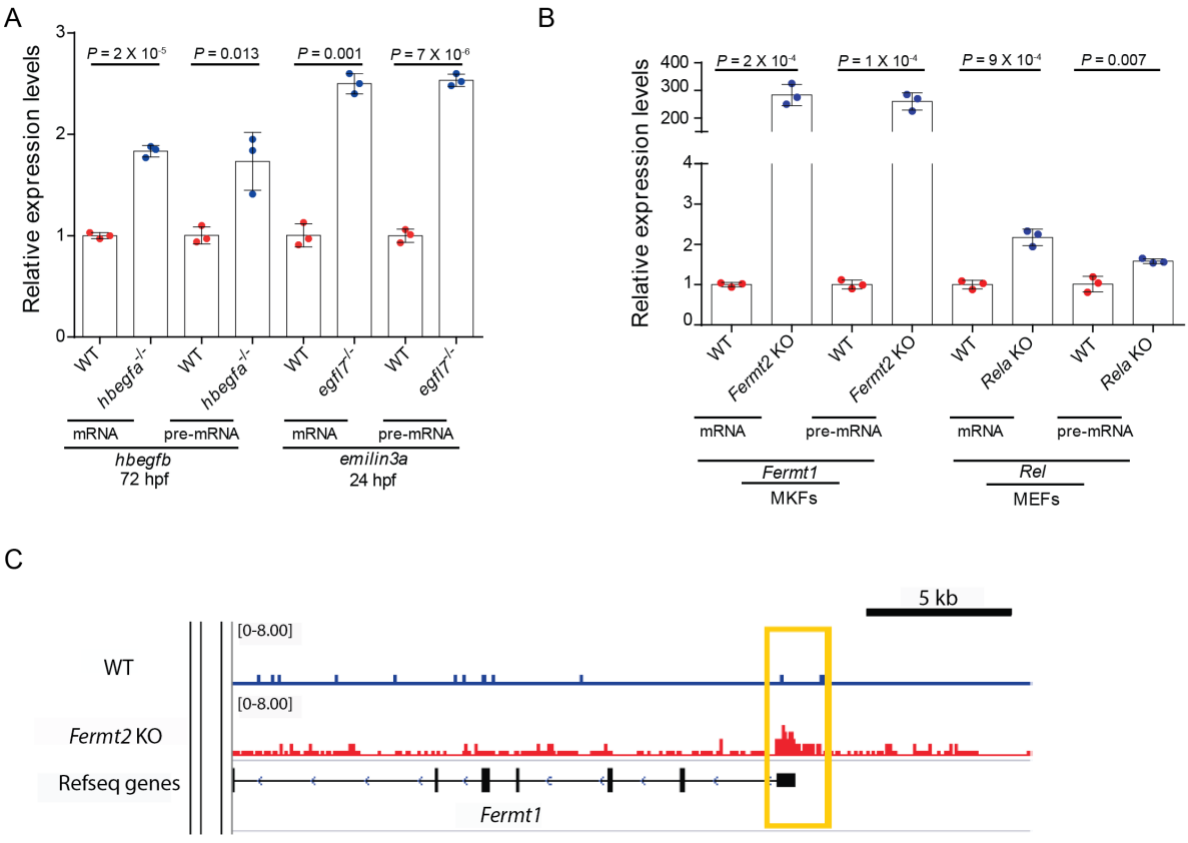


Figure 6. Transcriptional adaptation involves increased transcription of the adapting genes.

A) qPCR analysis of the mRNA and pre-mRNA expression levels of *hbegfb* and *emilin3a* in *hbegfa* and *egfl7* wild-type and mutant zebrafish. B) qPCR analysis of the mRNA and pre-mRNA expression levels of *Fermt1* and *Rel* in *Fermt2* and *Rela* wild-type and knockout

mouse cells. C) Screenshot from the integrated genome viewer (IGV) browser showing tracks of the *Fermt1* locus and displaying ATAC-seq signals in wild-type and *Fermt2* KO MKFs. n=3 independent biological samples. Control or wild-type expression levels were set at 1 for the analyses. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

6.3. Transcriptional adaptation is triggered by mutant mRNA degradation

6.3.1. DNA lesion is not the trigger for transcriptional adaptation

Given that loss of protein function is not the trigger for transcriptional adaptation, I ought to investigate two other possibilities: a) the DNA lesion b) consequences of the presence of mutant mRNA molecules. If DNA lesion is the trigger for transcriptional adaptation, then I should expect to observe upregulation of adapting genes in any kind of mutant alleles, including those that do not affect the protein function. However, inframe mutant alleles of *hbegfa* and *egfl7*, that are not expected to affect protein function, did not display a transcriptional adaptation response (Figure 7A), and similarly, mutations in the 5' untranslated region (5'UTR) of *vegfaa* and *egfl7* did not lead to upregulation of the adapting genes (Figure 7C). Furthermore, a mutation in the last exon of *vcla*, that doesn't affect RNA or protein integrity, did not lead to upregulation of *vclb* (Figure 7C). Altogether, these data indicated that the DNA lesion itself is not the trigger for transcriptional adaptation. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

6.3.2. Transcriptional adaptation correlates with mutant mRNA degradation

Interestingly, while analyzing different mutant alleles for some of the gene models I was studying, I observed that while *hbegfa*$^{\Delta 7}$ and *vcla*$^{\Delta 13}$ mutants upregulate *hbegfb* and *vclb*, respectively, two other PTC-containing alleles, *hbegfa*^{sa18135} and *vcla*^{sa14599}, did not exhibit a transcriptional adaptation response (Fig. 8A). To understand the underlying cause for this discrepancy, I analyzed the mutant mRNA levels in the different alleles and observed that while *hbegfa*$^{\Delta 7}$ and *vcla*$^{\Delta 13}$ mutants (that exhibit a transcriptional adaptation response) displayed a strong reduction in the mutant mRNA levels, *hbegfa*^{sa18135} and *vcla*^{sa14599} mutants displayed only a slight or no decrease in mutant mRNA levels (Fig. 8B). This reduction in mutant mRNA levels was observed in all models of transcriptional adaptation I was analyzing (Figure 8 C,

D) and notably, absent in the mutant alleles that did not display a transcriptional adaptation response (Figure 7A-C).

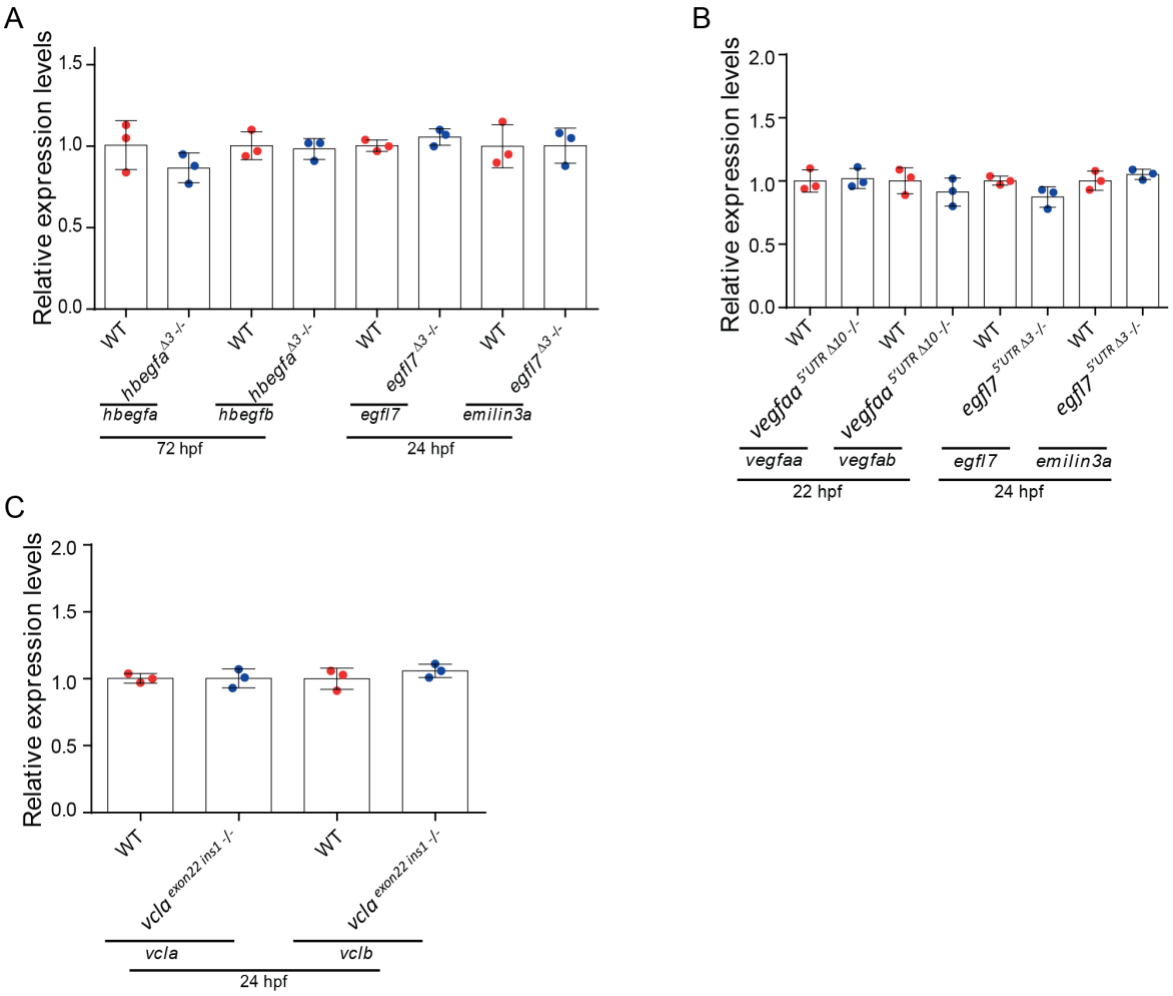


Figure 7. DNA lesion by itself is not sufficient to induce a transcriptional adaptation response.

A) qPCR analysis of the mRNA expression levels of *hbegfa*, *hbegfb* and *egfl7*, *emilin3a* in *hbegfa* and *egfl7* wild-type and inframe Δ3 mutant zebrafish. B) qPCR analysis of the mRNA expression levels of *vegfaa*, *vegfab* and *egfl7*, *emilin3a* in *vegfaa* and *egfl7* wild-type and 5'UTR mutant zebrafish. C) qPCR analysis of the mRNA expression levels of *vcla* and *vclb* in *vcla* wild-type last exon (exon 22) mutant zebrafish. n=3 independent biological samples. Control or wild-type expression levels were set at 1 for the analyses. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

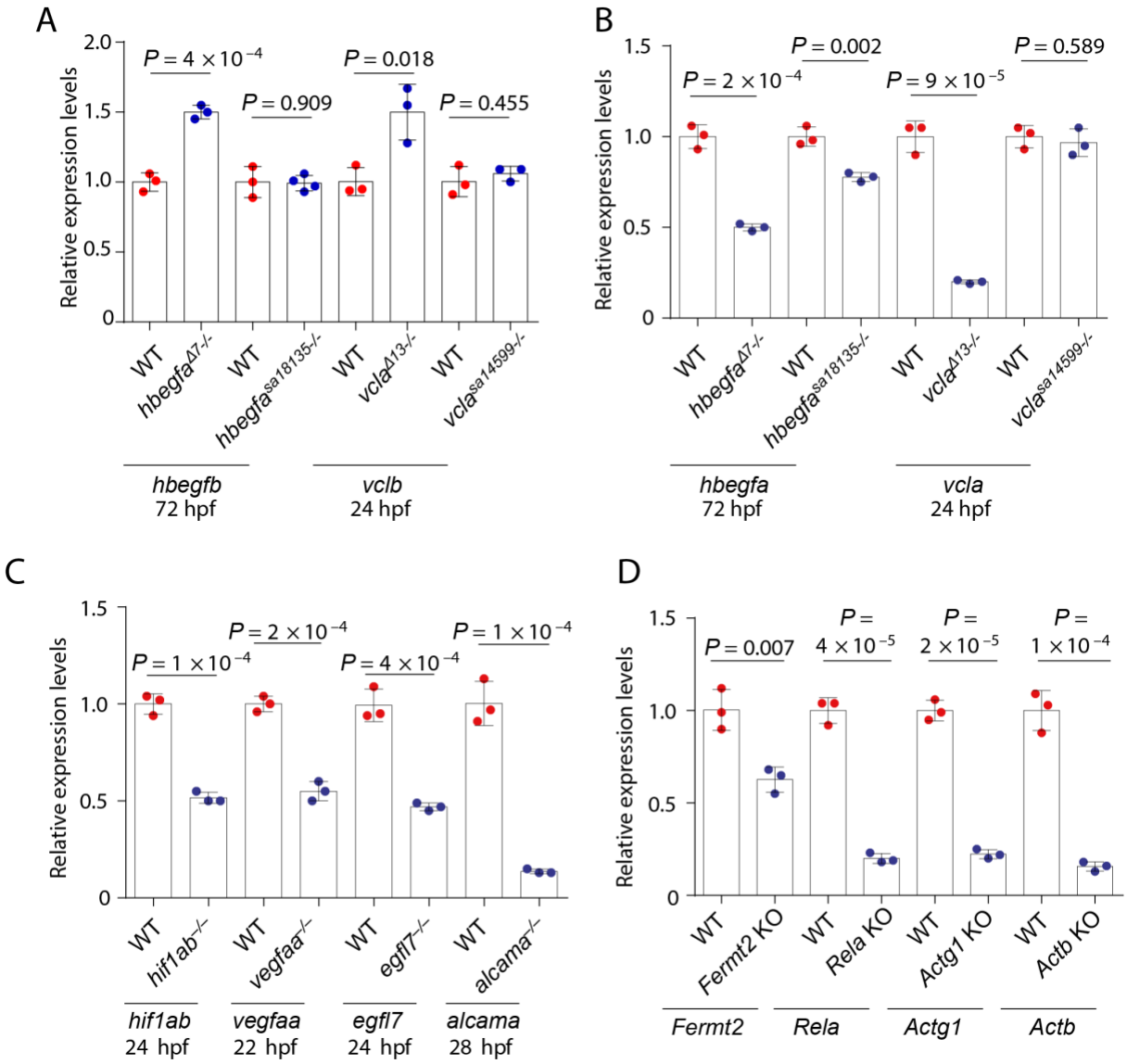


Figure 8. Transcriptional adaptation correlates with mutant mRNA degradation.

A) qPCR analysis of the mRNA expression levels of *hbegfb* and *vclb* in wild-type and the different indicated *hbegfa* and *vcla* mutant alleles. B) qPCR analysis of the mRNA expression levels of *hbegfa* and *vcla* in wild-type and the different indicated *hbegfa* and *vcla* mutant alleles. C) qPCR analysis of the mRNA expression levels of *hif1ab*, *vegfaa*, *egfl7* and *alcama* in *hif1ab*, *vegfaa*, *egfl7* and *alcama* wild-type and mutant zebrafish. D) qPCR analysis of the mRNA expression levels of *Fermt2*, *Rela*, *Actg1* and *Actb* in *Fermt2*, *Rela*, *Actg1* and *Actb* wild-type and knockout mouse cells. n=3 independent biological samples. Control or wild-type expression levels were set at 1 for the analyses. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

To identify if the decreased mutant mRNA expression levels were due to decreased transcription or decreased stability of the mutant mRNA, I analyzed levels of *hbegfa*, *egfl7* and *alcama* pre-mRNA in *hbegfa*$_{\Delta 7}$, *egfl7,* and *alcama* mutant zebrafish, and

observed that they were not changed, or slightly upregulated, compared to wildtype, unlike the mRNA levels that were downregulated (Figure 9A). Similarly, *Fermt2* and *Rela* pre-mRNA levels were unchanged in the knockout cells compared to wild-type (Figure 9B). In addition, I performed metabolic labeling experiments to assess newly transcribed RNAs in a short time frame (1 hour) and observed increased levels of *Fermt2*, *Rela* and *Actg1* mutant mRNAs and pre-mRNAs compared to wt (Figure 9C). Furthermore, transcription inhibition experiments using Actinomycin-D confirmed shorter half-lives of mutant *Fermt2*, *Rela* and *Actg1* mRNAs (Figure 9 D-F). Taken together, these data indicate that mutants displaying transcriptional adaptation exhibit decreased mutant mRNA expression levels due to mRNA degradation rather than decreased transcription. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

6.3.3. Inhibition of the mRNA surveillance machinery blunts the transcriptional adaptation response

Based on finding the correlation between transcriptional adaptation and mutant mRNA degradation, I ought to analyze the mRNA surveillance machinery's role in triggering transcriptional adaptation. I started by crossing a Upf1 mutant to some of the zebrafish mutant models of transcriptional adaptation, creating double mutants. Upf1 is a central factor in non-sense mediated decay as previously explained (Isken and Maquat, 2007) and as expected, mutating *upf1* in *hbegfa*$^{\Delta 7}$, *vegfaa* and *vcla*$^{\Delta 13}$ mutant backgrounds led to reduced decay and increased stability of the respective mutant mRNAs (Figure 10A). Interestingly, this reduction in mutant mRNA decay led to loss of the transcriptional adaptation response (Figure 10B). Similarly, knockdown of other proteins of the mRNA surveillance machinery, such as ERF1, SMG6, EXOSC4 (a component of the exosome complex) and XRN1, led to reduction, or loss, of the transcriptional adaptation response in *Rela* and *Actb* knockout cells (Figure 10C, D), in a manner that correlated with reduced mutant mRNA decay (Figure 10E, F). Moreover, pharmacological inhibition of nonsense-mediated decay in *hbegfa*$^{\Delta 7}$ mutant larvae stabilized *hbegfa* mutant mRNA and blocked the transcriptional adaptation response (Figure 10G, H). Blocking translation with cycloheximide (CHX), and thereby inhibiting the different mRNA surveillance machinery, in *Rela* knockout MEFs, also led to stabilizing *Rela* mutant mRNA and loss of *Rel* upregulation (Figure 10 I, J). Certain lines in this subsection have been

quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

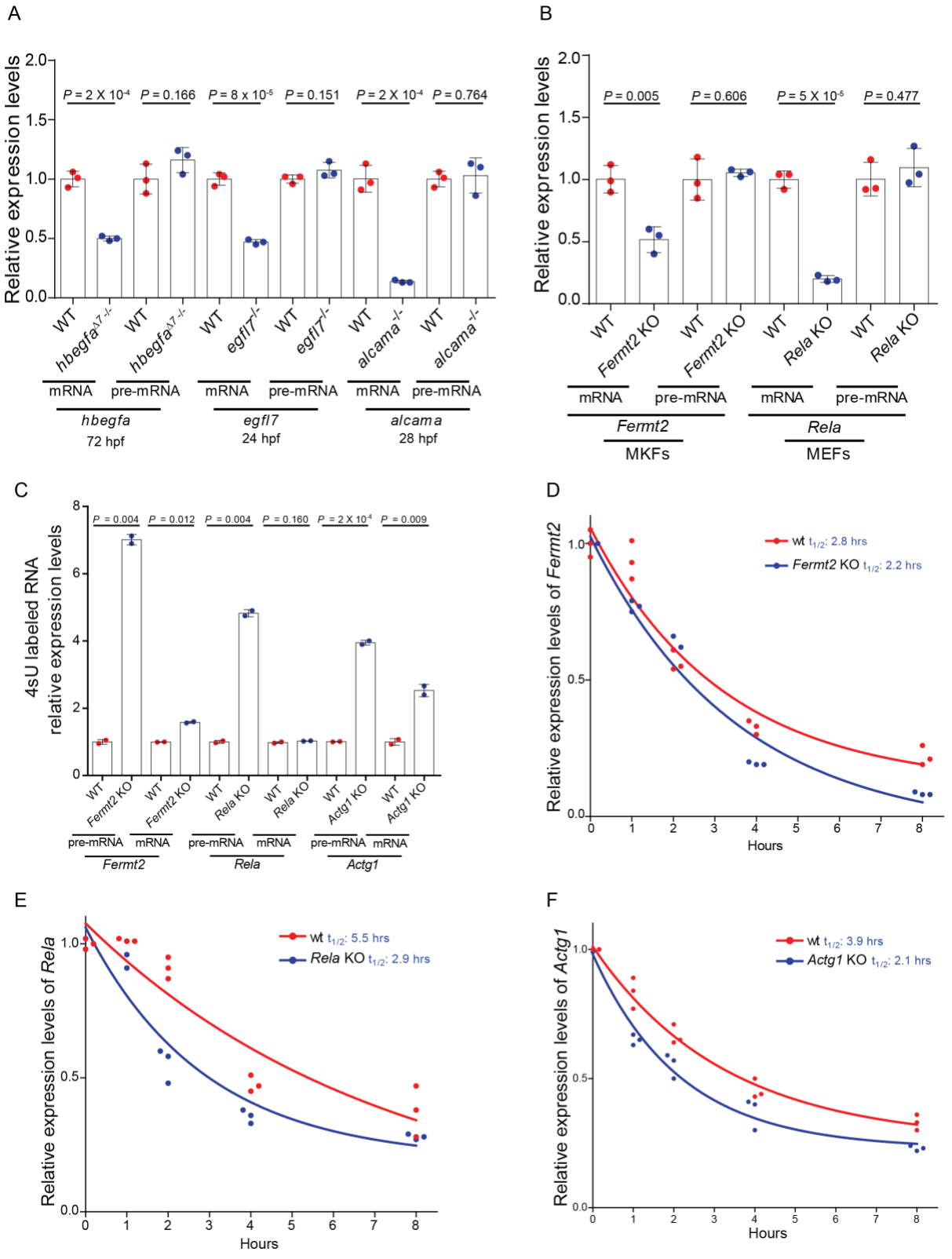


Figure 9. mRNA degradation underlies reduced mutant mRNA levels in models of transcriptional adaptation.

A) qPCR analysis of mRNA and pre-mRNA expression levels of *hbegfa*, *egfl7* and *alcama* in wild-type and *hbegfa*, *egfl7* and *alcama* mutant zebrafish. B) qPCR analysis of mRNA and pre-mRNA expression levels of *Fermt2* and *Rela* in wild-type and *Fermt2* and *Rela* knockout mouse cells. C) qPCR analysis of 4-thiouridine (4sU)-labeled mRNA and pre-mRNA expression levels of *Fermt2*, *Rela* and *Actg1* in *Fermt2*, *Rela* and *Actg1* wt and knockout cells. D) Fitted exponential degradation curves of mRNA levels of *Fermt2* in wild-type and *Fermt2* knockout MKFs. E) Fitted exponential degradation curves of mRNA levels of *Rela* in wild-type and *Rela* knockout MEFs. F) Fitted exponential degradation curves of mRNA levels of *Actg1* in wild-type and *Actg1* knockout MEFs. $t_{1/2}$: half-life. n=3 (A, B, D-F); 2 (C) independent biological samples. Control or wild-type expression levels were set at 1 for the analyses. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

6.3.4. Ectopic induction of RNA decay can trigger a transcriptional adaptation response

Reducing or eliminating mutant mRNA decay reduces the transcriptional adaptation response. I next asked if ectopic induction of RNA degradation through injection of uncapped transcripts, which get rapidly degraded inside through the 5' to 3' exonucleases (Mukherjee et al., 2012), in wild-type zebrafish embryos or mouse cells can trigger a transcriptional adaptation response. Interestingly, I observed that injection of uncapped *hif1ab* and *vegfaa* transcripts to 1-cell stage zebrafish embryos led to upregulation of *epas1a* and *vegfab*, respectively, as soon as 6 hours post-fertilization, concomitant with an almost-full degradation of the injected uncapped transcripts (Figure 11A), besides an increase in the endogenous *hif1ab* and *vegfaa* expression levels (Figure 11B). In addition, transfection of uncapped *Actb* transcripts into wild-type mouse embryonic stem cells induced a transcriptional adaptation response as fast as 6 hours post-transfection (Figure 11C). To confirm the requirement for degradation, I injected uncapped *hif1ab* and *vegfaa* transcripts with a 5' upstream viral sequence that renders them resistant to XRN1 5' to 3' exonuclease activity (Boehm et al., 2016). As expected, the injected embryos did not display a transcriptional adaptation response, in accordance with lack of degradation of the injected transcripts (Figure 11D). Altogether, these data indicate that mutant mRNA degradation plays a very central role in triggering transcriptional adaptation. Quite interestingly, injection of uncapped transcripts corresponding to the non-coding strand of *hif1ab* or *vegfaa* did not induce the upregulation of *epas1a* or *vegfab* (Figure 11E), suggesting that the RNA sequence itself might influence the transcriptional adaptation response. Certain lines in this subsection have been

quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

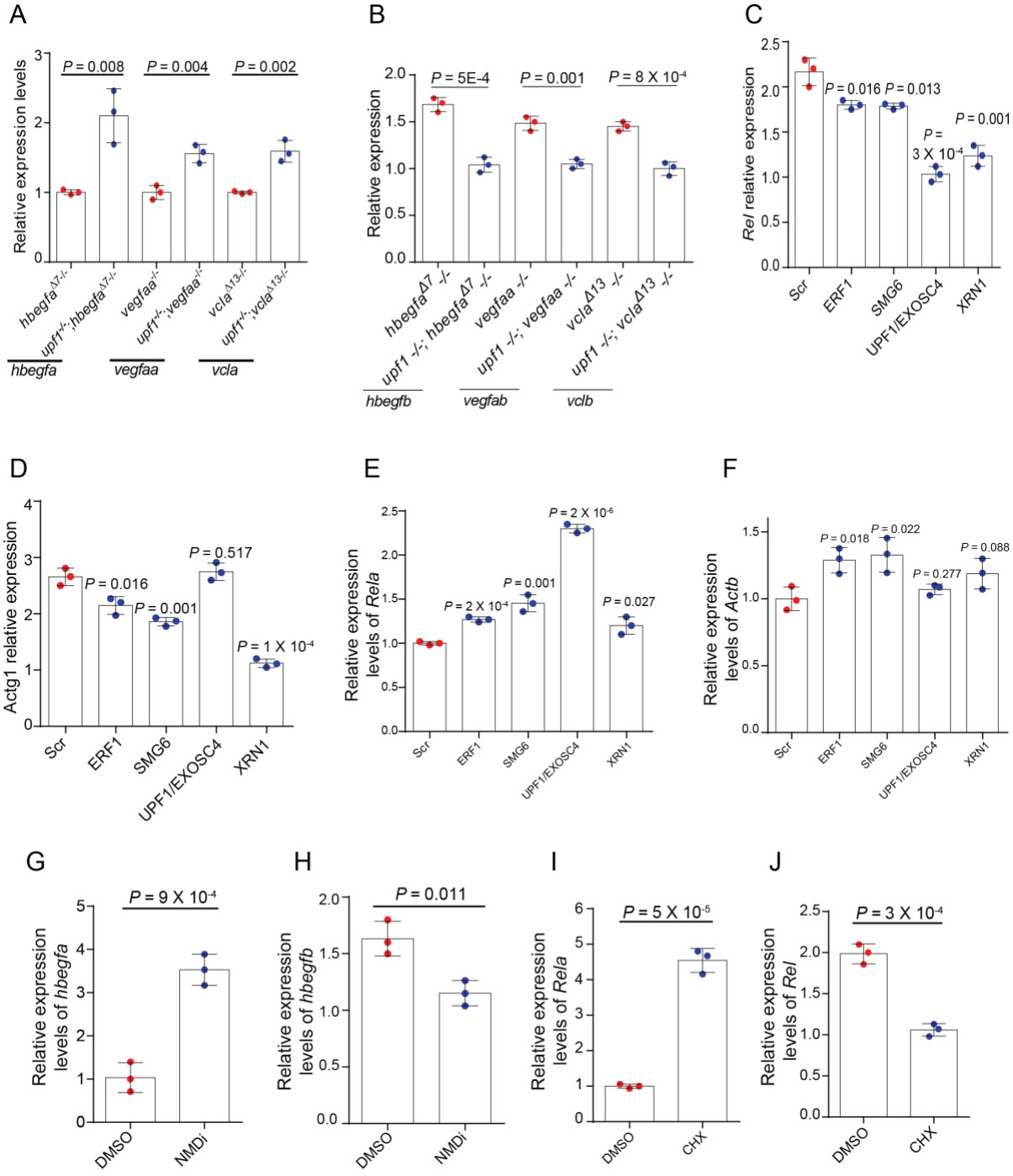


Figure 10. mutant mRNA decay is necessary for inducing transcriptional adaptation.

A) qPCR analysis of mRNA expression levels of *hbegfa*, *vegfaa*, and *vcla* in *upf1*;*hbegfa*, *upf1*;*vegfaa* and *upf1*;*vcla* wild-type and double mutant zebrafish. B) qPCR analysis of mRNA expression levels of *hbegfb*, *vegfab*, and *vclb* in *upf1*;*hbegfa*, *upf1*;*vegfaa* and *upf1*;*vcla* wild-type and double mutant zebrafish. C) qPCR analysis of mRNA expression levels of *Rel* following knockdown of the indicated proteins in *Rela* knockout MEFs using siRNAs. D) qPCR analysis of mRNA expression levels of *Actg1* following knockdown of the

indicated proteins in *Actb* knockout mESCs using siRNAs. E) qPCR analysis of mRNA expression levels of *Rela* following knockdown of the indicated proteins in *Rela* knockout MEFs using siRNAs. F) qPCR analysis of mRNA expression levels of *Actb* following knockdown of the indicated proteins in *Actb* knockout mESCs using siRNAs. G) qPCR analysis of mRNA levels of *hbegfa* in 6 days post fertilization (dpf) wild-type and *hbegfa* mutant zebrafish treated with NMD inhibitor (NMDi). H)) qPCR analysis of mRNA levels of *hbegfb* in 6 days post fertilization (dpf) wild-type and *hbegfa* mutant zebrafish treated with NMDi. I) qPCR analysis of mRNA levels of *Rela* in wild-type and *Rela* knockout MEFs treated with CHX. J) qPCR analysis of mRNA levels of *Rel* in wild-type and *Rela* knockout MEFs treated with CHX. n=3 independent biological samples. Control or wild-type expression levels were set at 1 for the analyses in panels A, E-G and I. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

6.3.5. Mutant alleles that fail to transcribe the mutated gene do not display transcriptional adaptation and exhibit stronger phenotypes

Next, I reasoned that if mutant mRNA degradation is required for transcriptional adaptation, then alleles that fail to transcribe the mutated gene should not exhibit a transcriptional adaptation response. I, therefore, made use of the CRISPR/Cas9 technology to engineer such alleles through either deleting the promoter region (Figure 12A) of the gene models I was studying, or the full gene locus (Figure 12B) (hereafter referred to as RNA-less alleles). Indeed, I observed that *hbegfa*, *vegfaa* and *alcama* RNA-less alleles do not display a transcriptional adaptation response (Figure 12C). Moreover, *egfl7* RNA-less alleles displayed a milder upregulation of the *emilin* genes compared to the *egfl7*$^{\Delta 4}$ allele (Rossi et al., 2015) that displays mutant mRNA decay (Figure 12D). Furthermore, in mouse cultured cells, *Rela*, *Actg1* and *Actb* RNA-less alleles failed to upregulate *Rel*, *Actg2* and *Actb*, respectively (Figure 12E). I attempted to engineer a *Fermt2* RNA-less allele in MKFs as well, however, the obtained clones displayed defects in proliferation that prevented me from expanding them for the analysis. Alternatively, I made use of the CRISPR interference (CRISPRi) technology (Qi et al., 2013; Thakore et al., 2015), and observed that decreasing the transcription of the mutated *Fermt2* gene in *Fermt2* knockout MKFs leads to a reduction in *Fermt1* expression levels (Figure 12F).

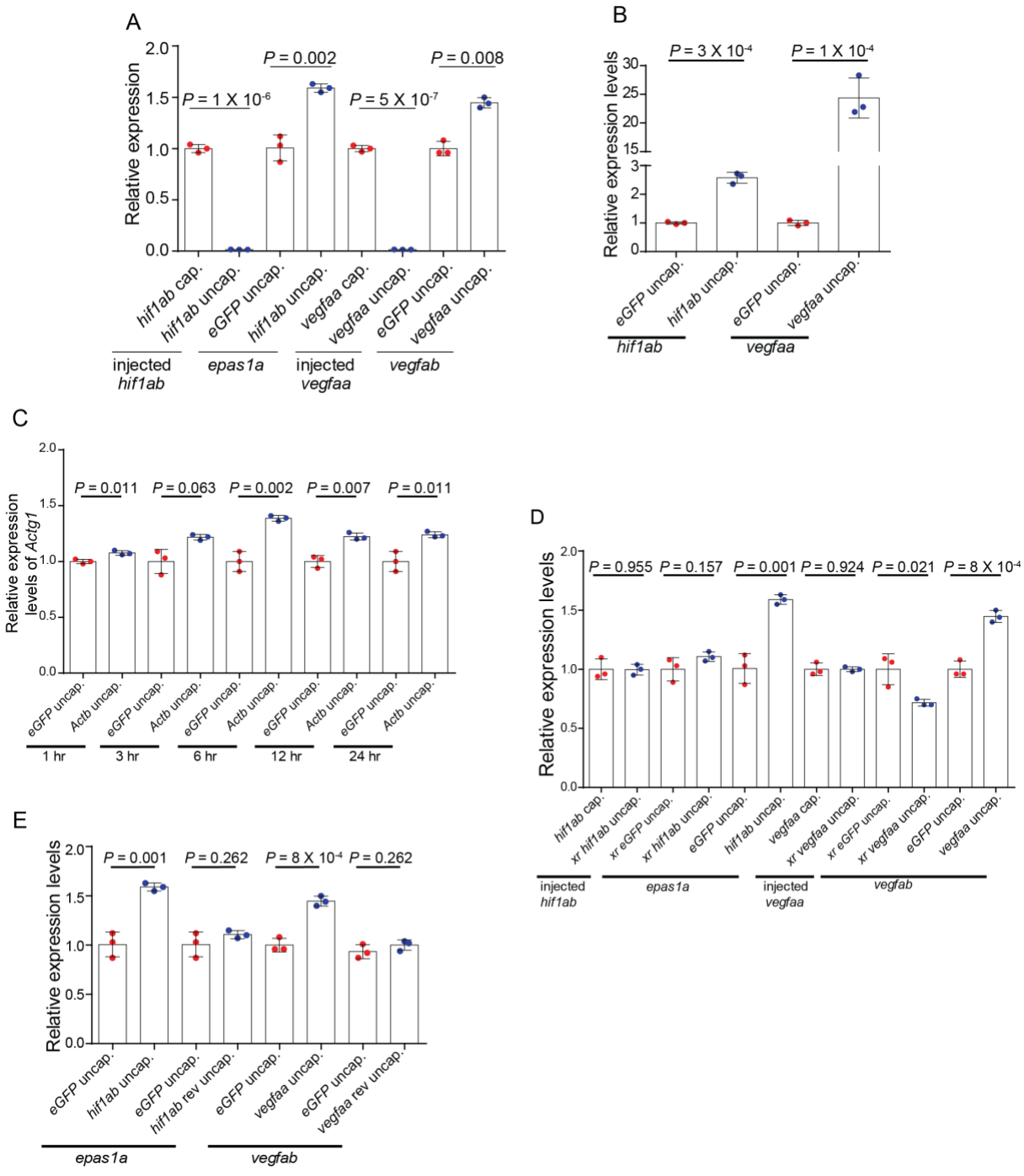


Figure 11. Ectopic induction of RNA decay can trigger transcriptional adaptation.

A) qPCR analysis of RNA levels of the injected *hif1ab*, *epas1a* and injected *vegfaa*, *vegfab* in 6 hpf wild-type zebrafish embryos injected with uncapped *eGFP* as control or *hif1ab* or *vegfaa* RNA. B) qPCR analysis of mRNA levels of the endogenous *hif1ab* and *vegfaa* in 6 hpf wild-type zebrafish embryos injected with uncapped *eGFP* as control or *hif1ab* or *vegfaa* RNA. C) qPCR analysis of mRNA levels of *Actg1* in wild-type mouse embryonic stem cells transfected with uncapped *eGFP* or *Actb* RNA at the indicated times post-transfection. hr: hours. D) qPCR analysis of RNA levels of injected *hif1ab*, *epas1a* and injected *vegfaa*, *vegfab* in 6 hpf wild-type zebrafish embryos injected with uncapped eGFP (as control) or hif1ab or *vegfaa* RNAs with or without a 5' upstream xrFRAG (xr) sequence. E) qPCR

analysis of mRNA levels of *epas1a* and *vegfab* in 6 hpf wild-type embryos injected with uncapped *hif1ab* or *vegfaa* RNAs corresponding to the coding or non-coding strand (rev). n=3 independent biological samples. Control or wild-type expression levels were set at 1 for the analyses in panels A, E-G and I. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

The lack of a transcriptional adaptation response in RNA-less alleles persuaded me to investigate if such alleles display stronger phenotypes than alleles displaying mutant mRNA decay. The NF-κB signaling pathways (which the genes *Rela* and *Rel* are part of) helps prevent tumor necrosis factor α (TNFα) from inducing apoptosis (Alcamo et al., 2001; Doi et al., 1999; Rosenfeld et al., 2000) and *Rel* upregulation in *Rela* knockout cells can help decrease TNFα-induced apoptosis in *Rela* knockout cells (Gapuzan et al., 2005). Interestingly, I found that promoter-less *Rela* mouse embryonic fibroblasts were more susceptible to apoptosis upon TNFα treatment than the *Rela* knockout cells, possibly due to lack of *Rel* upregulation (Figure 13A). Moreover, *Actb* full-locus deletion mouse embryonic stem cells exhibited reduced filamentous actin protrusive activity, in addition to stronger proliferation defects, than *Actb* knockout mESCs (Figure 13B, C). Notably, *egfl7* RNA-less mutant zebrafish embryos displayed strong vascular defects unlike the previously reported *egfl7*$^{\Delta 4}$ allele displaying mutant mRNA decay (Rossi et al., 2015) that was phenotypically wildtype (Figure 13D). Furthermore, promoter-less *vegfaa* mutant embryos exhibited a stronger central artery (CtA) sprouting phenotype compared to the *vegfaa*$^{\Delta 10}$ mutants that display mutant mRNA decay (Figure 13E). In addition, RNA-less *hbegfa* mutant larvae, but not the PTC-bearing *hbegfa*$^{\Delta 7}$ mutants, exhibited a slow blood-flow phenotype (Figure 13F). Promoter-less *alcama* mutant larvae as well displayed an elongated cardiac ventricle phenotype, a phenotype that was absent in *alcama*$^{\Delta 8}$ mutant fish (Figure 13G). These data indicated that generation of RNA-less alleles can help uncover phenotypes that were previously masked in mutant alleles displaying mutant mRNA-decay through transcriptional adaptation-induced genetic compensation. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

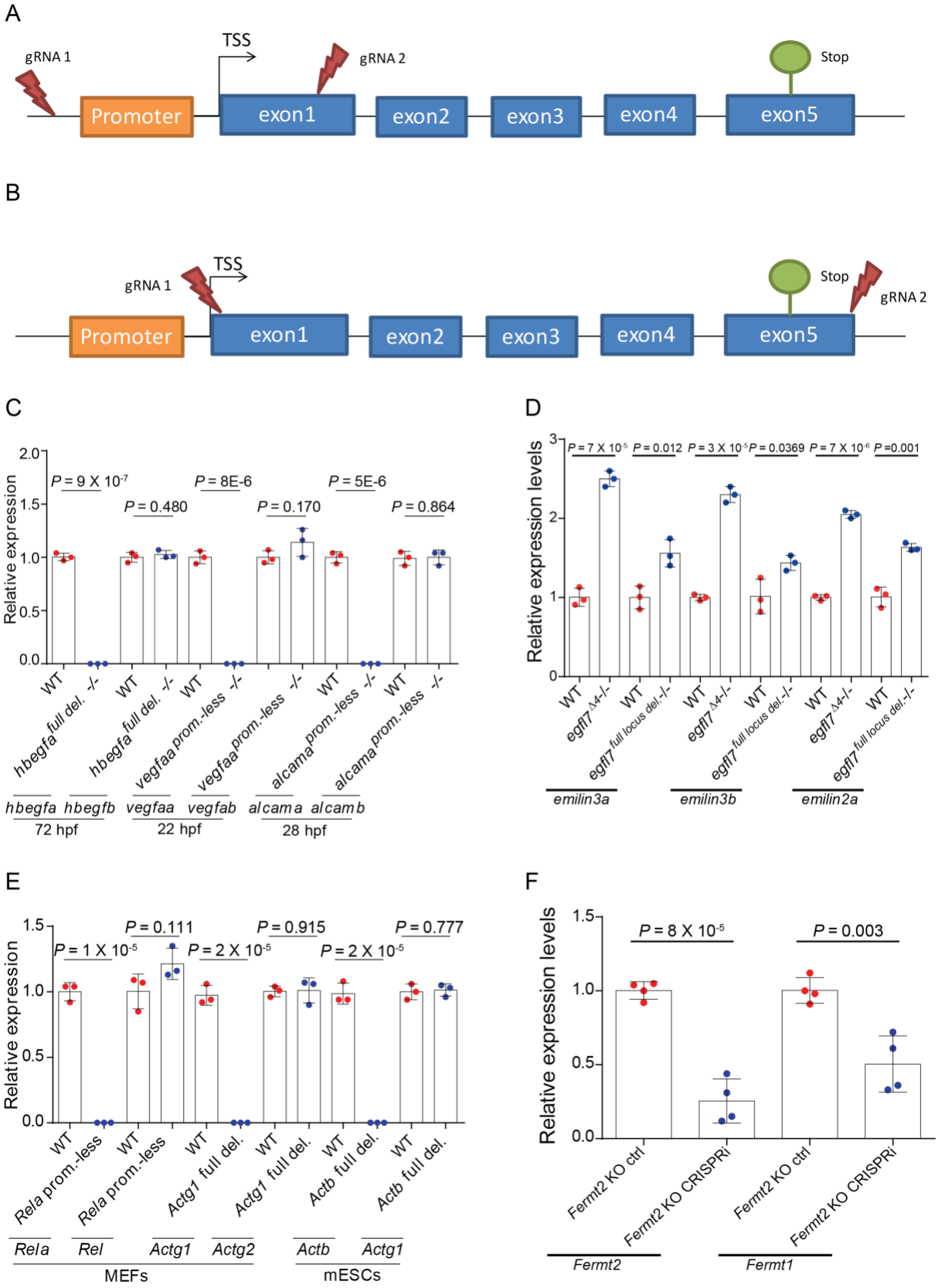


Figure 12. RNA-less alleles fail to induce a transcriptional adaptation response.

A, B) Schematic illustration of the strategy used to generate A) promoter-less or B) full-locus deletion alleles using the illustrated gRNAs and the CRISPR/Cas9 system. Red thunderbolt: gRNAs; TSS: transcription start site. C) qPCR analysis of mRNA expression levels of *hbegfa*

and *hbegfb*, *vegfaa* and *vegfab*, *alcama* and *alcamb* in wild-type and RNA-less mutant alleles of *hbegfa* (full locus deletion) or the *vegfaa* or *alcama* (promoterless alleles). D) qPCR analysis of the mRNA levels of *emilin3a*, *emilin3b* and *emilin2a* in wild-type, *egfl7*$^{\Delta 4}$ and *egfl7*$^{full\ locus\ del.}$ mutant zebrafish embryos. E) qPCR analysis of mRNA expression levels of *Rela* and *Rel*, *Actg1* and *Actg2*, *Actb* and *Actg1* in wild-type and RNA-less alleles of *Rela* (promoterless) or the *Actg1* or *Actb* (full locus deletions). F) qPCR analysis of mRNA expression levels of *Fermt2* and *Fermt1* following reduction of *Fermt2* transcription in *Fermt2* knockout MFKs by CRISPRi. n= 3(C-E), 4(F) independent biological samples. Control or wild-type expression levels were set at 1 for the analyses in panels C-F. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

6.4. Transcriptional adaptation and sequence similarity.

To understand the molecular mechanisms underlying transcriptional adaptation, I performed RNA-sequencing on *Fermt2*, *Actg1* and *Actb* wild-type, knockout and RNA-less cells. Hundreds of genes were upregulated or downregulated in the knockout alleles compared to wild-type (Table 30), however, only 81 genes were commonly upregulated in the three sequenced knockout models (Table 31, Figure 14A). Gene ontology analysis of those 81 genes revealed lack of a stress-induced response (Figure 14C).

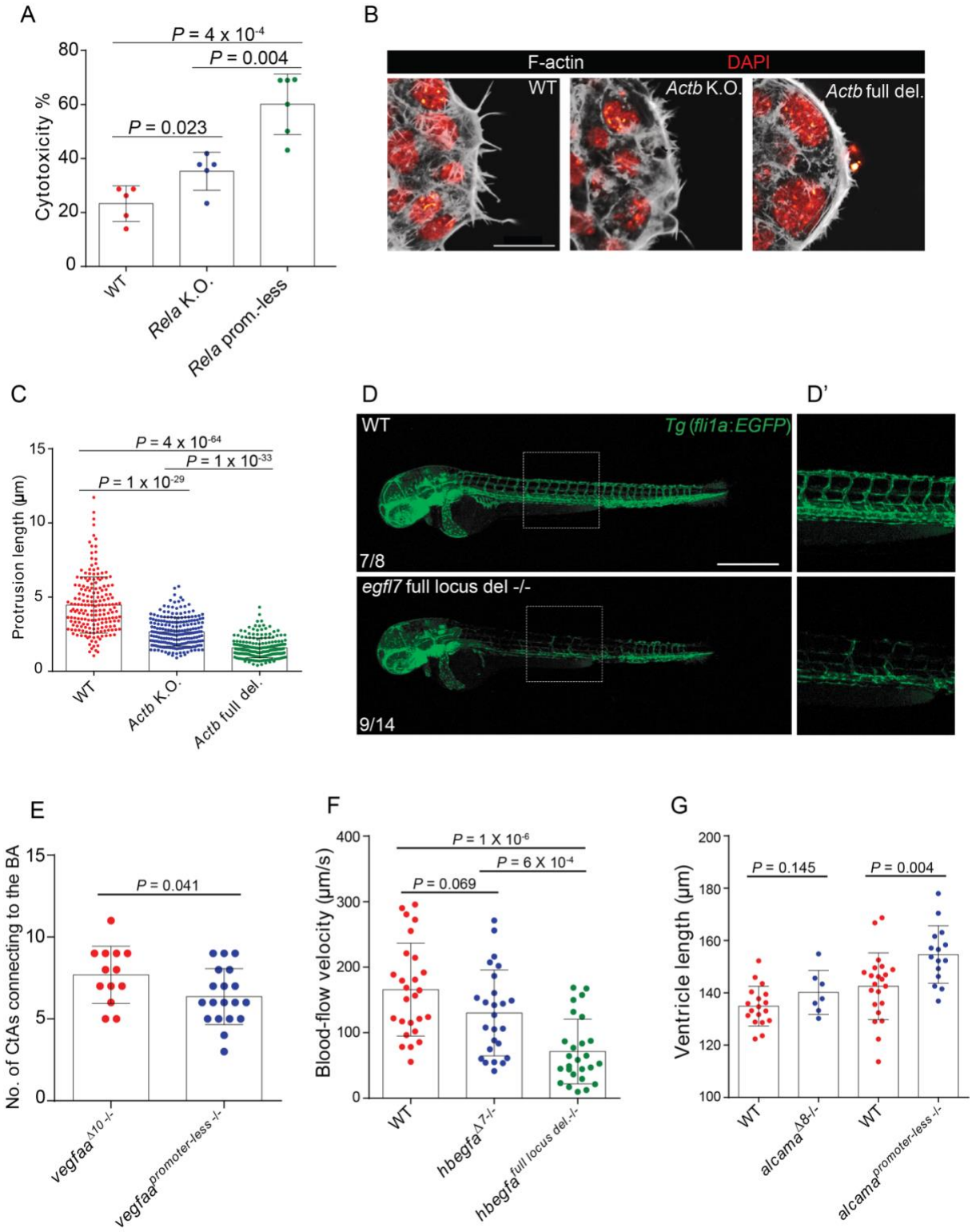


Figure 13. RNA-less display stronger phenotypes than alleles displaying mutant mRNA decay.

A) Cytotoxicity analysis after treatment with tumor necrosis factor α of wild-type, *Rela* knockout and *Rela* promoter-less mouse embryonic fibroblasts. Percentages are normalized to DMSO-treated cells. B) Confocal micrographs of wild-type, *Actb* knockout and *Actb* full locus deletion mouse embryonic stem cells. White: F-actin filaments; red: nuclei. Scale bar: 20 µm. C) Protrusion length of F-actin filaments in wild-type, *Actb* knockout and *Actb* full

locus deletion mouse embryonic stem cells. D) Confocal micrographs of 48 hours post-fertilization wild-type and *egfl7* full locus deletion mutant embryos in *Tg*(*fli1a*:*eGFP*) background. Lateral view with anterior to the left. D': higher magnification of the indicated dashed boxes. Scale bar: 500 µm. E) Number of connected central arteries (CtAs) to the basilar artery (BA) in *vegfaa*$^{\Delta 10}$ and *vegfaa*$_{promoter\text{-}less}$ mutant zebrafish embryos at 58 hours post-fertilization. F) Blood flow velocity in wild-type, and hbegfa$^{\Delta 7}$ and *hbegfa*$_{full\ locus\ del.}$ mutant zebrafish larvae at 78 hours post-fertilization. G) Cardiac ventricle length quantification in wild-type, and *alcama* $^{\Delta 8}$ and *alcama*$_{promoter\text{-}less}$ mutant zebrafish larvae at 100 hours post-fertilization. n= 5 (A), 189, 219 and 205 (C), 13 and 19 (E), 25 (F), and 18, 7, 22 and 15 (G) independent biological samples. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

	No. of upregulated genes L2F > 0.585	**No. of upregulated genes L2F KO > WT**	**No. of downregulated genes L2F < -0.585**	**No. of downregulated genes L2F KO < WT**
Fermt2 KO	2,002	3,634	2,237	3,935
Actg1 KO	2,081	3,509	2,225	2,652
Actb KO	584	1,155	703	1,110

Table 30. Number of differentially expressed genes in the indicated knockout cells compared to wildtype

L2F: $\log_2$ (fold change); $P \leq 0.05$. Table adapted and reprinted with permission from (El-Brolosy et al., 2019).

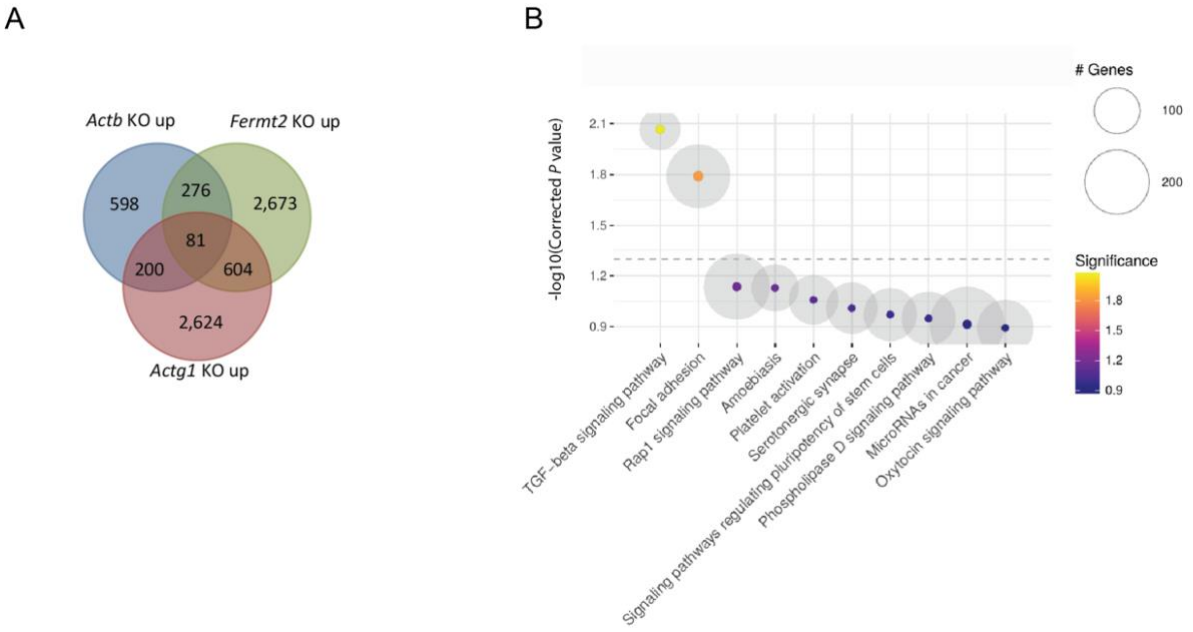


Figure 14. Cultured cells knockout models of transcriptional adaptation do not display a stress-induced response.

A) Venn diagram of the number of genes increased in expression in the three different knockout cultured cell models of transcriptional adaptation with Log_2 (fold change) (L2F) knockout > wildtype and $P \leq 0.05$. B) KEGG pathway enrichment analysis for the 81 genes upregulated commonly in *Fermt2, Actg1* and *Actb* knockout cells in comparison to wild-type cells. The top 10 KEGG pathways based upon *P* value are shown. Dashed line indicates the *P* value of 0.05. Size of circles provides scale; the total number of genes in a given pathway are represented by the outer gray circles while the number of genes in that specific pathway that are commonly upregulated is represented by the colored circles in the center. n = 2 independent biological samples. *P* values are not multiple testing corrected. DESeq2 tests for significance of coefficients in a negative binomial GLM (Generalized Linear Model) with the Wald test. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

Ensembl gene	KEGG PATHWAY terms
Imp4	Ribosome biogenesis in eukaryotes
Tmem131	
Eif5b	RNA transport
Pgap1	Metabolic pathways, Glycosylphosphatidylinositol(GPI)-anchor biosynthesis
Ankrd44	
Mars2	Selenocompound metabolism, Aminoacyl-tRNA biosynthesis
Ino80d	
Klf7	
Rpe	Metabolic pathways, Carbon metabolism, Biosynthesis of amino acids, Pentose phosphate pathway, Pentose and glucuronate interconversions
Epha4	Axon guidance
Trip12	Ubiquitin mediated proteolysis

Sned1	
Hdlbp	
Nifk	
Inhbb	TGF-beta signaling pathway, Cytokine-cytokine receptor interaction, Signaling pathways regulating pluripotency of stem cells
Ptgs2	Metabolic pathways, MicroRNAs in cancer, Pathways in cancer, TNF signaling pathway, Retrograde endocannabinoid signaling, Serotonergic synapse, Oxytocin signaling pathway, Ovarian steroidogenesis, Chemical carcinogenesis, NF-kappa B signaling pathway, Arachidonic acid metabolism, Leishmaniasis, Small cell lung cancer, Regulation of lipolysis in adipocytes, VEGF signaling pathway
Slc30a1	Mineral absorption
Elac2	RNA transport
Ntn1	Axon guidance
Trappc1	
Vps53	
Supt6	
Epn3	Endocytosis
Slc35b1	
Plekhm1	
Ubxn2a	
Odc1	Metabolic pathways, Glutathione metabolism, Arginine and proline metabolism
Id2	TGF-beta signaling pathway, Signaling pathways regulating pluripotency of stem cells, Transcriptional misregulation in cancer, Hippo signaling pathway
Lamb1	ECM-receptor interaction, Pathways in cancer, PI3K-Akt signaling pathway, Toxoplasmosis, Focal adhesion, Amoebiasis, Small cell lung cancer
Susd6	
Plk2	FoxO signaling pathway
Drosha	Ribosome biogenesis in eukaryotes, Proteoglycans in cancer
Tnfrsf21	Cytokine-cytokine receptor interaction
Epb41l4a	
Sema6a	Axon guidance
Ehbp1l1	
Cd82	p53 signaling pathway
Gm10800	
Thbs1	MicroRNAs in cancer, TGF-beta signaling pathway, ECM-receptor interaction, PI3K-Akt signaling pathway, Phagosome, p53 signaling pathway, Rap1 signaling pathway, Focal adhesion, Proteoglycans in cancer, Malaria, Bladder cancer
Plcb4	Metabolic pathways, Adrenergic signaling in cardiomyocytes, Insulin secretion, Alzheimer's disease, Platelet activation, Phospholipase D signaling pathway, Glutamatergic synapse, cGMP-PKG signaling pathway, Calcium signaling pathway, Inflammatory mediator regulation of TRP channels, Pathways in cancer, Circadian entrainment, Cholinergic synapse, Melanogenesis, Estrogen signaling pathway, Wnt signaling pathway, Sphingolipid signaling pathway, Vascular smooth muscle contraction, Retrograde endocannabinoid signaling, Serotonergic synapse, GnRH signaling pathway, Oxytocin signaling pathway, Renin secretion, Aldosterone synthesis and secretion, Chemokine signaling pathway, Glucagon signaling pathway, Dopaminergic synapse, Rap1 signaling pathway, Phosphatidylinositol signaling system, Long-term potentiation, Salivary secretion, Gastric acid secretion, Huntington's disease, Pancreatic secretion, Thyroid hormone synthesis, AGE-RAGE signaling pathway in diabetic complications, Chagas disease (American trypanosomiasis), African trypanosomiasis, Amoebiasis, Long-term depression, Gap junction, Inositol phosphate metabolism, Thyroid hormone signaling pathway, Endocrine and other factor-regulated calcium reabsorption

Car2	Gastric acid secretion, Pancreatic secretion, Bile secretion, Collecting duct acid secretion, Proximal tubule bicarbonate reclamation, Nitrogen metabolism
Nbea	
Tiparp	
Map9	
Trim2	
Pdlim5	
Adgrl2	
Tln1	Platelet activation, HTLV-I infection, Rap1 signaling pathway, Focal adhesion
Zfp462	
Dnajc25	
Nfib	
Mrpl37	
Ebna1bp2	
Zcchc17	
Id3	TGF-beta signaling pathway, Signaling pathways regulating pluripotency of stem cells
Lrpap1	
Wdr1	
Limch1	
Cdkl2	
Anxa3	
Tgfbr3	
Coro1c	
Ankrd13a	
Pdgfa	MicroRNAs in cancer, Phospholipase D signaling pathway, Regulation of actin cytoskeleton, Pathways in cancer, Cytokine-cytokine receptor interaction, PI3K-Akt signaling pathway, MAPK signaling pathway, EGFR tyrosine kinase inhibitor resistance, Ras signaling pathway, HTLV-I infection, Rap1 signaling pathway, Glioma, Transcriptional misregulation in cancer, Focal adhesion, Melanoma, Gap junction, Prostate cancer, Choline metabolism in cancer
Lnx2	
Hsph1	Protein processing in endoplasmic reticulum
Frmd4b	
Arhgdib	Neurotrophin signaling pathway, Vasopressin-regulated water reabsorption
Myadm	
U2af2	Spliceosome
Pla2g4c	Metabolic pathways, Platelet activation, Phospholipase D signaling pathway, Glutamatergic synapse, Inflammatory mediator regulation of TRP channels, MAPK signaling pathway, Vascular smooth muscle contraction, Serotonergic synapse, GnRH signaling pathway, Oxytocin signaling pathway, Ovarian steroidogenesis, Ras signaling pathway, Glycerophospholipid metabolism, Arachidonic acid metabolism, Fc epsilon RI signaling pathway, Long-term depression, Linoleic acid metabolism, VEGF signaling pathway, Choline metabolism in cancer, Ether lipid metabolism, alpha-Linolenic acid metabolism
Ltbp4	
Actn4	Regulation of actin cytoskeleton, Tight junction, Leukocyte transendothelial migration, Arrhythmogenic right ventricular cardiomyopathy (ARVC), Viral carcinogenesis, Systemic lupus erythematosus, Focal adhesion, Adherens junction, Amoebiasis
Nav2	
Arrdc4	
Nr2f2	

Folr1	Endocytosis
Mrpl17	Ribosome
Irs2	MicroRNAs in cancer, FoxO signaling pathway, cGMP-PKG signaling pathway, Type II diabetes mellitus, Longevity regulating pathway - multiple species, Insulin signaling pathway, Longevity regulating pathway, AMPK signaling pathway, Non-alcoholic fatty liver disease (NAFLD), Insulin resistance, Adipocytokine signaling pathway, Regulation of lipolysis in adipocytes
Ntm	
Rab27a	

Table 31. Genes upregulated commonly between *Fermt2*, *Actg1* and *Actb* knockout cells in comparison to wild-type cells and their associated KEGG pathway (if available).
Log_2 (fold change) (L2F) knockout > wildtype and $P \leq 0.05$. *P* values are not multiple testing corrected. DESeq2 tests for significance of coefficients in a negative binomial GLM (Generalized Linear Model) with the Wald test.

A first-pass analysis of the genes upregulated in each knockout cell line model showed that a disproportionate number of those genes displayed sequence similarity with the mRNA of the mutated gene. Multiple studies in the past decade have shown that mRNA decay and gene expression are coupled processes (Elkon et al., 2010; Hao and Baltimore, 2009; Sun et al., 2012), whereby following mRNA decay, certain mRNA decay factors can translocate back to the nucleus and induce gene expression through promoting transcription initiation and elongation (Haimovich et al., 2013). Besides, as mentioned in the introduction, several recent studies have sequenced RNA degradation intermediates and identified a range of fragments that vary widely in size (Ibrahim et al., 2018; Ibrahim and Mourelatos, 2019; Kurosaki et al., 2018; Peach et al., 2015; Pelechano et al., 2016; Ueno et al., 2018; Valen et al., 2011), suggesting a possible biological function for such fragments. I thereby hypothesized that following mutant mRNA degradation, the degradation intermediates can act as guides for certain decay factors or other RNA binding proteins, to bring them to loci of genes exhibiting sequence similarity such as paralogs in order to induce their expression. Indeed, exploring the relationship between sequence similarity and upregulation of genes in the RNA-seq data, based on multiple similarity thresholds, revealed a significant correlation at certain values that varied between the different models (Figure 15 and subsection **5.2.42** of the Methods section in this thesis).

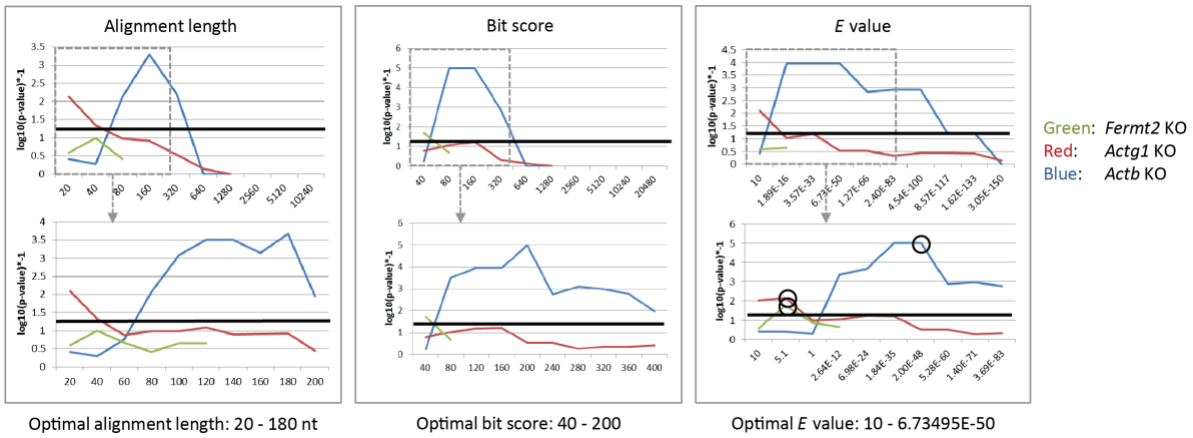


Figure 15. Analysis of parameters of sequence similarity parameters in models of transcriptional adaptation.

Impact of several values of three quality parameters for BLASTn alignment (alignment length, Bit score, *E* value) on the significance of the correlation observed between sequence similarity and up-regulation, and therefore identification of potential adapting genes. *E* value represents the probability of the alignment resulting by chance (the lower it is, the better), Bit score evaluates both alignment length and quality (the higher it is, the better). For each diagram, the Y-axis shows the negative $\log_{10}$ of the significance *P* value (the higher it is, the better), and the X-axis shows the respective parameter value. A black horizontal line marks the *P* value of 0.05. The *E* value thresholds that I used in the analyses are marked by a circle. Lines ending preliminarily imply the lack of further alignments following that value. The second raw of the diagrams is closer look-ups of the boxed regions in the first row of the diagrams. Depending on the analyzed gene, the optimal thresholds differ markedly. n = 2 independent biological samples. *P* values are not multiple testing corrected. DESeq2 tests for significance of coefficients in a negative binomial GLM (Generalized Linear Model) with the Wald test. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

Using the optimal *E*-values (a value that represents the probability of an alignment resulting by chance (the lower it is, the less likely it is to be due to chance)), of the sequence similarity alignments to identify what I hereafter refer to as 'similar' protein-coding genes (the values being 5.1 for both the *Fermt2* and *Actg1* models and 2×10^{-4} for *Actb* (the stricter *E* value for *Actb* was necessary due of its repetitive 3′UTR, that lead to 'noisy' misleading matches)), I observed that at least 50 to 60% of those similar genes were upregulated significantly in the knockout cells, compared to a maximum of 10 to 21% when analyzing randomly selected genes not exhibiting sequence similarity (Figure 16A, Table 32, 33, 34). Interestingly, out of the 12 upregulated 'similar' genes in *Actg1* knockout MEFs, 7 did not display an upregulation in *Actg1* RNA-less allele (Table 33). Similarly, 4 of the 6 upregulated 'similar' genes in *Actb* knockout mESCs were not increased in expression in the *Actb*

RNA-less mESCs (Table 34). Furthermore, I observed that 4 of the 9 not-upregulated 'similar' genes on the mRNA level in *Actg1* knockout MEFs were upregulated on the pre-mRNA level (Figure 16B), suggesting that transcriptional adaptation may involve the upregulation of more than 50-60% of 'similar' genes, whose levels might be further modulated post-transcriptionally by the cell to avoid potential harmful effect from overexpression of certain genes.

Zebrafish *actb1* and mouse *Actb* transcripts share a very high degree of sequence similarity (Appendix section **12.1**). To further validate the requirement of sequence similarity, I injected uncapped mouse *Actb* RNA into 1-cell stage zebrafish embryos and interestingly, observed an upregulation of the zebrafish *actb1* gene expression (Figure 17A). To further confirm the requirement of sequence similarity, I generated uncapped synthetic transcripts of *hif1ab* containing either only sequences that exhibiting sequence similarity with *epas1a* locus or sequences that do not align to *epas1a* (Figure 7b, Appendix sections **12.2**, **12.3**). Notably, only the uncapped transcripts that contain sequences exhibiting similarity with *epas1a* led to *epas1a* upregulation but not the uncapped transcript containing the non-similar sequences (Figure 17C). Taken together, these data suggest that sequence similarity influences transcriptional adaptation, at least in some cases.

In an attempt to try and understand whether the localization of the similarity influences the upregulation of 'similar' genes, I generated different uncapped synthetic transcripts composed of sequences of *hif1ab* that align to either *epas1a* promoter region, exons, introns, or 3'UTR. Interestingly, uncapped transcripts containing sequences exhibiting similarity with *epas1a* exons or introns led to the strongest induction of *epas1a* upregulation, while those sharing sequence similarity with the 3'UTR did not induce a transcriptional adaptation response and those corresponding to *epas1a* promoter region led to a milder upregulation response (Figure 7D). Notably, these data were accordant to the RNA-seq analysis of the knockout mouse cells lines (Tables 32, 33, 34) where genes sharing sequence similarity to the mutant gene's mRNA in their 3'UTR were not increased in expression, and those sharing sequence similarity to their promoter regions were mildly upregulated. These data indicated, that at least in some cases, the localization of sequence similarity can influence transcriptional adaptation. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

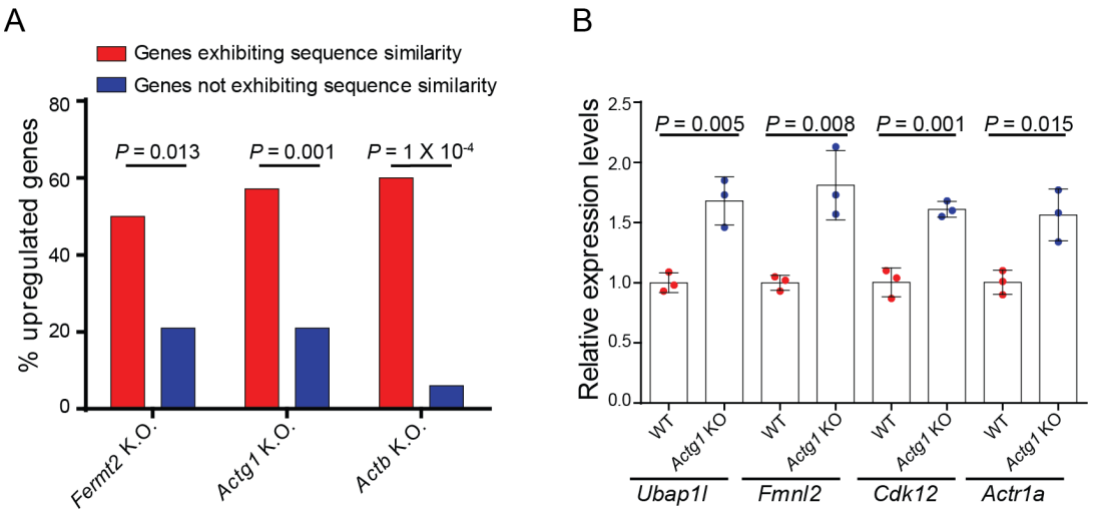


Figure 16. Transcriptome analysis of the mouse knockout cells reveals upregulation of a significant proportion of genes exhibiting sequence similarity with the mutant gene's mRNA.

A) Percentage of significantly upregulated protein-coding genes (Log_2 (Fold change) knockout > wild-type and *P* value ≤ 0.05) sharing sequence similarity with *Fermt*2, *Actg1* and *Actb* compared to the percentage of genes not sharing sequence similarity using the optimal *E* values. B) qPCR analysis of pre-mRNA expression levels of *Ubapl*, *Fmnl2*, *Cdk12* and *Actr1a* in wild-type and *Actg1* knockout MEFs. n= 2 (A), 3 (B) independent biological samples. Wild-type expression levels were set at 1 for the analyses in panels B. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

Gene	L2F *Fermt2* KO/wt	*P*val	*P*adj
Fermt1	8.055307305	5.19E-43	1.94E-41
Dchs2	3.501304705	4.92E-11	3.99E-10
Paqr5	1.679449433	2.46E-13	2.41E-12
Mlxipl	0.618877214	0.409259	0.573342
Pcnx	0.502076808	8.12E-06	3.76E-05
Nkain1	0.39205614	0.043898	0.097439
Vapb	0.386807243	2.47E-06	1.22E-05
Strbp	0.273294934	0.000195	0.000738
Prkar2b	0.271409666	0.006358	0.017971
Pou2f1	0.191075348	0.041339	0.092619
Depdc1b	0.097436998	0.581669	0.726092
Fermt3	0.035722034	0.948806	0.970063
Mfsd11	-0.085901914	0.621497	0.757444
Trim24	-0.255516547	0.00941	0.025447
Atm	-0.516886899	3.84E-13	3.70E-12
Slc25a13	-0.678800185	1.40E-06	7.14E-06
Abhd18	-0.82015185	0.024339	0.05872
Eda	-1.503214668	0.167483	0.294635

Table 32. RNA-sequencing analysis of genes sharing sequence similarity with *Fermt2* in knockout MKFs compared to wild-type.

L2F: Log_2(fold change). Bold refers to 'similar' genes upregulated significantly in *Fermt2* knockout MKFs relative to wild-type. Red refers to L2F>0 and blue refers to L2F<0. Green indicates a *P* value or *P* adjusted value ≤ 0.05. Violet refers to genes sharing sequence similarity with the mutant gene's mRNA in their promoter region. Non-colored genes share sequence similarity with the mutant gene's mRNA in their introns or exons. n = 2 independent biological samples. *P* values are not multiple testing corrected. DESeq2 tests for significance of coefficients in a negative binomial GLM (Generalized Linear Model) with the Wald test. Table adapted and reprinted with permission from (El-Brolosy et al., 2019).

Gene	L2F *Actg1* KO/wt	*P*val	*P*adj
Actg2	5.86	0	0
Ptp4a3	3.91	0.00000246	1.40E-05
Hcn1	3.47	1.84E-09	1.45E-08
Adamts12	2.14	4.91E-46	2.53E-44
Sec24d	0.81	4.36E-20	7.67E-19
Acta2	0.76	6.48E-45	3.22E-43
Hcfc2	0.72	2.91E-08	2.05E-07
Ubap1l	0.45	0.573535144	0.743877533
Pde10a	0.44	0.000216157	0.000933738
Nol11	0.43	9.8E-11	8.69E-10
Actr1b	0.29	0.000631774	0.002508725
Atf7	0.29	0.002212762	0.007853264
Actb	0.19	0.000162494	0.00071632
Znrf2	0.13	0.270233353	0.458179229
Lrrc58	-0.15	0.019456235	0.054271242
Fmnl2	-0.2	0.008174843	0.025396546
Cdk12	-0.24	0.000463124	0.001889135
Actr1a	-0.28	0.000780469	0.003040052
Fbxl17	-0.49	0.00000017	0.000284708
Grpel2	-0.58	0.0000601	0.45255246
Acta1	-1.91	1.01E-25	2.36E-24

Table 33. RNA-sequencing analysis of genes sharing sequence similarity with *Actg1* in knockout MEFs compared to wild-type.

L2F: Log_2(fold change). Bold refers to 'similar' genes upregulated significantly in *Actg1* knockout MEFs relative to wild-type. Red refers to L2F>0 and blue refers to L2F<0. Green indicates a *P* value or *P* adjusted value ≤ 0.05. Violet refers to 'similar' genes sharing sequence similarity with the mutant gene's mRNA in their promoter. Yellow refers to 'similar' genes sharing sequence similarity with the mutant gene's mRNA in their 3'UTR. Non-colored genes share sequence similarity with the mutant gene's mRNA in their introns or exons. Boxes refer to 'similar' genes being upregulated in the knockout cells but not the RNA-less allele. n = 2 independent biological samples. *P* values are not multiple testing corrected. DESeq2 tests for significance of coefficients in a negative binomial GLM (Generalized Linear Model) with the Wald test. Table adapted and reprinted with permission from (El-Brolosy et al., 2019).

Gene	L2F *Actb* KO/wt	*P*val	*P*adj
Acta2	7.948538835	5.34E-151	7.98E-147
Actg2	2.961946198	0.020429	1
Acta1	2.556443324	0.001842	0.036378
Actc1	2.284580096	0.017007	1
Rasl2-9	1.353877161	0.350521	1
Actg1	0.982444337	3.78E-11	7.63E-09
Pde10a	0.552114557	0.432741	0.833262
Fbxl17	0.524667754	0.032681	0.244749
Gm17087	0.158304112	0.936537	0.989226
Lrrc58	0.028544142	0.851949	0.974102

Table 34. RNA-sequencing analysis of genes sharing sequence similarity with *Actb* in knockout mESCs compared to wild-type.

L2F: Log_2(fold change). Bold refers to 'similar' genes upregulated significantly in *Actb* knockout mESCs relative to wild-type. Red refers to L2F>0 and blue refers to L2F<0. Green indicates a *P* value or *P* adjusted value ≤ 0.05. Violet refers to 'similar' genes sharing sequence similarity with the mutant gene's mRNA in their promoter. Yellow refers to 'similar' genes sharing sequence similarity with the mutant gene's mRNA in their 3'UTR. Non-colored genes share sequence similarity with the mutant gene's mRNA in their introns or exons. Boxes refer to 'similar' genes being upregulated in the knockout cells but not the RNA-less allele. n = 2 independent biological samples. *P* values are not multiple testing corrected. DESeq2 tests for significance of coefficients in a negative binomial GLM (Generalized Linear Model) with the Wald test. Table adapted and reprinted with permission from (El-Brolosy et al., 2019).

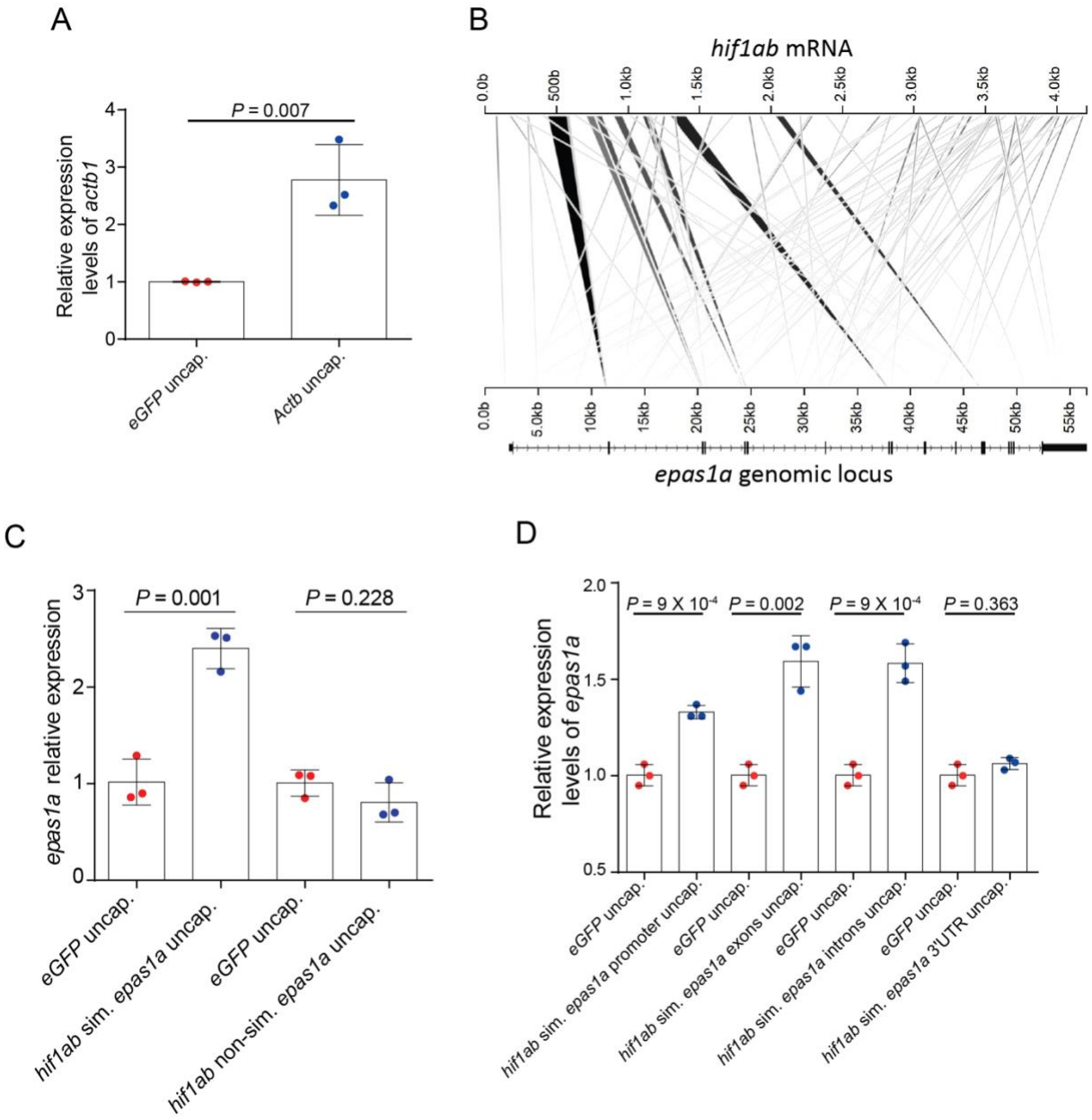


Figure 17. Sequence similarity plays a role in inducing transcriptional adaptation.

A) qPCR analysis of mRNA expression levels *actb1* in 6 hpf wild-type zebrafish embryos injected with uncapped *eGFP* as a control or mouse *Actb1* RNA. B) Schematic representation of sequence similarity regions between the *epas1a* locus (gene body + 2kb upstream of TSS) and *hif1ab* mRNA. The maximum *E* value was set to 25 to allow better visibility. The grey triangles represent the alignments and their intensities refer to the quality of the alignment, while their width at the base refers to the similarity region's length. TSS: transcription start site. C) qPCR analysis of mRNA expression levels of *epas1a* in 6 hpf wild-type zebrafish injected with uncapped *eGFP* as control or uncapped RNAs composed only of *hif1ab* mRNA sequences that share sequence similarity *epas1a* locus (*hif1ab* sim.) or only *hif1ab* mRNA sequences that do not exhibit sequence similarity to *epas1a* (*hif1ab* non-sim.). D) qPCR analysis of *epas1a* mRNA expression levels in wild-type embryos injected with uncapped transcripts composed exclusively of *hif1ab* sequences exhibiting sequence similarity to *epas1a* promoter, exons, introns, or 3'UTR at 6hpf. n= 3 (A, C, D) independent biological

samples. Wild-type or control expression levels were set at 1 for the analyses in panels A, C and D. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

6.5. Epigenetic remodeling at adapting genes transcription start-sites.

As previously mentioned, several studies in the past decade have reported that mRNA degradation and gene expression are coupled processes (Elkon et al., 2010; Hao and Baltimore, 2009; Sun et al., 2012), an effect that is possibly mediated through translocation of decay factors to the nucleus (Haimovich et al., 2013). Having identified a role for sequence similarity in transcriptional adaptation, it allowed me to build a model whereby following mutant mRNA decay, degradation intermediates/fragments can translocate back to the nucleus with some decay factors, or other RNA binding proteins, to guide them to loci of the adapting genes, through homology-mediated base-pairing, and induce their expression (Figure 18). Several studies have reported interactions between decay factors and histone modifiers or chromatin remodelers (Berretta et al., 2008; Collins et al., 2007; Pinskaya et al., 2009). In order to understand, how decay factors can contribute to inducing the adapting genes' expression, I performed a small RNAi screen in the *Rela* knockout MEFs targeting major chromatin remodelers and histone modifiers that are known to be involved in transcription activation (Table 21; section **5.1.18.1** of the thesis). Knockdown of the lysine demethylases JMJD2 and KDM6, which remove the repressive histone marks H3K9me3 and H3K27me3, respectively, led to a decrease in *Rel* upregulation levels (Figure 19A). *Rel* upregulation was however completely dampened upon knockdown of WDR5, a component of the COMPASS complex that deposits the permissive H3K4me3 histone mark (Ruthenburg et al., 2006; Shilatifard, 2012; Sims et al., 2003; Wysocka et al., 2005) (Figure 19A). Accordingly, chromatin immunoprecipitation experiments (ChIP) revealed an increased enrichment of WDR5 and its associated H3K4me3 mark at the transcription start sites of *Fermt1*, *Rel,* and *Actg2* in *Fermt2*, *Rela* and *Actg1* knockout cells, respectively (Figure 19 B-D). Interestingly, knockdown of the decay factors UPF1 with EXOSC4 or XRN1 led to a reduction in H3K4me3 levels at the transcription start site of *Rel* in *Rela* knockout MEFs (Figure 19E). Taken together, these data proposed a model whereby following mutant mRNA degradation, certain decay factors translocate back to the nucleus to bind to the adapting genes' loci,

possibly guided by the mRNA degradation intermediates, and recruit chromatin remodelers or histone modifiers, that can lead to enhancing the adapting genes' transcription levels. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

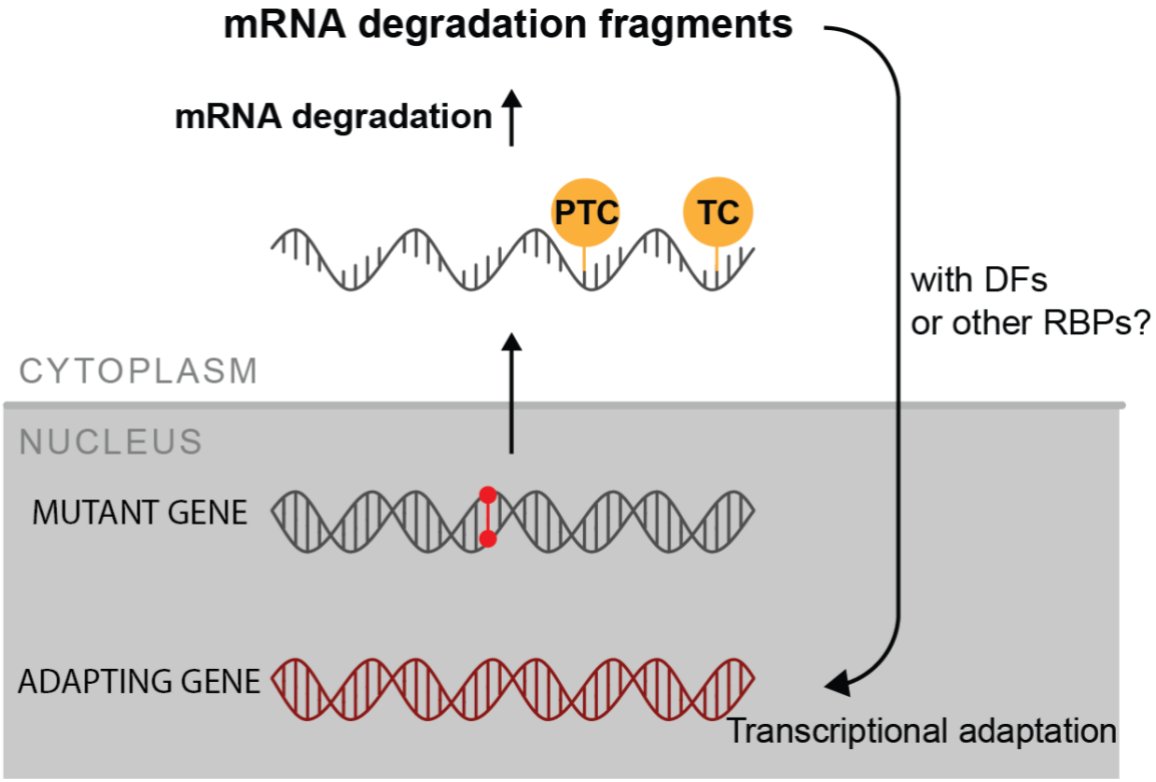


Figure 18. Model of transcriptional adaptation.

A simplified putative model for the transcriptional adaptation phenomenon. PTC: premature termination (stop) codon; TC: termination (stop) codon; RBPs: RNA binding proteins; DFs: decay factors. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

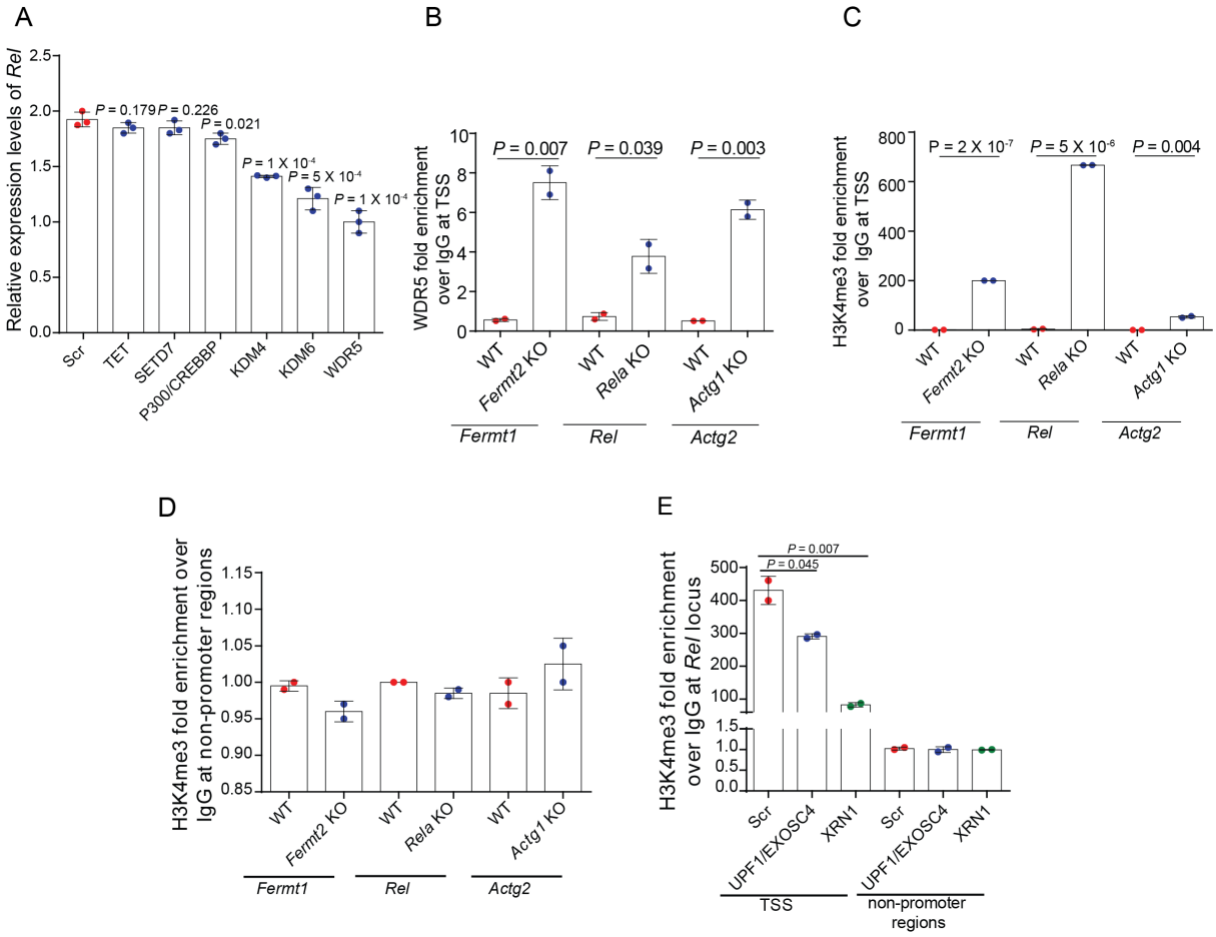


Figure 19. Decay factor-dependent epigenetic remodeling in transcriptional adaptation.

A) qPCR analysis of mRNA expression levels *Rel* in *Rela* knockout MEFs after knockdown of the indicated genes using siRNA. B) ChIP-qPCR analysis of the occupancy of WDR5 close to the transcription start sites (TSS) of *Fermt1*, *Rel* and *Actg2* in wild-type cells or *Fermt*2, *Rela* and *Actg1* knockout cells, respectively. C) ChIP-qPCR analysis of the occupancy of H3K4me3 marks close to the transcription start sites (TSS) of *Fermt1*, *Rel* and *Actg2* in wild-type cells or *Fermt*2, *Rela* and *Actg1* knockout cells, respectively. D) ChIP-qPCR analysis of the occupancy of H3K4me3 marks at non-promoter regions (within the gene body; as a control) of *Fermt1*, *Rel* and *Actg2* in wild-type cells or *Fermt*2, *Rela* and *Actg1* knockout cells, respectively. E) ChIP-qPCR analysis of the occupancy of H3K4me3 marks close to the transcription start sites (TSS) of *Rel* or at a non-promoter region (within the gene body; as a control) following knockdown of UPF1/EXOSC4 or XRN1 using siRNA in *Rela* knockout MEFs. A, E) Scr: Scrambled siRNA control. n= 3 (A), 2 (B-E) independent biological samples. Wild-type or control expression levels were set at 1 for the analyses in panels B, C and D. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

6.6. Potential role for antisense RNAs in transcriptional adaptation.

Natural antisense transcripts are non-coding RNAs that are transcribed from the opposite strand of the DNA in a region encoding a given gene. In most of the cases they act as negative regulators of the sense gene's expression and multiple models have been proposed to explain their inhibitory effects (Faghihi and Wahlestedt, 2009; Wery et al., 2018). According to one model, they can form a double-stranded RNA duplex with the sense transcript triggering an RNAi response that leads to degradation of the sense transcript, or they can alter the processing of the sense RNA through interfering with its capping, polyadenylation or export, which ultimately affects the sense gene's expression levels (Borsani et al., 2005; Faghihi and Wahlestedt, 2009; Osato et al., 2007). According to another model, the antisense RNAs can act as scaffolds for DNA methyltransferases and histone modifiers, leading to silencing of the sense gene's expression through DNA methylation (Tufarelli et al., 2003) or creating a less-permissive chromatin environment, for example through increased deposition of H3K27me3 marks (Modarresi et al., 2012; Morris et al., 2008; Yu et al., 2008), respectively.

Interestingly, a previous study reported that transfection of short RNA fragments corresponding to the *Cdk9* and *Sox9* mRNAs leads to increased expression of those genes (Ghanbarian et al., 2017). Mechanistically, the authors reported that such fragments led to the downregulation of antisense transcripts present at the *Cdk9* or *Sox9* loci which normally represses the expression of the sense RNAs. Accordingly, I observed an upregulation of *Cdk9* and *Sox9* mRNA expression levels upon transfection of uncapped *Cdk9* or *Sox9* transcripts, respectively, into wild-type cells (Figure 20A). Moreover, transfection of uncapped *BDNF* transcripts into wild-type HEK293T cells led to downregulation of *BDNF* antisense transcript levels and concomitant upregulation of *BDNF* mRNA (sense RNA) expression levels (Figure 20B), in agreement with a previous report that showed that knockdown of *BDNF* antisense transcript leads to increased expression of the sense transcript, a response that was linked to decreased H3K27me3 marks (Modarresi et al., 2012). Notably, antisense transcripts at the adapting genes *vclb* and *hbegfb* loci were also downregulated in *vcla*$^{\Delta 13}$ and *hbegfa*$^{\Delta 7}$ mutant zebrafish (Figure 20C, D), and interestingly these antisense transcripts shared sequence similarity with *vcla* and *hbegfa* mRNAs, respectively (Figure 20C). Taken together, these data suggest that another mechanism underlying transcriptional adaptation may involve mRNA decay

fragments acting on antisense RNAs within the adapting genes' loci, in a sequence-specific manner. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

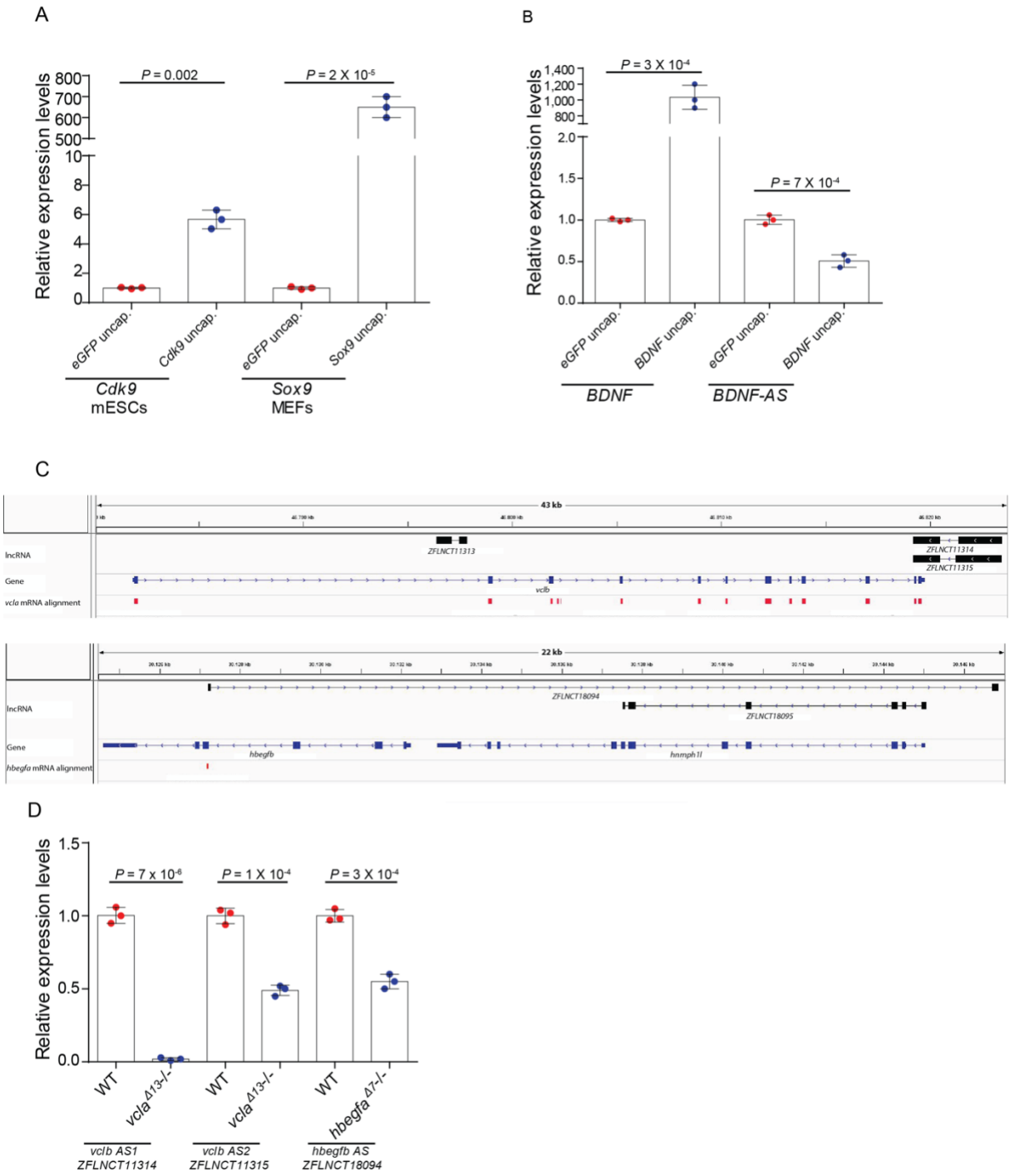


Figure 20. A potential role for antisense transcripts in transcriptional adaptation.

A) qPCR analysis of the mRNA expression levels *Cdk9* and *Sox9* in mESCs or MEFs transfected with uncapped *eGFP* (as control) or *Cdk9* or *Sox9* transcripts. B) qPCR analysis of the mRNA expression levels *BDNF* and *BDNF antisense* (*BDNF-AS)* in wild-type HEK293T cells transfected with uncapped *eGFP* (as control) or *BDNF* transcripts. C)

Screenshot from the integrated genome viewer (IGV) browser showing tracks of the *vclb* and *hbegfb* loci and the corresponding annotated antisense RNAs (lncRNAs). Two alignments of a length of 147 and 105 bp between *vcla* mRNA and *vclb* antisense transcripts were observed, and a single alignment of a length of 39 bp between *hbegfa* mRNA and *hbegfb* antisense transcript (red boxes in the third tracks of the IGV screenshots). Antisense transcripts shown in the lncRNA track were obtained from the GSE32898 dataset (Pauli et al., 2012). D) qPCR analysis of the expression levels of two *vclb* antisense transcripts and the *hbegfb* antisense transcript in wild-type and *vcla*$^{\Delta 13}$ or *hbegfa*$^{\Delta 7}$ mutant zebrafish, respectively. A, B, D) n= 3 independent biological samples. Wild-type or control expression levels were set at 1 for the analyses. *P* values were calculated by a two-tailed Student's t-test and data are mean ± s.d. Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

7. Discussion

Despite being a widespread phenomenon (El-Brolosy and Stainier, 2017) and its relevant implications, the molecular mechanisms underlying transcriptional adaptation were not clearly understood. In this thesis, I show that mutant mRNA degradation triggers the upregulation of related genes in a sequence-specific manner. My data suggested two potential models: a) following mutant mRNA decay, degradation factors, or other RNA binding proteins, translocate back to the nucleus along with decay intermediates, guiding them in a sequence-dependent manner to the adapting genes' loci to induce gene expression through recruiting chromatin remodelers and histone modifiers. According to the other model b) decay intermediates may repress antisense transcripts and thereby allow for derepression and increased expression of the sense mRNA (Figure 21). It is possible that other mechanisms, that are yet to be identified, also exist, and which might be even gene-specific.

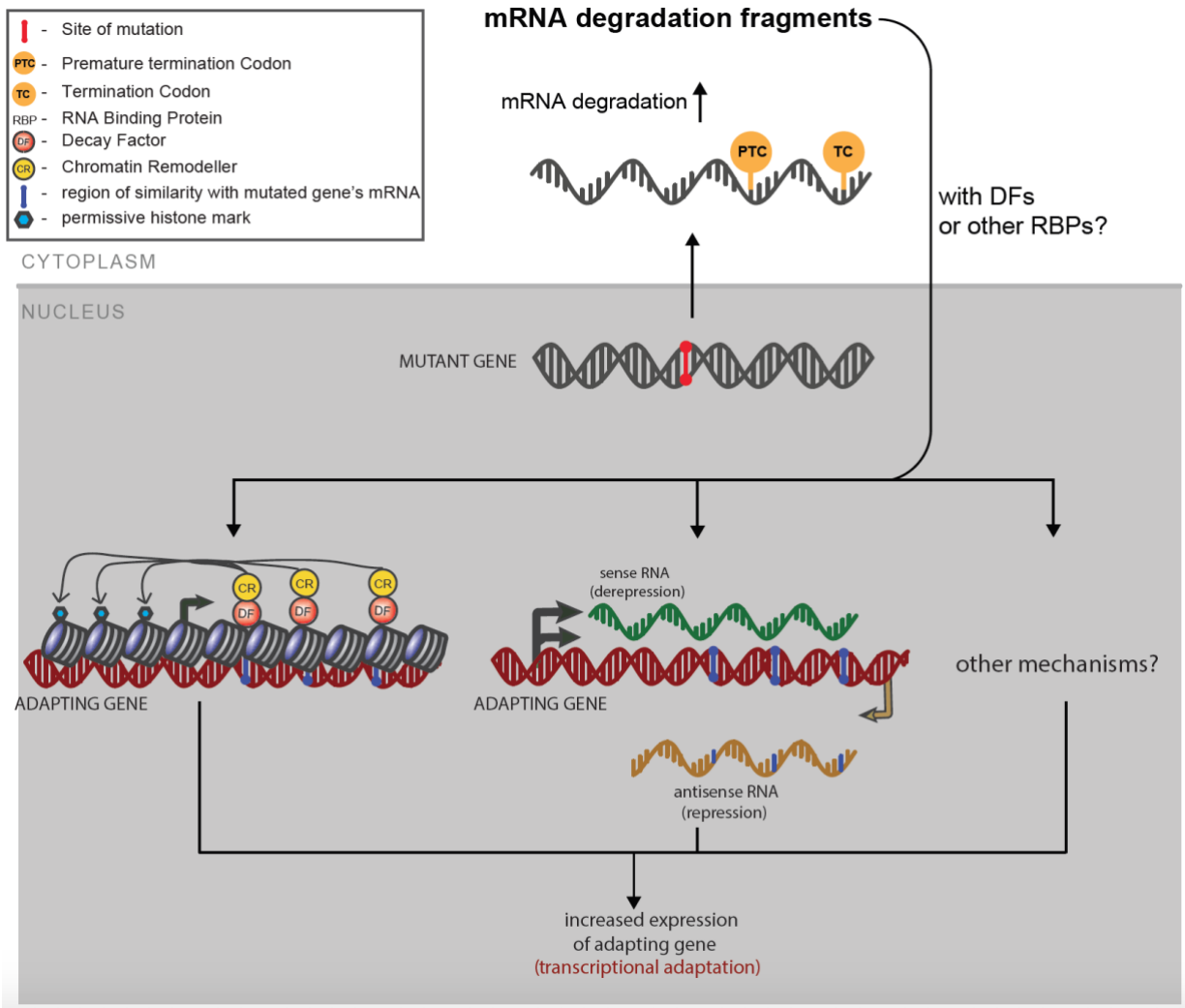


Figure 21. Expanded model transcriptional adaptation.

Figure adapted and reprinted with permission from (El-Brolosy et al., 2019).

7.1. Transcriptional adaptation modifies expression of related genes

I focused my initial analyses on paralogous genes as they are the most likely to compensate for each other's functions. In zebrafish, I identified 5 models (*hbegfa*, *vcla*, *hif1ab*, *egfl7,* and *alcama*) whereby an indel mutation leads to the upregulation of a paralogue, and 4 in mouse cell lines (*Fermt2*, *Rela*, *Actg1,* and *Actb*) (Figure 4A, 5A). However, I also studied another model where the adapting gene was not a paralog (*egfl7-emilin3a*) (Figure 4A) and found through transcriptome analysis of the mouse cell line models that transcriptional adaptation is not limited to paralogs and involves the upregulation of other genes exhibiting sequence similarity with the mutant gene's mRNA (Figure 16A, Table 32, 33 and 34). While, I have observed that hundreds of genes are upregulated in the knockout mouse cells, most of which are potentially due to direct and indirect effects of the loss of the protein function, it is important to note that most of the genes exhibiting sequence similarity where upregulated only in alleles displaying mutant mRNA decay but not the RNA-less alleles (Table 33, 34), suggesting that they are direct targets of transcriptional adaptation. A third group of genes might be possibly upregulated as an indirect consequence of transcriptional adaptation. For example, due to being involved in a transcriptional network with the upregulated genes. Further temporal studies of transcriptional adaptation would be required to differentiate between direct and indirect targets of the response.

Previous reports studying heterozygous mutations reported increased expression of the wild-type allele due to disruption of an inhibitory feedback loop (Guidi et al., 2004; Trieu et al., 2003). I have also observed that transcriptional adaptation involves modulation of the mutated gene. Metabolic labeling experiments revealed increased transcription of the mutant gene (Figure 9C) and heterozygous fish models displayed an upregulation of the wild-type allele (Figure 4D), indicating that transcriptional adaptation might explain cases of haplosufficiency. In agreement with my data, a previous study reported increased accumulation of pre-mRNAs harboring a PTC near the transcription sites of the mutant genes (Muhlemann et al., 2001).

It is, however, important to note that not all cases of transcriptional adaptation can lead to a functional compensation response; for example, *vegfaa* mutants still display strong vascular defects despite the upregulation of its paralog *vegfab* (Rossi et al., 2016; Rossi et al., 2015). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

7.2. Mutant mRNA decay and transcriptional adaptation

All of the mutant models of transcriptional adaptation studied in this thesis displayed mutant mRNA decay (Figure 8 B-D, 9 A-F). Interestingly, ENU mutant alleles of *hbegfa* and *vcla* did not display a transcriptional adaptation response, correlating with absence, or very minor levels, of mutant mRNA decay in those alleles (Figure 8A, B). Absence of a transcriptional adaptation response in *hbegfa* and *vcla* point mutants (ENU-induced) but not in the indel alleles (CRISPR-induced) is not due to the type of mutation or the mutagen, as data from the lab showed that *alcama* and *hif1ab* ENU-induced point mutants displayed mutant mRNA decay and a transcriptional adaptation response. Presence of a PTC doesn't always lead to nonsense-mediated decay (e.g., *hbegfa* and *vcla* ENU alleles do harbor a PTC) and a number of reasons can explain that: 1) The nature of the stop codon can influence NMD efficiency, for example, amber stop codons (UAG) are more likely to be readthrough during translation, and thereby transcripts containing it are more likely to escape non-sense mediated decay, than an ochre stop codon (UAA) (Dabrowski et al., 2015) (the *hbegfa*^{sa18135} allele (ENU-induced) has an amber stop codon while the *hbegfa*$^{\Delta 7}$ allele (CRISPR-induced) has an ochre one (Figure 3)); 2) NMD efficiency is also highly determined by the distance from the next exon-junction complex (a PTC needs to be at least 50-55 bp away from the next exon-junction complex for NMD to be induced efficiently (e.g., the *vcla*$^{exon22\ ins1}$ allele has the PTC in the last exon, after which there is no EJC, and thereby such allele doesn't display NMD)). Indel-induced frameshift mutations may also lead to introduction of rare codons at which the ribosome is more likely to stall, and thereby eliciting stronger mRNA decay (Presnyak et al., 2015).

Consistent with my observations, a zebrafish study on Metallothionein 2 (*mt2*) reported that mutant alleles of *mt2* with the strongest levels of mutant mRNA decay displayed milder vascular phenotypes compared to alleles with lower levels of mRNA decay (Schuermann et al., 2015). Moreover, morpholino-mediated knockdown of the decay factors Smg1 and Upf1 in the *mt2* mutant allele displaying strong mutant

mRNA decay led to increased number of embryos displaying vascular defects. Studying the *hbegfa* gene, I observed a correlation between transcriptional adaptation and the extent of mRNA decay. The *hbegfa*$^{\Delta 7}$ allele displayed 50% reduction of the mutant mRNA levels and an upregulation of the paralogue *hbegfb* while the *hbegfa*^{sa18135} allele displayed only around 20% reduction in the mutant transcript levels and no transcriptional adaptation response (Figure 8A, B). In addition, I observed, in zebrafish, that upregulation of the adapting genes was more significant at the stages where the mutant gene had the highest expression in early development (Figure 4A). Moreover, heterozygous conditions displayed lower upregulation levels of the adapting genes than the homozygous counterparts (Figure 4C). CRISPRi-mediated knockdown of the mutant *Fermt2* expression levels also led to a decrease in the expression levels of the paralogous gene *Fermt1* in *Fermt2* knockout cells (Figure 12F). Taken together, these data indicate that mutant mRNA degradation levels influence the extent of inducing a transcriptional adaptation response. It shall be interesting to further investigate the minimal levels of degradation required, and whether increasing the expression levels of the mutant gene, for example using CRISPR activation (CRISPRa) (Konermann et al., 2015), can lead to increased upregulation levels of the adapting genes.

Knockdown of some degradation factors had different effects on stabilization of the mutant mRNA in *Rela* and *Actb* knockout models, in addition to zebrafish models. For example, knockdown of UPF1 and EXOSC4 in *Rela* knockout cells led to stabilization of the mutant *Rela* transcript, however, no stabilization of the *Actb* mutant transcript was observed upon knocking them down in *Actb* knockout mESCs (Figure 10 E, F). Similarly, while *upf1* mutations in *hbegfa*, *vcla,* and *vegfaa* mutant background led to stabilization of the respective mutant transcripts (Figure 10A), *upf1*; *egfl7* mutants did not display increased stability of the *egfl7* mutant mRNA. Such difference can be explained by the fact that different mutant transcripts may be degraded through different RNA decay machinery. For example, the *egfl7*$^{\Delta 4}$ transcript is potentially subjected to no-go decay, rather than nonsense-mediated decay, as prediction of the mRNA secondary structure (Gruber et al., 2008) revealed severe changes in its secondary structure and introduction of new loops that may trigger no-go decay (Figure22). Thereby, potentially explaining its non-sensitivity to loss of *upf1*. It is however important to note that I have always observed, whenever there was a partial or full stabilization of the mutant mRNA, a reduced or diminished

transcriptional adaptation response. Accordingly, I also observed that ectopic induction of RNA decay in wild-type embryos or cells, through introduction of uncapped transcripts, induced a transcriptional adaptation response (Figure 11A, C). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

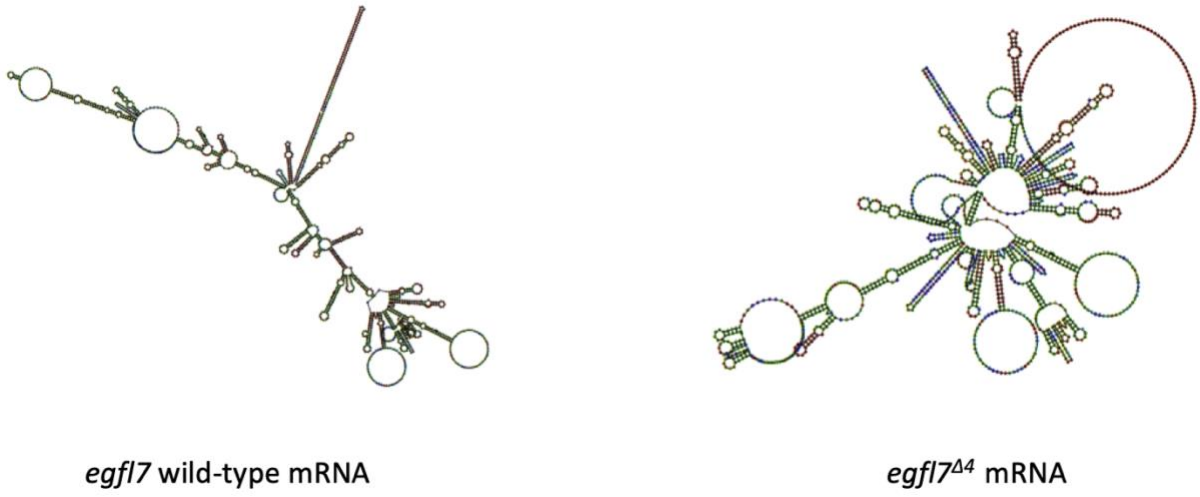


Figure 22. Predicted secondary structures for *egfl7* wild-type and mutant mRNAs.

7.3. Guidelines for generation of mutant alleles

In this thesis, I also provide guidelines on how to efficiently generate mutant alleles that can uncover phenotypes masked through transcriptional adaptation-induced compensation. I showed that alleles that fail to transcribe the mutant gene (RNA-less alleles; generated by deletion of promoter regions or the entire gene locus), do not display a transcriptional adaptation response and exhibit stronger phenotypes than alleles displaying mutant mRNA decay (Figure 12 C, E, Figure 13 A-G). Engineering RNA-less alleles, however, can require the deletion of relatively large regions of the DNA which can influence the expression of other genes (e.g., neighboring genes). I, hereby, recommend starting with generation of RNA-less alleles, and if a phenotype is observed, one can confirm it by small indel mutations. Alternatively, if the protein of interest has clearly defined and conserved important domain, in-frame mutations targeting such domain can be a good alternative as they should not lead to degradation of the mutant transcript in most of the cases (Figure 7A).

7.4. Transcriptional adaptation involves upregulation of a significant proportion of genes exhibiting sequence similarity

Transcriptome analyses of the knockout mouse cell line models displaying mutant mRNA decay revealed that at least 50 to 60% of genes exhibiting sequence similarity with the mutant gene's mRNA get upregulated (Figure 16A). Interestingly as well, most of those genes were not upregulated in the corresponding RNA-less alleles (Table 33, 34). Injection of synthetic *hif1ab* transcripts containing sequences that exhibit sequence similarity to *epas1a* induced a transcriptional adaptation response but not transcripts containing sequences not exhibiting similarity (Figure 17C). Similarly, injection of mouse *Actb* uncapped transcripts into zebrafish embryos induced the upregulation of the highly-similar zebrafish ortholog *actb1* (Figure 17A). Taken together, these data suggested that transcriptional adaptation is induced in a sequence-similarity mediated mechanism.

One major question is why are not all genes exhibiting sequence similarity upregulated and what defines whether a given 'similar' gene gets upregulated or not. A number of explanations can be proposed. First, I observed that the proportion of upregulated genes varies depending on the similarity thresholds set, such as sequence alignment length and quality, and the corresponding E-values (Figure 15). Further systematic unbiased studies would be important to identify the optimal thresholds of the different criteria of similarity required for a given gene to be upregulated. I have, for example, observed that the localization of similarity influences the transcriptional adaptation response. Through injection of uncapped synthetic *hif1ab* transcripts containing sequences sharing sequence similarity with *epas1a* promoter, exons, introns, or 3'UTR, I observed that embryos injected with the transcripts exhibiting sequence similarity to *epas1a* exons or introns displayed the strongest upregulation levels of *epas1a* mRNA (Figure 17D), findings that were also consistent with the transcriptome data (Table 32, 33, 34).

Feedback loops may provide a simpler explanation, it is likely that some of the 'similar' genes are components of genetic networks that include other genes that are being upregulated in the knockout cells, and thereby influencing their expression in a transcriptional adaptation-independent manner. To avoid the effects of genes upregulated due to loss of protein function, future studies may benefit from using the uncapped RNA approach, or transgenes expressing a mutated gene, to study the criteria of similarity required for transcriptional adaptation-mediated upregulation.

Temporal studies of the response shall be also useful to identify if some 'similar' genes are fast responders compared to others, and might, therefore, influence the expression of the slower responding 'similar' genes. Another explanation might be due to the ability of the cell to fine-tune its transcriptome. I observed that in *Actg1* knockout cells, some 'similar' genes were upregulated on the pre-mRNA but not the mRNA level. It is likely, thereby, that some post-transcriptional mechanisms fine-tune the expression of the upregulated 'similar' genes to avoid any overt overexpression effects. A recent study, however, has reported what they referred to as 'genetic over-compensation', where zebrafish mutants of *marcksb*, a gene implicated in BMP signaling, displayed increased BMP signaling compared to wildtype owing to the upregulation of other Marcks genes as *marcksa*, *marcksl1a,* and *marcksl1b* (Ye et al., 2019).

Finally, I have shown that transcriptional adaptation involves modulation of the epigenetic landscape and proposed a potential role for antisense RNAs in the response. It is thereby likely that criteria such as the chromatin landscape and presence of regulatory RNAs at a given 'similar' gene locus may influence the response. It might be also compelling to investigate if transcriptional adaptation can also lead to downregulation of specific genes. Further understanding the molecular mechanisms underlying transcriptional adaptation would be essential to further investigate these hypotheses.

7.5. Molecular mechanisms underlying transcriptional adaptation

I proposed a model, whereby following mutant mRNA decay, mRNA degradation intermediates can translocate back to the nucleus to modulate gene expression in a sequence-specific manner (Figure 18). One of the models I proposed involves, RNA degradation intermediates guiding decay factors to the adapting genes' loci, which can then recruit the COMPASS complex to induce gene expression. These data agree with the growing number of studies that report evidence of ´cross-talk between mutant mRNA decay and gene expression. For example, some transcription factors, components of RNA polymerase II, and promoter elements have been associated with mRNA decay in the cytoplasm (Bregman et al., 2011; Goler-Baron et al., 2008; Lotan et al., 2005; Lotan et al., 2007). On the other hand, cytoplasmic mRNA decay is also known to influence gene expression (reviewed in (Hartenian and Glaunsinger, 2019). A previous study showed that wild-type but not catalytically inactive Xrn1 (the

5’ to 3’ exonuclease) can shuttle back to the nucleus, in an mRNA-decay dependent manner) and bind to chromatin to promote transcription initiation and elongation (Haimovich et al., 2013). The authors additionally showed that other decay factors such as Lsm1 and Dcp2 also bind to chromatin and influence transcription. In my study, I have also identified a requirement of decay factors for transcriptional adaptation (Figure 10 A-F). Interestingly, one of those decay factors was XRN1. Knockdown of XRN1 in *Rela* and *Actb* knockout cells ablate the transcriptional adaptation response (Figure 10 C, D) and injection of uncapped transcripts that had a 5’ sequence that renders them resistant to XRN1-mediated decay did not induce transcriptional adaptation (Figure 11D). It shall be interesting to investigate if XRN1 associates with chromatin at the adapting genes’ loci. Those previous reports have mainly studied the cross-talk between mRNA decay and gene expression as a way of buffering gene expression, i.e., reduce transcription when decay is minimal and increase transcription when decay is increased. And while they were mainly studied in yeast, a recent study has made the same observations in human cultured cells (Singh et al., 2019). Besides regulating the adapting genes’ expression levels, I have also shown that transcriptional adaptation modulates the mutant gene’s expression levels (Figure 4D, Figure 9C). Other studies have reported modulation of the mutated gene in response to a nonsense mutation. For example, UPF1 was reported to promote alternative splicing of the TCRbeta gene in response to nonsense mutations acquired during lymphocyte development (Mendell et al., 2002). Interestingly, such response was reported to be signaled upon translation of PTC-containing TCRbeta transcripts (Wang et al., 2002).

The other model I proposed involves mRNA degradation intermediates repressing antisense transcripts at the adapting genes’ loci and thereby allowing for increases sense RNA expression, and that model has been discussed thoroughly throughout this thesis. It however remains to be determined, whether the decrease in the adapting genes’ antisense transcript levels observed in the different models of transcriptional adaptation (Figure 20 A-D) is the cause for the increased sense mRNA expression or a consequence of such response.

The current model of transcriptional adaptation suggests that small RNA degradation intermediates translocate back to the nucleus to induce the upregulation of the adapting genes. A recent study in *C. elegans* has reported a requirement for factors involved in small RNA maturation and transport to the nucleus, such as Argonaute

proteins and DICER, in triggering transcriptional adaptation (Serobyan et al., 2020). Notably, Argonaute proteins are components of the RNAi machinery, and it shall be interesting to investigate if they are involved in inducing transcriptional adaptation in a manner similar to their role in RNAa (RNA activation) discussed in section **3.3.2.3.** of this thesis.

In agreement with my findings, a study from Zhejiang University, China, reported similar results (Ma et al., 2019). Studying two zebrafish mutant models, they proposed that mRNAs harboring a PTC can form a complex with the COMPASS complex and Upf3a to induce upregulation of paralogous genes. While similar to my findings, they implicate the COMPASS complex, and in particular WDR5, in the response, they proposed that response is induced by the full-length PTC-containing transcript rather than degradation products. Two paralogs of Upf3 exist in vertebrates, with Upf3b being more active in triggering NMD than Upf3a (Chan et al., 2009; Kim et al., 2001; Kunz et al., 2006; Lykke-Andersen et al., 2000), and a study suggesting that UPF3a is an inhibitor of NMD (Shum et al., 2016). The authors proposed a model whereby upon stalling of a ribosome at a PTC, if Upf3a joins the EJC, it will recruit WDR5 (and the COMPASS complex) and together with the mutant mRNA, it will form a complex that can increase H3K4me3 marks at the TSS of paralogous genes and activate their transcription. If Upf3b joins, however, it will induce NMD through interacting with Upf1 and Upf2. The slight discrepancy between my findings and their observations stem mainly from the fact that they do not observe loss of the transcriptional adaptation response upon inactivating Upf1 in their mutant zebrafish models. It remains to be understood whether that discrepancy is due to different approaches of inactivating Upf1 or whether the mechanisms underlying transcriptional adaptation can differ between different genes. For example, while I analyzed *Upf1* mutants, latest at 3 dpf, a stage where the *upf1* mutant fish still look healthy, they analyzed *Upf1* mutants at 4 dpf (*Upf1* mutants develop strong edema at 5 dpf). Moreover, they did not observe a transcriptional adaptation response upon injection of uncapped transcripts corresponding to the studied genes. It is important to note, however, that they analyzed the injected embryos at 1.5 dpf (approximately 36 hpf), and given how unstable these transcripts are, an earlier analysis of the injected embryos is required. Supporting this claim, transfection of *Actb* uncapped transcripts triggered a transcriptional adaptation response at 6 hours post-transfection but not after 24 hours (Figure 11C). However, in agreement with my

data, they observed that premature termination codons that do not trigger nonsense-mediated decay do not induce a transcriptional adaptation response. My observations on uncapped transcripts (Figure 11) and no-stop decay models (Figure 3) triggering transcriptional adaptation, and the response involving upregulation of genes with limited sequence similarity to the mutated gene's mRNA (Table 32, 33, 34), favor a model whereby mRNA degradation intermediates induce the response. However, I cannot rule out a potential contribution from the undegraded mutant transcripts in triggering transcriptional adaptation. Future studies, such as an unbiased genetic screen, will be required to better understand the mechanism and nature of the RNAs inducing the response.

7.6. Discrepancy between global and conditional knockout studies, and incomplete penetrance: a potential role for transcriptional adaptation?

The zebrafish in vivo models I used in this thesis were all germline mutants assessed for the transcriptional adaptation response at early developmental stages. One interesting question is whether transcriptional adaptation can be induced at later developmental stages or if it needs to be established during embryonic development or germline maturation. My data from *Fermt2* and *Actg1* knockout cells (where the mutations were not germline mutations) and the observations that *egfl7* CRISPants (embryos that are injected with a Cas9 and a gRNA targeting *egfl7* that are mosaics for the mutation) display a transcriptional adaptation response (Rossi et al., 2015), suggest that the mutation does not need to go through the germline to induce the response, however, a number of studies reported stronger phenotypes in conditional knockout mice compared to global knockouts. For instance, while *Pkm2* global mutants are viable and fertile (Dayton et al., 2016), conditional knockout of *Pkm2* in mouse embryonic fibroblasts displays cell-cycle arrest due to limited nucleotide synthesis (Lunt et al., 2015). Moreover, hepatocyte-conditional knockout *Sirt1* mice display fatty liver, while the global mutants display no liver defects (Wang et al., 2010). Furthermore, *Fgfr3* specific-knockout in chondrocytes leads to increased and stronger chondrona-like lesions compared to global knockouts (Zhou et al., 2015b). Besides, while *Cd44* keratinocyte-conditional knockout mice exhibit delayed wound healing, decreased stiffness of the epidermis and decreased proliferation of keratinocytes following exposure to 12-O-tetradecanoylphorbol-13-acetate (Shatirishvili et al., 2016), global knockouts display mild defects (Protin et al., 1999; Schmits et al., 1997). Notably, conditional knockout of the tumor suppressor gene

Rb1 in mouse embryonic fibroblasts allowed quiescent cells to re-enter the cell cycle, while quiescent cells isolated from the global *Rb1* knockout mice were not able to re-enter cell cycle due to the compensatory upregulation of the tumor suppressor protein RBL1 (Sage et al., 2003). Cell non-autonomous and acute vs chronic loss of the protein function (Cerikan et al., 2016) effects can explain many of these findings, however, it would be interesting to further investigate if different transcriptional adaptation responses are observed between the conditional and global knockouts as it may have implications on our understanding of the transcriptional adaptation machinery.

On a different note, multiple factors have been proposed to explain the observations of incomplete penetrance of mutant phenotypes such as genetic background, environmental factors (Nadeau, 2001; Raj and van Oudenaarden, 2008). A recent mouse study has however reported that incomplete penetrance is also observed in animals of similar genetic backgrounds (Dickinson et al., 2016). Interestingly, a *C. elegans* study attributed the incomplete penetrance of the *skn-1* mutants gut phenotype (Bowerman et al., 1992) to the variability of the compensating gene *end-1* expression levels (Raj et al., 2010). It shall be interesting to assess whether incomplete penetrance can be explained by transcriptional adaptation (e.g., different mutants exhibiting different decay levels of the mutant gene, and thereby leading to differences in transcriptional adaptation-mediated upregulation of compensating genes). Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy and Stainier, 2017).

7.7. Implications on human genetics

In a number of genetic diseases, mutations that are less likely to cause mRNA decay, such as missense and in-frame mutations, are more represented in affected individuals than mutations that are more likely to induce mRNA decay such as nonsense and out-of-frame indel mutations (Chuzhanova et al., 2003; Eisensmith and Woo, 1992; Ferec and Cutting, 2012; Genschel and Schmidt, 2000; Myerowitz, 1997; Zhou et al., 2014). Notably, a study on Marfan syndrome patients reported that the mildest form of the syndrome was observed in individuals displaying very low levels of the *FBN1* mutant gene (due to an out-of-frame indel mutation that introduced a PTC in the mutant transcript) when compared to individuals displaying no decrease in mutant mRNA levels due to a missense mutation (Dietz et al., 1993).

Similar results were observed in individuals with heterozygous nonsense mutations in the *HBB* gene; individuals who displayed decay of the mutant *HBB* transcripts were asymptomatic while individuals displaying no decay developed beta thalassemia-intermedia (Hall and Thein, 1994). The current understanding is that missense mutations are more represented as disease-causing mutations in some diseases as they can be translated into toxic proteins that can act as dominant-negative or constitutively-active proteins, thereby leading to the phenotype. I, however, propose that mutations predicted to cause mutant mRNA decay, such as nonsense mutations, might be less represented in affected individuals as they can induce a transcriptional adaptation response that can ameliorate the severity of the phenotype. Future analyses of the transcriptomes of such individuals can help test this hypothesis. Notably, for example, *ACTA1* nonsense mutations in muscle biopsies of nemaline myopathy patients display upregulation of the paralog ACTC1, an upregulation that the authors suggested to be a potential determinant of the disease severity (Nowak et al., 2007).

Recent whole-genome sequencing studies reported homozygous loss-of-function mutations (LoF) in healthy individuals in several genes, including *EGFL7* and *RELA* that I studied in this thesis and other disease-associated genes (Chen et al., 2016b; Lek et al., 2016; MacArthur et al., 2012; Sulem et al., 2015; Wall et al., 2019). Studies have estimated that each individual may harbor around 100 heterozygous LoF and 20 homozygous LoF mutations in protein-coding genes (Durbin et al., 2010; MacArthur et al., 2012; Ng et al., 2008). The Genome100KAsia consortium identified protein-truncating variants (PTVs) in 8,766 protein-coding genes, most being heterozygous, but interestingly they identified 856 homozygous PTVs (Wall et al., 2019). Analysis of more than 500,00 genomes has also identified 13 individuals harboring disease-causing mutations in 8 different genes (Chen et al., 2016b). Transcriptome analysis shall be helpful to understand if such individuals display mutant mRNA decay and whether it is associated with upregulation of a compensating gene. These transcriptome analyses can help make us understand why some mutations are deleterious, and cause disease, while others not. They can also help identify new modifier genes that may be modulated for therapeutic purposes. Further investigating the molecular mechanisms underlying transcriptional adaptation may lead to the development of more effective therapies, ones that enhance an individual's robustness to a mutation rather than correct the mutation.

For example, I have shown that injection of uncapped RNAs can induce a transcriptional adaptation response (Figure 11A) and (Ma et al., 2019) proposed that introduction of transgenes harboring a PTC can similarly introduce a transcriptional adaptation response. Future studies will be required to assess the therapeutic potential of such approaches, in addition to potentially introducing a PTC in missense alleles associated with genetic diseases. Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

8. Conclusion

In conclusion, I have identified a novel role for mutant mRNA degradation in triggering transcriptional adaptation. I achieved my study aims as follows:

Aim 1: Aim 1: Identify transcriptional adaptation's molecular trigger.
Analyzing different mutant alleles for *hbegfa* and *vcla* in zebrafish, I observed a correlation between levels of mutant mRNA degradation and upregulation of related genes. Through pharmacological and genetic approaches, I was able to confirm the requirement of the mRNA surveillance machinery in triggering transcriptional adaptation.

Aim 2: Provide guidelines for the generation of mutant alleles with minimal transcriptional adaptation-derived compensation.
Mutant alleles engineered not to transcribe the mutated gene fail to show a transcriptional adaptation response, further suggesting the necessity of mutant mRNA degradation for triggering the response. Such mutants can be generated through deletion of the promoter region or the full genomic locus of the genes of interest (referred to in the thesis as RNA-less alleles). Interestingly, I was able to show that such RNA-less alleles display stronger phenotypes than alleles displaying mutant mRNA decay. Inframe mutations targeting well-characterized and conserved domains can be also a good alternative, as they are less likely to lead to mutant mRNA decay.

Aim 3: Investigate the molecular mechanisms underlying transcriptional adaptation.
Transcriptome analysis of knockout mouse cell line models revealed upregulation of a substantial proportion of genes sharing sequence similarity with the decaying mutant mRNA, suggesting a sequence-dependent mechanism. Mechanistically, I proposed that mutant mRNA decay intermediates may act as guides to bring decay factors, or other RNA binding proteins, to loci of adapting genes in order to induce their expression through modifying the chromatin environment. Alternatively, I also proposed that mRNA decay intermediates may target antisense RNAs at the adapting genes' loci to induce transcriptional adaptation.

9. Summary (Zusammenfassung)

9.1. English summary

Investigating the molecular mechanisms underlying genetic compensation

Introduction

The development of a single-cell zygote to a complex organism with different cell types is a fascinating process that has been optimized over millions of years of evolution. To ensure similar developmental outcomes despite small changes in the environmental surroundings or genetic makeup, multiple buffering systems evolved; giving rise to the term: genetic robustness. The increasing number of generated mutant animals showing no obvious phenotype (Bouche and Bouchez, 2001; El-Brolosy and Stainier, 2017; Giaever et al., 2002; Kok et al., 2015) and the recent sequencing studies of healthy individuals with homozygous loss-of-function mutations (Chen et al., 2016b; Lek et al., 2016; Sulem et al., 2015) have revived interest in the concept of genetic robustness. Multiple mechanisms were proposed to explain the phenomenon such as functional genetic redundancies (Tautz, 1992), rewiring of cellular networks including transcriptional networks and protein feedback loops (Barabasi and Oltvai, 2004; Davidson and Levin, 2005), and for rapidly proliferating organisms like yeast, accumulation adaptive mutations (Teng et al., 2013) (reviewed in(El-Brolosy and Stainier, 2017)). The recent advances in reverse genetic tools, such as CRISPR, have not only reinforced the findings that several generated mutants do not display an obvious phenotype but has also revealed big discrepancies with knockdown models (such as antisense-treated animals (or cells) e.g., morpholinos). Scientists have surprisingly observed, counter-intuitively, in several model organisms that knockdown of specific genes leads to stronger phenotypes than knockout of the same gene ((Daude et al., 2012; De Souza et al., 2006; Gao et al., 2015; Hall et al., 2013; Kok et al., 2015) (reviewed in(El-Brolosy and Stainier, 2017)). While some studies have attributed such discrepancy to off-target effects of the antisense reagents, Rossi and colleagues proposed genetic compensation by upregulation of related genes to be underlying some cases of discrepancy between knockdown and knockout models (Rossi et al., 2015). Studying *egfl7*, an endothelial extracellular matrix gene, they observed strong

vasculature defects upon knockdown of the gene using antisense morpholinos (morphants), while mutants displayed no obvious phenotype. Transcriptomic and proteomic analyses revealed upregulation of another family of extracellular matrix genes, the *emilins*, in mutant animals but not morphants. Furthermore, they reported upregulation of *vegfab* in *vegfaa* mutants but not morphants. They also suggested that these transcriptional adaptation responses are induced upstream of the loss of protein function as overexpression of a dominant-negative version of Vegfaa did not lead to *vegfab* upregulation in the injected embryos.

Several other studies have reported genetic compensation in response to deleterious mutations ((Sztal et al., 2018; Tondeleir et al., 2012; Zhu et al., 2017), reviewed in(El-Brolosy and Stainier, 2017)). The underlying molecular mechanism, however, remained unclear, a question that I aimed to identify during my Ph.D. studies.

Results

Transcriptional adaptation is independent of loss of protein function and involves enhanced transcription

I started by analyzing different zebrafish and mouse cell line mutants harboring a premature termination codon (PTC) or have their last exon deleted (Figure 3). I studied six different zebrafish genes: *hbegfa*, *vcla*, *hif1ab*, *vegfaa*, *egfl7* and *alcama*, besides four mouse genes: *Fermt2*, *Rela*, *Actg1*, and *Actb* and observed increased mRNA levels of a paralog or a family member of such genes (hereafter referred to as adapting genes), namely *hbegfb*, *vclb*, *epas1*, *vegfab*, *emilin3a* and *alcamb* in the homozygous zebrafish mutants, and *Fermt1*, *Rel*, *Actg2*, and *Actg1* in the mouse cell line mutants (Figure 4A and Figure 5A). Interestingly, heterozygous alleles also displayed upregulation of the adapting genes and upregulation of the wild-type allele of the mutated gene (Figure 4C, D and Figure 5E). Rescue of the mutant zebrafish and mouse cell lines by injection or transfection of the wild-type version of the mutated gene did not dampen the transcriptional adaptation response, further confirming that transcriptional adaptation is triggered upstream of the loss of protein function (Figure 4B and Figure 5B, C). To identify if the upregulation of the adapting genes transcript levels is due to enhanced transcription or increased mRNA stability, I analyzed pre-mRNA levels of the adapting genes in some of the mutants and observed that they were also upregulated (Figure 6A, B). Moreover, *Fermt2*

knockout cells displayed increased chromatin opening at the transcription start site of the adapting gene, *Fermt1*, as assessed by ATAC-seq (Figure 6C).

Mutant mRNA degradation triggers transcriptional adaptation

Since loss of protein function is not the underlying trigger for transcriptional adaptation, I investigated two other potential triggers; the DNA lesion itself or the mutant mRNA molecules. I reasoned that if the DNA lesion is the trigger for the response, then any kind of mutation, including those that do not affect the transcript and protein integrity, should induce the response. However, it was not the case as in-frame mutations and mutations in un-translated regions did not display a transcriptional adaptation response (Figure 7 A-C). PTC-containing mRNAs and mutations that lead to a transcript lacking a stop codon (such as those studied in this thesis where the last exon was deleted) are subjected to degradation by the mRNA surveillance machinery through non-sense mediated decay and no-stop decay pathways, respectively, or potentially other pathways (Isken and Maquat, 2007; Wolin and Maquat, 2019). While analyzing two different PTC-containing mutant alleles for *hbegfa* and *vcla*, I observed a correlation between levels of mutant mRNA decay and the upregulation of the adapting gene. *hbegfa* or *vcla* mutant alleles that displayed limited, or no, mutant mRNA decay didn't display transcriptional adaptation (Figure 8A, B). Genetic inactivation of *upf1* (a key component of the non-sense mediated decay machinery) in zebrafish mutant alleles displaying mutant mRNA decay, led to stabilization of the mutant transcript and loss of the transcriptional adaptation response (Figure 10A, B). Moreover, knockdown of other components of the mRNA surveillance machinery such as SMG6, ERF1, and XRN1 in *Rela* and *Actb* knockout mouse cell lines also dampened the transcriptional adaptation response (Figure 10 C, D). Furthermore, injection or transfection of uncapped RNAs, that are rapidly degraded inside the cell, into wild-type embryos or cells induced transcriptional adaptation (Figure 11 A-C), further confirming that mRNA degradation induces the response.

Mutant alleles that fail to transcribe the mutated gene do not induce transcriptional adaptation and display stronger phenotypes than alleles displaying mutant mRNA decay

I reasoned that if mutant mRNA degradation is important to induce transcriptional adaptation, then mutant alleles that do not transcribe the mutated gene should also

not be able to induce the response. To this end, I generated such mutants through deletion of either the promoter region or the entire genomic locus (hereafter referred to as RNA-less alleles). Indeed, zebrafish and mouse RNA-less alleles failed to induce upregulation of adapting genes (Figure 12 C, E). More interestingly, such mutant alleles displayed stronger phenotypes than those observed in alleles displaying mutant mRNA decay (Figure 13 A-G). For example, the generated *egfl7* RNA-less mutant displayed vascular defects akin to those observed in the morphants, a phenotype that was lacking in the mutant allele displaying mutant mRNA decay (Rossi et al., 2015). Thereby, generation of RNA-less alleles can uncover previously masked phenotypes, enabling better understanding of gene function.

Transcriptional adaptation is induced in a sequence dependent manner

To further understand the molecular mechanisms underlying transcriptional adaptation, I performed transcriptome analysis of *Fermt2*, *Actg1,* and *Actb* knockout mouse cells. I observed that at least 50 to 60% of genes sharing sequence similarity with the mutated gene's mRNA (based on multiple similarity thresholds, as described in section 5.2.42 in this thesis; hereafter referred to as 'similar genes') were upregulated in the knockout alleles, compared to a maximum of 21% of genes not sharing sequence similarity (Figure 16A). More interestingly, many of those upregulated similar genes were upregulated in alleles displaying mutant mRNA decay but not in RNA-less alleles (Table 33, 34). To further confirm the importance of sequence similarity for transcriptional adaptation, I injected uncapped transcripts composed solely of *hif1ab* mRNA sequences either sharing or not sharing, sequence similarity with *epas1a* genomic locus (Figure 17B). Only the version containing sequences sharing sequence similarity with *epas1a* was able to induce transcriptional adaptation (Figure 17C). Such data suggested a model, whereby following mutant mRNA degradation, decay intermediates would induce the response in a sequence-dependent fashion (Figure 18).

Transcriptional adaptation involves epigenetic remodeling

Several studies in the past decade reported that mRNA decay and gene expression are interconnected processes (Elkon et al., 2010; Hao and Baltimore, 2009; Sun et al., 2012). Following mRNA decay, decay factors were reported to translocate back to the nucleus and interact with epigenetic remodelers to induce gene expression

(Collins et al., 2007; Haimovich et al., 2013). I thereby performed a small siRNA screen on *Rela* knockout cells to identify chromatin remodelers or epigenetic modifiers involved in transcriptional adaptation. The strongest effect was observed upon knockdown of WDR5, which resulted in loss of *Rel* upregulation in *Rela* knockout cells (Figure 19A). WDR5 is a component of the COMPASS complex that deposits the permissive H3K4me3 histone mark. Chromatin immunoprecipitation experiments revealed enrichment of both WDR5 and H3K4me3 at the transcription start sites of *Fermt1*, *Rel,* and *Actg2* in *Fermt2*, *Rela,* and *Actg1* knockout cells respectively (Figure 19B, C). Interestingly, knockdown of XRN1 or UPF1/EXOSC4 in *Rela* knockout cells led to depletion of H3K4me3 marks at *Rel* transcription start site (Figure 19E), suggesting a model whereby following mRNA degradation, decay factors translocate back to the nucleus along with the decay intermediates that can guide them to the adapting genes' loci to induce gene expression through recruiting epigenetic remodelers (Figure 21).

A potential role for antisense RNAs in transcriptional adaptation

Antisense RNAs can act as negative regulators of gene expression (Faghihi and Wahlestedt, 2009; Modarresi et al., 2012). I observed downregulation of antisense transcripts in the *vclb* and *hbegfb* loci in *vcla* and *hbegfa* mutant alleles displaying mutant mRNA decay (Figure 20 D). Moreover, transfection of uncapped *BDNF* transcripts in HEK cells led to decreased expression levels of *BDNF* antisense transcript and increased expression of the sense one (Figure 20B). Altogether, these data indicate that targeting antisense RNAs can be another way through which mRNA decay intermediates induce transcriptional adaptation (Figure 21).

Conclusion

In conclusion, I have identified a novel role for the mRNA surveillance machinery in maintaining genetic robustness. I proposed that following mutant mRNA degradation, decay intermediates can translocate back to the nucleus to induce the response in a sequence-dependent manner. I have shown that the response involves epigenetic remodeling at the adapting gene's loci and may involve targeting antisense transcripts. Future analyses shall help identify further mechanistic details behind transcriptional adaptation. The findings shall help design mutant alleles with minimal transcriptional adaptation-derived compensation, thereby facilitating studying gene function. I have shown that mutant alleles that fail to transcribe the mutated

gene do not display a transcriptional adaptation response, and can uncover previously masked phenotypes. The findings have also huge implications on our understanding of disease-causing mutations. Recent whole-genome sequencing studies have identified individuals with homozygous loss of function mutations in several genes, including genes such as RELA and EGFL7 studied in this thesis (Lek et al., 2016; Sulem et al., 2015). It shall be interesting to investigate whether such individuals display mutant mRNA decay and if it is associated with a transcriptional adaptation response that can protect them. Further investigating the molecular mechanisms underlying transcriptional adaptation can thereby hold a great translational potential.

9.2. Zusammenfassung (German summary)

Untersuchung der molekularen Mechanismen der genetischen Kompensation

Einführung

Die Entwicklung einer einzelligen Zygote zu einem komplexen Organismus mit unterschiedlichen Zelltypen ist ein faszinierender Prozess, der über Millionen von Jahren der Evolution optimiert wurde. Um ähnliche Entwicklungsergebnisse trotz geringfügiger Änderungen der Umweltbedingungen oder der genetischen Verfassung zu gewährleisten, entwickelten sich mehrere Puffersysteme, die unter dem Begriff genetische Robustheit zusammengefasst werden. Die zunehmende Anzahl erzeugter tierischer Mutanten, die keinen offensichtlichen Phänotyp zeigten (Bouche und Bouchez, 2001; El-Brolosy und Stainier, 2017; Giaever et al., 2002; Kok et al., 2015) und die jüngsten Sequenzierungsstudien von Personen mit homozygoten Funktionsverlustmutationen ohne klinische Symptome (Chen et al., 2016; Lek et al., 2016; Sulem et al., 2015) haben das Interesse am Konzept der genetischen Robustheit wiederbelebt. Es wurden mehrere Mechanismen vorgeschlagen, um das Phänomen zu erklären: funktionelle genetische Redundanzen (Tautz, 1992), Neuverdrahtung von zellulären Netzwerken, einschließlich Transkriptionsnetzwerken und Proteinrückkopplungsschleifen (Barabasi und Oltvai, 2004; Davidson und Levin, 2005) und für sich schnell vermehrende Organismen wie Hefe, die Akkumulation adaptiver Mutationen (Teng et al., 2013) (Übersicht in (El-Brolosy und Stainier, 2017)). Die jüngsten Fortschritte bei reversen genetischen Instrumenten wie CRISPR haben nicht nur die Erkenntnisse bestärkt, dass viele erzeugte Mutanten keinen offensichtlichen Phänotyp aufweisen, sondern auch große Diskrepanzen mit Knockdown-Modellen (wie mit Antisense behandelten Tieren (oder Zellen) aufgedeckt, z.B. Morpholinos). Wissenschaftler haben überraschenderweise in mehreren Modellorganismen kontraintuitiv beobachtet, dass der Knockdown bestimmter Gene zu stärkeren Phänotypen führt als das Ausschalten derselben ((Daude et al., 2012; De Souza et al., 2006; Gao et al., 2015; Hall et al., 2013; Kok et al., 2015) (besprochen in (El-Brolosy und Stainier, 2017). Während einige Studien eine solche Diskrepanz den Nebenwirkungen der Antisense-Reagenzien zuschrieben, haben Rossi und Kollegen eine genetische Kompensation durch Hochregulierung verwandter Gene vorgeschlagen, um einige Fälle von Diskrepanzen

zwischen Knockdown- und Knockout-Modellen zu erklären (Rossi et al., 2015). Sie studierten *egfl7*, ein endotheliales extrazelluläres Matrixgen und beobachteten starke Gefäßdefekte beim Knockdown des Gens unter Verwendung von Antisense Morpholinos (Morphanten), während Mutanten keinen offensichtlichen Phänotyp aufwiesen. Bei Transkriptom- und Proteomanalysen stellten sie eine Hochregulation einer anderen Familie extrazellulärer Matrixgene, der Emiline in mutierten Tieren, jedoch nicht bei Morphanten fest. Sie berichteten über eine Hochregulierung von *vegfab* in *vegfaa*-Mutanten, jedoch nicht in Morphanten. Sie schlugen zudem vor, dass diese Transkriptionsanpassung vor dem Verlust der Proteinfunktion induziert werden, da eine Überexpression einer dominanten negativen Version von Vegfaa nicht zu einer *vegfab*-Hochregulation in den injizierten Embryonen führte.

In mehreren anderen Studien wurde eine genetische Kompensation als Reaktion auf Mutationen berichtet, die zum Funktionsverlust des Gens führen ((Sztal et al., 2018; Tondeleir et al., 2012; Zhu et al., 2017), besprochen in (El-Brolosy und Stainier, 2017)). Der zugrunde liegende molekulare Mechanismus blieb jedoch unklar. Diesen Mechanismus aufzuklären war das Ziel meiner Promotion.

Ergebnisse

Die Transkriptionsanpassung ist unabhängig vom Verlust der Proteinfunktion und bringt eine gesteigerte Transkriptionrate mit sich

Ich begann mit der Analyse verschiedener Zebrafisch- und Mauszelllinien-Mutanten, die ein vorzeitiges Terminationscodon (PTC) enthalten oder deren letztes Exon gelöscht wurde (Abbildung 3). Ich habe sechs verschiedene Zebrafischgene untersucht: *hbegfa*, *vcla*, *hif1ab*, *vegfaa*, *egfl7* und *alcama* sowie vier Mausgene: *Fermt2*, *Rela*, *Actg1* und *Actb*. Ich beobachtete erhöhte mRNA-Spiegel eines Paralogs oder eines Familienmitglieds solcher Gene (nachstehend bezeichnet) als adaptierende Gene), nämlich *hbegfb*, *vclb*, *epas1*, *vegfab*, *emilin3a* und *alcamb* in den homozygoten Zebrafischmutanten und *Fermt1*, *Rel*, *Actg2* und *Actg1* in den Mauszelllinienmutanten (Abbildung 4A und Abbildung 5A). Interessanterweise zeigten heterozygote Allele auch eine Hochregulation der adaptierenden Gene und eine Hochregulation des Wildtyp-Allels des mutierten Gens (Abbildung 4C, D und Abbildung 5E). Die Rettung der mutierten Zebrafisch- und Mauszelllinien durch Injektion oder Transfektion der Wildtyp-Version des mutierten Gens dämpfte die Antwort auf die Transkriptionsanpassung nicht und bestätigte weiter, dass die

Transkriptionsanpassung vor dem Verlust der Proteinfunktion ausgelöst wird (Abbildung 4B und Abbildung 5B, C). Um festzustellen, ob die Hochregulation der Transkriptionsniveaus der adaptierenden Gene auf eine erhöhte Transkription oder erhöhte mRNA-Stabilität zurückzuführen ist, analysierte ich die Prä-mRNA-Niveaus der adaptierenden Gene in einigen Mutanten und stellte fest, dass diese ebenfalls hochreguliert waren (Abbildung 6A, B). Darüber hinaus zeigten ATAC-seq Resultate für *Fermt2*-Knockout-Zellen eine erhöhte Chromatinöffnung an der Transkriptionsstartstelle des adaptierenden Gens *Fermt1* (Abbildung 6C).

Der Abbau mutierter mRNA löst eine Transkriptionsanpassung aus.

Da der Verlust der Proteinfunktion nicht der zugrunde liegende Auslöser für die Transkriptionsanpassung ist, habe ich zwei weitere mögliche Auslöser untersucht: die DNA-Läsion selbst und die mutierten mRNA-Moleküle. Wenn die DNA-Läsion der Auslöser für die Reaktion ist, sollte jede Art von Mutation, einschließlich jener, die das Transkript und die Proteinintegrität nicht beeinflussen, die Reaktion auslösen. Dies war jedoch nicht der Fall, da In-Frame-Mutationen und Mutationen in nicht translatierten Regionen keine Antwort auf die Transkriptionsanpassung zeigten (Abbildung 7A-C). PTC-haltige mRNAs und Mutationen, die zu einem Transkript führen, dem ein Stoppcodon fehlt (wie jene in dieser Arbeit untersuchten, in denen das letzte Exon deletiert wurde), werden durch den mRNA-Überwachungsapparat durch Nonsense-vermittelten Zerfall und No-Stop-Zerfall bzw. potenziell andere Wege abgebaut (Isken und Maquat, 2007; Wolin und Maquat, 2019). Während ich zwei verschiedene PTC-haltige mutierte Allele auf *hbegfa* und *vcla* analysierte, beobachtete ich eine Korrelation zwischen dem Ausmaß des mutierten mRNA-Zerfalls und der Hochregulation des adaptierenden Gens. Mutierte *hbegfa* oder *vcla* Allele, die einen begrenzten oder keinen mutationsinduzierten mRNA-Zerfall verursachten, zeigten keine Transkriptionsanpassung (Abbildung 8A, B). Die genetische Inaktivierung von *upf1* (einer Schlüsselkomponente der Nonsense-vermittelten Zerfallsmaschinerie) in mutierten Zebrafisch-Allelen, die mutierten mRNA-Zerfall zeigten, führte zur Stabilisierung des mutierten Transkripts und zum Verlust der Transkriptionsadaptionsantwort (Abbildung 10A, B). Darüber hinaus wurde die Transkriptionsanpassungsreaktion durch das Herunterfahren anderer Komponenten der mRNA-Überwachungsmaschinerie wie SMG6, ERF1 und XRN1 in *Rela*- und *Actb*-Knockout-Mauszelllinien ebenfalls gedämpft (Abbildung 10C-F).

Darüber hinaus induzierte die Injektion oder Transfektion von RNAs ohne 5‘-Cap Struktur, die innerhalb der Zelle schnell abgebaut werden, in Wildtyp-Embryonen oder -Zellen eine Transkriptionsanpassung, was weiter bestätigt, dass der mRNA-Abbau die Reaktion induziert (Abbildung 11 A-C).

Mutierte Allele, die das mutierte Gen nicht transkribieren, induzieren keine Transkriptionsanpassung und zeigen stärkere Phänotypen als Allele, die einen mutierten mRNA-Zerfall zeigen

Wenn der Abbau mutierter mRNA wichtig ist, um die Transkriptionsanpassung zu induzieren, sollten mutierte Allele, die das mutierte Gen nicht transkribieren, auch nicht in der Lage sein, die Reaktion zu induzieren. Zu diesem Zweck erzeugte ich solche Mutanten entweder durch Deletion der Promotorregion oder des gesamten genomischen Locus (nachstehend als RNA-lose Allele bezeichnet). Tatsächlich induzierten RNA-freie Allele in Zebrafisch und Maus keine Hochregulation der adaptierenden Gene (Abbildung 12C, E). Interessanter ist, dass solche mutierten Allele stärkere Phänotypen aufwiesen als solche, die bei Allelen mit mutiertem mRNA-Zerfall beobachtet wurden (Abbildung 13 A-G). Beispielsweise wies die erzeugte egfl7-RNA-freie Mutante Gefäßdefekte auf, die denen der Morphanten ähnelten, ein Phänotyp, der im mutierten Allel fehlte, das einen Nonsense-induzierten mRNA-Zerfall aufwies (Rossi et al., 2015). Dadurch kann die Erzeugung von RNA-freien Allelen zuvor maskierte Phänotypen aufdecken und ein besseres Verständnis der Genfunktion ermöglichen.

Die Transkriptionsanpassung wird sequenzabhängig induziert

Um die molekularen Mechanismen zu verstehen, die der Transkriptionsanpassung zugrunde liegen, führte ich eine Transkriptomanalyse von *Fermt2*-, *Actg1*- und *Actb*-Knockout-Mauszellen durch. Ich beobachtete, dass mindestens 50 bis 60% der Gene, die die Sequenzähnlichkeit mit der mRNA des mutierten Gens teilen (basierend auf mehreren Ähnlichkeitsschwellenwerten, wie in Abschnitt 5.2.42 dieser Arbeit beschrieben; im Folgenden als "ähnliche Gene" bezeichnet), hochreguliert wurden, verglichen mit maximal 21% der Gene, die keine Sequenzähnlichkeit aufweisen (Abbildung 16A). Interessanterweise waren viele dieser hochregulierten ähnlichen Gene in Allelen hochreguliert, die einen Nonsense-induzierten mRNA-Zerfall zeigten, jedoch nicht in RNA-freien Allelen (Tabelle 33, 34). Um die Bedeutung der Sequenzähnlichkeit für die Transkriptionsanpassung weiter zu

bestätigen, injizierte ich Transkripte ohne 5‘-Cap Struktur, die ausschließlich aus *hif1ab*-mRNA-Sequenzen bestanden, die entweder die Sequenzähnlichkeit mit dem *epas1a*-Genomlocus teilen oder nicht teilen (Abbildung 17B). Nur die Version, die Sequenzen enthält, die Sequenzähnlichkeit mit *epas1a* teilen, konnte eine Transkriptionsanpassung induzieren (Abbildung 17C). Solche Daten deuteten auf ein Modell hin, bei dem Zerfallsintermediate nach Abbau mutierter mRNA die Reaktion in sequenzabhängiger Weise induzieren (Abbildung 18).

Die Transkriptionsanpassung beinhaltet einen epigenetischen Umbau

In mehreren Studien des letzten Jahrzehnts wurde berichtet, dass der mRNA-Zerfall und die Genexpression miteinander verbunden sind (Elkon et al., 2010; Hao and Baltimore, 2009; Sun et al., 2012). Es wurde berichtet, dass nach dem mRNA-Zerfall Zerfallsfaktoren in den Zellkern zurücktranslozieren und mit epigenetischen Umbaufaktoren interagieren, um die Genexpression zu induzieren (Collins et al., 2007; Haimovich et al., 2013). Ich führte dabei ein kleines siRNA-Screening an *Rela*-Knockout-Zellen durch, um Chromatin-Remodeler oder epigenetische Modifikatoren zu identifizieren, die an der Transkriptionsanpassung beteiligt sind. Der stärkste Effekt wurde beim Herunterfahren von WDR5 beobachtet, was zum Verlust der *Rel*-Hochregulation in *Rela*-Knockout-Zellen führte (Abbildung 19A). WDR5 ist eine Komponente des COMPASS-Komplexes, der die zulässige Histonmarke H3K4me3 hinterlegt. Chromatin-Immunpräzipitationsexperimente ergaben eine Anreicherung von WDR5 und H3K4me3 an den Transkriptionsstartstellen von *Fermt1*, *Rel* und *Actg2* in *Fermt2*-, *Rela*- und *Actg1*-Knockout-Zellen (Abbildung 19B, C). Interessanterweise führte der Abbau von XRN1 oder UPF1 / EXOSC4 in *Rela*-Knockout-Zellen zu einem Abbau der H3K4me3-Markierungen an der *Rel*-Transkriptionsstartstelle (Abbildung 19E), was auf ein Modell hindeutet, bei dem nach dem Abbau der mRNA die Zerfallsfaktoren zusammen mit den Zerfallszwischenprodukten in den Zellkern zurücktranslozieren können und sie zu den Loci der adaptierenden Gene führen, um die Genexpression durch Rekrutierung von epigenetischen Umbaufaktoren zu induzieren (Abbildung 21).

Eine mögliche Rolle für Antisense-RNAs bei der Transkriptionsanpassung

Antisense-RNAs können als negative Regulatoren der Genexpression wirken (Faghihi und Wahlestedt, 2009; Modarresi et al., 2012). Ich beobachtete eine Herunterregulation von Antisense-Transkripten in den *vclb*- und *hbegfb*-Loci in *vcla*-

und *hbegfa*-mutierten Allelen, die einen mutierten mRNA-Zerfall zeigten (Abbildung 20 C, D). Darüber hinaus führte die Transfektion von nicht abgedeckten BDNF-Transkripten in HEK-Zellen zu verringerten Expressionsniveaus des BDNF-Antisense-Transkripts und zu einer erhöhten Expression des Sense-Transkripts (Abbildung 20B). Insgesamt deuten diese Daten darauf hin, dass das Targeting von Antisense-RNAs ein weiterer Weg sein kann, durch den mRNA-Zerfallsintermediate die Transkriptionsanpassung induzieren (Abbildung 21).

Fazit

Zusammenfassend habe ich eine neue Rolle für die mRNA-Überwachungsmaschinerie bei der Aufrechterhaltung der genetischen Robustheit identifiziert. Ich schlug vor, dass nach dem Abbau mutierter mRNA Zerfallsintermediate in den Kern zurücktransloziert werden können, um die Reaktion in sequenzabhängiger Weise zu induzieren. Ich habe gezeigt, dass die Reaktion eine epigenetische Umstrukturierung an den Loci des adaptierenden Gens und möglicherweise das Targeting von Antisense-Transkripten umfasst. Zukünftige Analysen sollen helfen, weitere mechanistische Details der Transkriptionsanpassung zu identifizieren. Die Ergebnisse sollen dazu beitragen, mutierte Allele mit minimaler Kompensation aufgrund der Transkriptionsanpassung zu entwerfen und so die Untersuchung der Genfunktion zu erleichtern. Ich habe gezeigt, dass mutierte Allele, die das mutierte Gen nicht transkribieren, keine Transkriptionsanpassungsreaktion zeigen und zuvor maskierte Phänotypen aufdecken können. Die Ergebnisse haben auch enorme Auswirkungen auf unser Verständnis von krankheitsverursachenden Mutationen. Jüngste Sequenzierungsstudien des gesamten Genoms haben Personen mit homozygoten, zum Funktionsverlust führenden Mutationen in mehreren Genen identifiziert, einschließlich Genen wie RELA und EGFL7, die in dieser Arbeit untersucht wurden (Lek et al., 2016; Sulem et al., 2015). Es wird interessant sein zu untersuchen, ob solche Individuen einen Zerfall mutierter mRNA aufweisen und ob er mit einer Transkriptionsanpassungsreaktion verbunden ist, die sie schützen kann. Eine weitere Untersuchung der molekularen Mechanismen, die der Transkriptionsanpassung zugrunde liegen, kann dabei ein großes Translationspotential bergen.

10. References

Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E., and Helin, K. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature *449*, 731-734.

Akimitsu, N. (2008). Messenger RNA surveillance systems monitoring proper translation termination. Journal of biochemistry *143*, 1-8.

Alcamo, E., Mizgerd, J.P., Horwitz, B.H., Bronson, R., Beg, A.A., Scott, M., Doerschuk, C.M., Hynes, R.O., and Baltimore, D. (2001). Targeted mutation of TNF receptor I rescues the RelA-deficient mouse and reveals a critical role for NF-kappa B in leukocyte recruitment. Journal of immunology (Baltimore, Md : 1950) *167*, 1592-1600.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of molecular biology *215*, 403-410.

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112-118.

Amrein, H., and Axel, R. (1997). Genes expressed in neurons of adult male Drosophila. Cell *88*, 459-469.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Anderson, J.L., Mulligan, T.S., Shen, M.C., Wang, H., Scahill, C.M., Tan, F.J., Du, S.J., Busch-Nentwich, E.M., and Farber, S.A. (2017). mRNA processing in mutant zebrafish lines generated by chemical and CRISPR-mediated mutagenesis produces unexpected transcripts that escape nonsense-mediated decay. PLoS Genet *13*, e1007105.

Baek, S.T., Kerjan, G., Bielas, S.L., Lee, J.E., Fenstermaker, A.G., Novarino, G., and Gleeson, J.G. (2014). Off-target effect of doublecortin family shRNA on neuronal

migration associated with endogenous microRNA dysregulation. Neuron *82*, 1255-1262.

Barabasi, A.L., and Oltvai, Z.N. (2004). Network biology: understanding the cell's functional organization. Nat Rev Genet *5*, 101-113.

Barr, M.L., and Bertram, E.G. (1949). A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. Nature *163*, 676.

Barreto, G., Schafer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Doderlein, G., Maltry, N., Wu, W., Lyko, F.*, et al.* (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature *445*, 671-675.

Basak, S., Shih, V.F., and Hoffmann, A. (2008). Generation and activation of multiple dimeric transcription factors within the NF-kappaB signaling system. Molecular and cellular biology *28*, 3139-3150.

Becker, K.A., Ghule, P.N., Lian, J.B., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2010). Cyclin D2 and the CDK substrate p220(NPAT) are required for self-renewal of human embryonic stem cells. Journal of cellular physiology *222*, 456-464.

Berretta, J., Pinskaya, M., and Morillon, A. (2008). A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. Genes & development *22*, 615-626.

Blecher-Gonen, R., Barnett-Itzhaki, Z., Jaitin, D., Amann-Zalcenstein, D., Lara-Astiaso, D., and Amit, I. (2013). High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. Nature protocols *8*, 539-554.

Boehm, V., Gerbracht, J.V., Marx, M.C., and Gehring, N.H. (2016). Interrogating the degradation pathways of unstable mRNAs with XRN1-resistant sequences. Nature communications *7*, 13691.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

References

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell *123*, 1279-1291.

Bouche, N., and Bouchez, D. (2001). Arabidopsis gene knockout: phenotypes wanted. Current opinion in plant biology *4*, 111-117.

Bowerman, B., Eaton, B.A., and Priess, J.R. (1992). skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo. Cell *68*, 1061-1075.

Braun, N., Wyrzykowska, J., Muller, P., David, K., Couch, D., Perrot-Rechenmann, C., and Fleming, A.J. (2008). Conditional repression of AUXIN BINDING PROTEIN1 reveals that it coordinates cell division and cell expansion during postembryonic shoot development in Arabidopsis and tobacco. The Plant cell *20*, 2746-2762.

Bregman, A., Avraham-Kelbert, M., Barkai, O., Duek, L., Guterman, A., and Choder, M. (2011). Promoter elements regulate cytoplasmic mRNA decay. Cell *147*, 1473-1483.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature methods *10*, 1213-1218.

Bunnell, T.M., Burbach, B.J., Shimizu, Y., and Ervasti, J.M. (2011). beta-Actin specifically controls cell growth, migration, and the G-actin pool. Molecular biology of the cell *22*, 4047-4058.

Burgess, D.L., Biddlecome, G.H., McDonough, S.I., Diaz, M.E., Zilinski, C.A., Bean, B.P., Campbell, K.P., and Noebels, J.L. (1999). beta subunit reshuffling modifies N- and P/Q-type Ca2+ channel subunit compositions in lethargic mouse brain. Molecular and cellular neurosciences *13*, 293-311.

Caceres, A., and Kosik, K.S. (1990). Inhibition of neurite polarity by tau antisense oligonucleotides in primary cerebellar neurons. Nature *343*, 461-463.

References

Caceres, A., Potrebic, S., and Kosik, K.S. (1991). The effect of tau antisense oligonucleotides on neurite formation of cultured cerebellar macroneurons. The Journal of neuroscience : the official journal of the Society for Neuroscience *11*, 1515-1523.

Cadigan, K.M., Grossniklaus, U., and Gehring, W.J. (1994). Functional redundancy: the respective roles of the two sloppy paired genes in Drosophila segmentation. Proceedings of the National Academy of Sciences of the United States of America *91*, 6324-6328.

Carr, M.S., Yevtodiyenko, A., Schmidt, C.L., and Schmidt, J.V. (2007). Allele-specific histone modifications regulate expression of the Dlk1-Gtl2 imprinted domain. Genomics *89*, 280-290.

Celik, A., He, F., and Jacobson, A. (2017). NMD monitors translational fidelity 24/7. Current genetics *63*, 1007-1010.

Cerikan, B., Shaheen, R., Colo, G.P., Glasser, C., Hata, S., Knobeloch, K.P., Alkuraya, F.S., Fassler, R., and Schiebel, E. (2016). Cell-Intrinsic Adaptation Arising from Chronic Ablation of a Key Rho GTPase Regulator. Developmental cell *39*, 28-43.

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 research *39*, e82.

Chambeyron, S., and Bickmore, W.A. (2004). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes & development *18*, 1119-1130.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nature structural & molecular biology *15*, 85-93.

Chan, W.K., Bhalla, A.D., Le Hir, H., Nguyen, L.S., Huang, L., Gecz, J., and Wilkinson, M.F. (2009). A UPF3-mediated regulatory switch that maintains RNA surveillance. Nature structural & molecular biology *16*, 747-753.

References

Chandrasekaran, V., Juszkiewicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. Nature structural & molecular biology *26*, 1132-1140.

Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z., and Rowley, J.D. (2004). Over 20% of human transcripts might form sense-antisense pairs. Nucleic acids research *32*, 4812-4820.

Chen, P., Wang, D., Chen, H., Zhou, Z., and He, X. (2016a). The non-essentiality of essential genes in yeast provides therapeutic insights into a human disease. Genome research.

Chen, R., Shi, L., Hakenberg, J., Naughton, B., Sklar, P., Zhang, J., Zhou, H., Tian, L., Prakash, O., Lemire, M.*, et al.* (2016b). Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. Nature biotechnology *34*, 531-538.

Chen, X., Grandont, L., Li, H., Hauschild, R., Paque, S., Abuzeineh, A., Rakusova, H., Benkova, E., Perrot-Rechenmann, C., and Friml, J. (2014). Inhibition of cell expansion by rapid ABP1-mediated auxin effect on microtubules. Nature *516*, 90-93.

Chu, Y., Yue, X., Younger, S.T., Janowski, B.A., and Corey, D.R. (2010). Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. Nucleic acids research *38*, 7736-7748.

Chuzhanova, N.A., Anassis, E.J., Ball, E.V., Krawczak, M., and Cooper, D.N. (2003). Meta-analysis of indels causing human genetic disease: mechanisms of mutagenesis and the role of local DNA sequence complexity. Human mutation *21*, 28-44.

Ciemerych, M.A., Kenney, A.M., Sicinska, E., Kalaszczynska, I., Bronson, R.T., Rowitch, D.H., Gardner, H., and Sicinski, P. (2002). Development of mice expressing a single D-type cyclin. Genes & development *16*, 3277-3289.

Cloos, P.A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen, K.H., and Helin, K. (2006). The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. Nature *442*, 307-311.

References

Cohen, R., Yokoi, T., Holland, J.P., Pepper, A.E., and Holland, M.J. (1987). Transcription of the constitutively expressed yeast enolase gene ENO1 is mediated by positive and negative cis-acting regulatory sequences. Molecular and cellular biology *7*, 2753-2761.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M.*, et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

d'Adda di Fagagna, F. (2014). A direct role for small non-coding RNAs in DNA damage response. Trends in cell biology *24*, 171-178.

D'Orazio, K.N., Wu, C.C., Sinha, N., Loll-Krippleber, R., Brown, G.W., and Green, R. (2019). The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. eLife *8*.

Dabrowski, M., Bukowy-Bieryllo, Z., and Zietkiewicz, E. (2015). Translational readthrough potential of natural termination codons in eucaryotes--The impact of RNA sequence. RNA biology *12*, 950-958.

Daude, N., Wohlgemuth, S., Brown, R., Pitstick, R., Gapeshina, H., Yang, J., Carlson, G.A., and Westaway, D. (2012). Knockout of the prion protein (PrP)-like Sprn gene does not produce embryonic lethality in combination with PrP(C)-deficiency. Proceedings of the National Academy of Sciences of the United States of America *109*, 9035-9040.

Davidson, E., and Levin, M. (2005). Gene regulatory networks. Proceedings of the National Academy of Sciences of the United States of America *102*, 4935.

Davies, E.K., Peters, A.D., and Keightley, P.D. (1999). High frequency of cryptic deleterious mutations in Caenorhabditis elegans. Science (New York, NY) *285*, 1748-1751.

Davis, M.P., van Dongen, S., Abreu-Goodger, C., Bartonicek, N., and Enright, A.J. (2013). Kraken: a set of tools for quality control and analysis of high-throughput sequence data. Methods (San Diego, Calif) *63*, 41-49.

References

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.W., Page, D.C.*, et al.* (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell stem cell *9*, 166-175.

Dayton, T.L., Gocheva, V., Miller, K.M., Israelsen, W.J., Bhutkar, A., Clish, C.B., Davidson, S.M., Luengo, A., Bronson, R.T., Jacks, T.*, et al.* (2016). Germline loss of PKM2 promotes metabolic distress and hepatocellular carcinoma. Genes & development *30*, 1020-1033.

De Grassi, A., Lanave, C., and Saccone, C. (2008). Genome duplication and gene-family evolution: the case of three OXPHOS gene families. Gene *421*, 1-6.

De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G., and Natoli, G. (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. Cell *130*, 1083-1094.

De Souza, A.T., Dai, X., Spencer, A.G., Reppen, T., Menzie, A., Roesch, P.L., He, Y., Caguyong, M.J., Bloomer, S., Herweijer, H.*, et al.* (2006). Transcriptional and phenotypic comparisons of Ppara knockout and siRNA knockdown mice. Nucleic acids research *34*, 4486-4494.

Deconinck, A.E., Rafael, J.A., Skinner, J.A., Brown, S.C., Potter, A.C., Metzinger, L., Watt, D.J., Dickson, J.G., Tinsley, J.M., and Davies, K.E. (1997). Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. Cell *90*, 717-727.

Deniaud, A., Karuppasamy, M., Bock, T., Masiulis, S., Huard, K., Garzoni, F., Kerschgens, K., Hentze, M.W., Kulozik, A.E., Beck, M.*, et al.* (2015). A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. Nucleic acids research *43*, 7600-7611.

Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H.*, et al.* (2016). High-throughput discovery of novel developmental phenotypes. Nature *537*, 508-514.

Dietz, H.C., McIntosh, I., Sakai, L.Y., Corson, G.M., Chalberg, S.C., Pyeritz, R.E., and Francomano, C.A. (1993). Four novel FBN1 mutations: significance for mutant

transcript level and EGF-like domain calcium binding in the pathogenesis of Marfan syndrome. Genomics *17*, 468-475.

Diss, G., Ascencio, D., DeLuna, A., and Landry, C.R. (2014). Molecular mechanisms of paralogous compensation and the robustness of cellular networks. Journal of experimental zoology Part B, Molecular and developmental evolution *322*, 488-499.

Diss, G., Dube, A.K., Boutin, J., Gagnon-Arsenault, I., and Landry, C.R. (2013). A systematic approach for the genetic dissection of protein complexes in living cells. Cell reports *3*, 2155-2167.

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.

Doi, T.S., Marino, M.W., Takahashi, T., Yoshida, T., Sakakura, T., Old, L.J., and Obata, Y. (1999). Absence of tumor necrosis factor rescues RelA-deficient mice from embryonic lethality. Proceedings of the National Academy of Sciences of the United States of America *96*, 2994-2999.

Dostie, J., and Dreyfuss, G. (2002). Translation is required to remove Y14 from mRNAs in the cytoplasm. Current biology : CB *12*, 1060-1067.

Downs, J.A., Nussenzweig, M.C., and Nussenzweig, A. (2007). Chromatin dynamics and the preservation of genetic information. Nature *447*, 951-958.

Doyle, E.L., Hummel, A.W., Demorest, Z.L., Starker, C.G., Voytas, D.F., Bradley, P., and Bogdanove, A.J. (2013). TAL effector specificity for base 0 of the DNA target is altered in a complex, effector- and assay-dependent manner by substitutions for the tryptophan in cryptic repeat -1. PloS one *8*, e82120.

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M.H., and Chandrasegaran, S. (2005). Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic acids research *33*, 5978-5990.

Durand, S., Franks, T.M., and Lykke-Andersen, J. (2016). Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. Nature communications *7*, 12434.

References

Durbin, R.M., Altshuler, D., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Collins, F.S., De La Vega, F.M., Donnelly, P.*, et al.* (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061-1073.

Eberle, A.B., Lykke-Andersen, S., Muhlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nature structural & molecular biology *16*, 49-55.

Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Muhlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS biology *6*, e92.

Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics *5*, 113.

Eisensmith, R.C., and Woo, S.L. (1992). Molecular basis of phenylketonuria and related hyperphenylalaninemias: mutations and polymorphisms in the human phenylalanine hydroxylase gene. Human mutation *1*, 13-23.

El-Brolosy, M.A., Kontarakis, Z., Rossi, A., Kuenne, C., Gunther, S., Fukuda, N., Kikhi, K., Boezio, G.L.M., Takacs, C.M., Lai, S.L.*, et al.* (2019). Genetic compensation triggered by mutant mRNA degradation. Nature *568*, 193-197.

El-Brolosy, M.A., and Stainier, D.Y.R. (2017). Genetic compensation: A phenomenon in search of mechanisms. PLoS Genet *13*, e1006780.

Elkon, R., Zlotorynski, E., Zeller, K.I., and Agami, R. (2010). Major role for mRNA stability in shaping the kinetics of gene induction. BMC genomics *11*, 259.

Evers, B., Jastrzebski, K., Heijmans, J.P., Grernrum, W., Beijersbergen, R.L., and Bernards, R. (2016). CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. Nature biotechnology *34*, 631-633.

Faghihi, M.A., and Wahlestedt, C. (2009). Regulatory roles of natural antisense transcripts. Nat Rev Mol Cell Biol *10*, 637-643.

Ferec, C., and Cutting, G.R. (2012). Assessing the Disease-Liability of Mutations in CFTR. Cold Spring Harbor Perspectives in Medicine *2*, a009480.

References

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet *9*, 102-114.

Fodor, B.D., Kubicek, S., Yonezawa, M., O'Sullivan, R.J., Sengupta, R., Perez-Burgos, L., Opravil, S., Mechtler, K., Schotta, G., and Jenuwein, T. (2006). Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. Genes & development *20*, 1557-1562.

Fournier, C., Goto, Y., Ballestar, E., Delaval, K., Hever, A.M., Esteller, M., and Feil, R. (2002). Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. The EMBO journal *21*, 6560-6570.

Francia, S., Michelini, F., Saxena, A., Tang, D., de Hoon, M., Anelli, V., Mione, M., Carninci, P., and d'Adda di Fagagna, F. (2012). Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature *488*, 231-235.

Franke, A., and Baker, B.S. (1999). The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Molecular cell *4*, 117-122.

Freudenberg, J.M., Ghosh, S., Lackford, B.L., Yellaboina, S., Zheng, X., Li, R., Cuddapah, S., Wade, P.A., Hu, G., and Jothi, R. (2012). Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. Nucleic acids research *40*, 3364-3377.

Frischmeyer, P.A., van Hoof, A., O'Donnell, K., Guerrerio, A.L., Parker, R., and Dietz, H.C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. Science (New York, NY) *295*, 2258-2261.

Fukunaga, Y., Liu, H., Shimizu, M., Komiya, S., Kawasuji, M., and Nagafuchi, A. (2005). Defining the roles of beta-catenin and plakoglobin in cell-cell adhesion: isolation of beta-catenin/plakoglobin-deficient F9 cells. Cell structure and function *30*, 25-34.

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.

Gao, Y., Zhang, Y., Zhang, D., Dai, X., Estelle, M., and Zhao, Y. (2015). Auxin binding protein 1 (ABP1) is not required for either auxin signaling or Arabidopsis development. Proceedings of the National Academy of Sciences *112*, 2275-2280.

Gapuzan, M.E., Schmah, O., Pollock, A.D., Hoffmann, A., and Gilmore, T.D. (2005). Immortalized fibroblasts from NF-kappaB RelA knockout mice show phenotypic heterogeneity and maintain increased sensitivity to tumor necrosis factor alpha after transformation by v-Ras. Oncogene *24*, 6574-6583.

Gapuzan, M.E., Yufit, P.V., and Gilmore, T.D. (2002). Immortalized embryonic mouse fibroblasts lacking the RelA subunit of transcription factor NF-kappaB have a malignantly transformed phenotype. Oncogene *21*, 2484-2492.

Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W., and Kulozik, A.E. (2003). Y14 and hUpf3b form an NMD-activating complex. Molecular cell *11*, 939-949.

Genschel, J., and Schmidt, H.H. (2000). Mutations in the LMNA gene encoding lamin A/C. Human mutation *16*, 451-459.

Gerber, A.P., Herschlag, D., and Brown, P.O. (2004). Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. PLoS biology *2*, E79.

Gerri, C., Marin-Juez, R., Marass, M., Marks, A., Maischein, H.M., and Stainier, D.Y.R. (2017). Hif-1alpha regulates macrophage-endothelial interactions during blood vessel development in zebrafish. Nature communications *8*, 15492.

Ghahramani Seno, M.M., Graham, I.R., Athanasopoulos, T., Trollet, C., Pohlschmidt, M., Crompton, M.R., and Dickson, G. (2008). RNAi-mediated knockdown of dystrophin expression in adult mice does not lead to overt muscular dystrophy pathology. Human molecular genetics *17*, 2622-2632.

Ghanbarian, H., Wagner, N., Michiels, J.F., Cuzin, F., Wagner, K.D., and Rassoulzadegan, M. (2017). Small RNA-directed epigenetic programming of embryonic stem cell cardiac differentiation. Scientific reports *7*, 41799.

References

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B.*, et al.* (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387-391.

Glasauer, S.M., and Neuhauss, S.C. (2014). Whole-genome duplication in teleost fishes and its evolutionary consequences. Molecular genetics and genomics : MGG *289*, 1045-1060.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS letters *582*, 1977-1986.

Glover, M.L., Burroughs, A.M., Egelhofer, T.A., Pule, M.N., Aravind, L., and Arribere, J.A. (2019). NONU-1 encodes a conserved endonuclease required for mRNA translation surveillance. bioRxiv, 674358.

Goler-Baron, V., Selitrennik, M., Barkai, O., Haimovich, G., Lotan, R., and Choder, M. (2008). Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes. Genes & development *22*, 2022-2027.

González-Gaitán, M., Rothe, M., Wimmer, E.A., Taubert, H., and Jäckle, H. (1994). Redundant functions of the genes knirps and knirps-related for the establishment of anterior Drosophila head structures. Proceedings of the National Academy of Sciences of the United States of America *91*, 8567-8571.

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R., and Hofacker, I.L. (2008). The Vienna RNA websuite. Nucleic acids research *36*, W70-74.

Gu, Z., Steinmetz, L.M., Gu, X., Scharfe, C., Davis, R.W., and Li, W.H. (2003). Role of duplicate genes in genetic robustness against null mutations. Nature *421*, 63-66.

Guidi, C.J., Veal, T.M., Jones, S.N., and Imbalzano, A.N. (2004). Transcriptional compensation for loss of an allele of the Ini1 tumor suppressor. The Journal of biological chemistry *279*, 4180-4185.

Hagelkruys, A., Lagger, S., Krahmer, J., Leopoldi, A., Artaker, M., Pusch, O., Zezula, J., Weissmann, S., Xie, Y., Schofer, C.*, et al.* (2014). A single allele of Hdac2 but not Hdac1 is sufficient for normal mouse brain development in the absence of its paralog. Development (Cambridge, England) *141*, 604-616.

References

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millan-Zambrano, G., Barkai, O., Chavez, S., Perez-Ortin, J.E., Darzacq, X., and Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000-1011.

Hall, E.A., Keighren, M., Ford, M.J., Davey, T., Jarman, A.P., Smith, L.B., Jackson, I.J., and Mill, P. (2013). Acute versus chronic loss of mammalian Azi1/Cep131 results in distinct ciliary phenotypes. PLoS Genet *9*, e1003928.

Hall, G.W., and Thein, S. (1994). Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia. Blood *83*, 2031-2037.

Hao, S., and Baltimore, D. (2009). The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. Nat Immunol *10*, 281-288.

Harada, A., Oguchi, K., Okabe, S., Kuno, J., Terada, S., Ohshima, T., Sato-Yoshitake, R., Takei, Y., Noda, T., and Hirokawa, N. (1994). Altered microtubule organization in small-calibre axons of mice lacking tau protein. Nature *369*, 488-491.

Hardwick, K.G., Boothroyd, J.C., Rudner, A.D., and Pelham, H.R. (1992). Genes that allow yeast cells to grow in the absence of the HDEL receptor. The EMBO journal *11*, 4187-4195.

Harigaya, Y., and Parker, R. (2010). No-go decay: a quality control mechanism for RNA in translation. Wiley interdisciplinary reviews RNA *1*, 132-141.

Hartenian, E., and Glaunsinger, B.A. (2019). Feedback to the central dogma: cytoplasmic mRNA decay and transcription are interdependent processes. Critical reviews in biochemistry and molecular biology *54*, 385-398.

Heard, E., and Disteche, C.M. (2006). Dosage compensation in mammals: fine-tuning the expression of the X chromosome. Genes & development *20*, 1848-1867.

References

Hebbard, L.W., Maurer, J., Miller, A., Lesperance, J., Hassell, J., Oshima, R.G., and Terskikh, A.V. (2010). Maternal embryonic leucine zipper kinase is upregulated and required in mammary tumor-initiating cells in vivo. Cancer Res *70*, 8863-8873.

Hinnebusch, A.G., Dever, T.E., and Asano, K. (2007). 9 Mechanism of Translation Initiation in the Yeast Saccharomyces cerevisiae.

Hoek, T.A., Khuperkar, D., Lindeboom, R.G.H., Sonneveld, S., Verhagen, B.M.P., Boersma, S., Vermeulen, M., and Tanenbaum, M.E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. Molecular cell *75*, 324-339.e311.

Hoffmann, A., Leung, T.H., and Baltimore, D. (2003). Genetic analysis of NF-kappaB/Rel transcription factors defines functional specificities. The EMBO journal *22*, 5530-5539.

Hoffmann, F.M. (1991). Drosophila abl and genetic redundancy in signal transduction. Trends in genetics : TIG *7*, 351-355.

Housden, B.E., Muhar, M., Gemberling, M., Gersbach, C.A., Stainier, D.Y., Seydoux, G., Mohr, S.E., Zuber, J., and Perrimon, N. (2017). Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. Nat Rev Genet *18*, 24-40.

Hsiao, T.L., and Vitkup, D. (2008). Role of duplicate genes in robustness against deleterious human mutations. PLoS Genet *4*, e1000014.

Hu, J., Chen, Z., Xia, D., Wu, J., Xu, H., and Ye, Z.Q. (2012). Promoter-associated small double-stranded RNA interacts with heterogeneous nuclear ribonucleoprotein A2/B1 to induce transcriptional activation. The Biochemical journal *447*, 407-416.

Huard, J.M., Forster, C.C., Carter, M.L., Sicinski, P., and Ross, M.E. (1999). Cerebellar histogenesis is disturbed in mice lacking cyclin D2. Development (Cambridge, England) *126*, 1927-1935.

Huntzinger, E., Kashima, I., Fauser, M., Sauliere, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. Rna *14*, 2609-2617.

References

Ibrahim, F., Maragkakis, M., Alexiou, P., and Mourelatos, Z. (2018). Ribothrypsis, a novel process of canonical mRNA decay, mediates ribosome-phased mRNA endonucleolysis. Nature structural & molecular biology *25*, 302-310.

Ibrahim, F., and Mourelatos, Z. (2019). Capturing 5' and 3' native ends of mRNAs concurrently with Akron sequencing. Nature protocols *14*, 1578-1602.

Ikeuchi, K., Tesina, P., Matsuo, Y., Sugiyama, T., Cheng, J., Saeki, Y., Tanaka, K., Becker, T., Beckmann, R., and Inada, T. (2019). Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. The EMBO journal *38*.

Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes & development *21*, 1833-1856.

Ito-Harashima, S., Kuroha, K., Tatematsu, T., and Inada, T. (2007). Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. Genes & development *21*, 519-524.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. The EMBO journal *27*, 736-747.

Jackson, A.L., and Linsley, P.S. (2010). Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. Nature reviews Drug discovery *9*, 57-67.

Jackson, J.G., and Pereira-Smith, O.M. (2006). Primary and compensatory roles for RB family members at cell cycle gene promoters that are deacetylated and downregulated in doxorubicin-induced senescence of breast cancer cells. Molecular and cellular biology *26*, 2501-2510.

Jacobsen, A., Silber, J., Harinath, G., Huse, J.T., Schultz, N., and Sander, C. (2013). Analysis of microRNA-target interactions across diverse cancer types. Nature structural & molecular biology *20*, 1325-1332.

References

Jagannathan, S., and Bradley, R.K. (2016). Translational plasticity facilitates the accumulation of nonsense genetic variants in the human population. Genome research *26*, 1639-1650.

Janowski, B.A., Younger, S.T., Hardy, D.B., Ram, R., Huffman, K.E., and Corey, D.R. (2007). Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. Nature chemical biology *3*, 166-173.

Jao, L.E., Wente, S.R., and Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proceedings of the National Academy of Sciences of the United States of America *110*, 13904-13909.

Jin, Q., Yu, L.R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.E., Wang, C., Brindle, P.K., Dent, S.Y., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. The EMBO journal *30*, 249-262.

Jost, A.P., and Weiner, O.D. (2015). Probing Yeast Polarity with Acute, Reversible, Optogenetic Inhibition of Protein Function. ACS synthetic biology *4*, 1077-1085.

Jurkin, J., Zupkovitz, G., Lagger, S., Grausenburger, R., Hagelkruys, A., Kenner, L., and Seiser, C. (2011). Distinct and redundant functions of histone deacetylases HDAC1 and HDAC2 in proliferation and tumorigenesis. Cell Cycle *10*, 406-412.

Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., and Hegde, R.S. (2018). ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Molecular cell *72*, 469-481.e467.

Juszkiewicz, S., and Hegde, R.S. (2017). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. Molecular cell *65*, 743-750.e744.

Karakas, B., Weeraratna, A.T., Abukhdeir, A.M., Konishi, H., Gustin, J.P., Vitolo, M.I., Bachman, K.E., and Park, B.H. (2007). P21 gene knock down does not identify genetic effectors seen with gene knock out. Cancer biology & therapy *6*, 1025-1030.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3

complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes & development *20*, 355-367.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J.*, et al.* (2005). Antisense transcription in the mammalian transcriptome. Science (New York, NY) *309*, 1564-1566.

Keene, J.D. (2007). RNA regulons: coordination of post-transcriptional events. Nat Rev Genet *8*, 533-543.

Keene, J.D., and Lager, P.J. (2005). Post-transcriptional operons and regulons co-ordinating gene expression. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *13*, 327-337.

Keene, J.D., and Tenenbaum, S.A. (2002). Eukaryotic mRNPs may represent posttranscriptional operons. Molecular cell *9*, 1161-1167.

Kervestin, S., Li, C., Buckingham, R., and Jacobson, A. (2012). Testing the faux-UTR model for NMD: analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p. Biochimie *94*, 1560-1571.

Kettleborough, R.N., Busch-Nentwich, E.M., Harvey, S.A., Dooley, C.M., de Bruijn, E., van Eeden, F., Sealy, I., White, R.J., Herd, C., Nijman, I.J.*, et al.* (2013). A systematic genome-wide analysis of zebrafish protein-coding gene function. Nature *496*, 494-497.

Kim, V.N., Kataoka, N., and Dreyfuss, G. (2001). Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. Science (New York, NY) *293*, 1832-1836.

Kiyosawa, H., Yamanaka, I., Osato, N., Kondo, S., Riken Ger Group Laboratory for Genome Exploration Research Group, R.G.S.C.R.Y.I.S.-c.T.-k.Y.K.J., Gslmembers Genome Science Laboratory, R.H.W.S.J., and Hayashizaki, Y. (2003). Antisense Transcripts With FANTOM2 Clone Set and Their Implications for Gene Regulation. Genome research *13*, 1324-1334.

References

Klauer, A.A., and van Hoof, A. (2012). Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. Wiley interdisciplinary reviews RNA *3*, 649-660.

Kok, F.O., Shin, M., Ni, C.W., Gupta, A., Grosse, A.S., van Impel, A., Kirchmaier, B.C., Peterson-Maduro, J., Kourkoulis, G., Male, I.*, et al.* (2015). Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. Developmental cell *32*, 97-108.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H.*, et al.* (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583-588.

Kunz, J.B., Neu-Yilik, G., Hentze, M.W., Kulozik, A.E., and Gehring, N.H. (2006). Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation. Rna *12*, 1015-1022.

Kurosaki, T., Li, W., Hoque, M., Popp, M.W., Ermolenko, D.N., Tian, B., and Maquat, L.E. (2014). A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes & development *28*, 1900-1916.

Kurosaki, T., Miyoshi, K., Myers, J.R., and Maquat, L.E. (2018). NMD-degradome sequencing reveals ribosome-bound intermediates with 3'-end non-templated nucleotides. Nature structural & molecular biology *25*, 940-950.

Kurosaki, T., Popp, M.W., and Maquat, L.E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. Nat Rev Mol Cell Biol *20*, 406-420.

Kuwano, Y., Rabinovic, A., Srikantan, S., Gorospe, M., and Demple, B. (2009). Analysis of nitric oxide-stabilized mRNAs in human fibroblasts reveals HuR-dependent heme oxygenase 1 upregulation. Molecular and cellular biology *29*, 2622-2635.

Kwon, H.B., Wang, S., Helker, C.S., Rasouli, S.J., Maischein, H.M., Offermanns, S., Herzog, W., and Stainier, D.Y. (2016). In vivo modulation of endothelial polarization by Apelin receptor signalling. Nature communications *7*, 11805.

References

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*, W272-276.

Lagger, G., O'Carroll, D., Rembold, M., Khier, H., Tischler, J., Weitzer, G., Schuettengruber, B., Hauser, C., Brunmeir, R., Jenuwein, T.*, et al.* (2002). Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. The EMBO journal *21*, 2672-2681.

Lalonde, S., Stone, O.A., Lessard, S., Lavertu, A., Desjardins, J., Beaudoin, M., Rivas, M., Stainier, D.Y.R., and Lettre, G. (2017). Frameshift indels introduced by genome editing can lead to in-frame exon skipping. PloS one *12*, e0178700.

Lam, E.W., Glassford, J., Banerji, L., Thomas, N.S., Sicinski, P., and Klaus, G.G. (2000). Cyclin D3 compensates for loss of cyclin D2 in mouse B-lymphocytes activated via the antigen receptor and CD40. The Journal of biological chemistry *275*, 3479-3484.

Lan, F., Bayliss, P.E., Rinn, J.L., Whetstine, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E.*, et al.* (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. Nature *449*, 689-694.

Law, D.J., Allen, D.L., and Tidball, J.G. (1994). Talin, vinculin and DRP (utrophin) concentrations are increased at mdx myotendinous junctions following onset of necrosis. Journal of cell science *107 (Pt 6)*, 1477-1483.

Law, S.H., and Sargent, T.D. (2014). The serine-threonine protein kinase PAK4 is dispensable in zebrafish: identification of a morpholino-generated pseudophenotype. PloS one *9*, e100268.

Lawson, N.D., and Weinstein, B.M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. Developmental biology *248*, 307-318.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. The EMBO journal *20*, 4987-4997.

References

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. The EMBO journal *19*, 6860-6869.

Le Hir, H., Sauliere, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. Nat Rev Mol Cell Biol *17*, 41-54.

Lee, S.R., Pratt, G.A., Martinez, F.J., Yeo, G.W., and Lykke-Andersen, J. (2015). Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity. Molecular cell *59*, 413-425.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B.*, et al.* (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285-291.

Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H.Q., Ong, C.T., Cubenas-Potts, C., Hu, M., Lei, E.P., Bosco, G.*, et al.* (2015). Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. Molecular cell *58*, 216-231.

Li, L.C., Okino, S.T., Zhao, H., Pookot, D., Place, R.F., Urakami, S., Enokida, H., and Dahiya, R. (2006). Small dsRNAs induce transcriptional activation in human cells. Proceedings of the National Academy of Sciences of the United States of America *103*, 17337-17342.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Licht, K., and Jantsch, M.F. (2016). Rapid and dynamic transcriptome regulation by RNA editing and RNA modifications. J Cell Biol *213*, 15-22.

Lin, A., Giuliano, C.J., Sayles, N.M., and Sheltzer, J.M. (2017). CRISPR/Cas9 mutagenesis invalidates a putative cancer dependency targeted in on-going clinical trials. eLife *6*.

Lin, M.L., Park, J.H., Nishidate, T., Nakamura, Y., and Katagiri, T. (2007). Involvement of maternal embryonic leucine zipper kinase (MELK) in mammary

carcinogenesis through interaction with Bcl-G, a pro-apoptotic member of the Bcl-2 family. Breast cancer research : BCR *9*, R17.

Lindahl, T., and Barnes, D.E. (2000). Repair of endogenous DNA damage. Cold Spring Harbor symposia on quantitative biology *65*, 127-133.

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes & development *27*, 2125-2138.

Lotan, R., Bar-On, V.G., Harel-Sharvit, L., Duek, L., Melamed, D., and Choder, M. (2005). The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. Genes & development *19*, 3004-3016.

Lotan, R., Goler-Baron, V., Duek, L., Haimovich, G., and Choder, M. (2007). The Rpb7p subunit of yeast RNA polymerase II plays roles in the two major cytoplasmic mRNA decay mechanisms. J Cell Biol *178*, 1133-1143.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Lu, J., and Deutsch, C. (2008). Electrostatics in the ribosomal tunnel modulate chain elongation rates. Journal of molecular biology *384*, 73-86.

Lunt, S.Y., Muralidhar, V., Hosios, A.M., Israelsen, W.J., Gui, D.Y., Newhouse, L., Ogrodzinski, M., Hecht, V., Xu, K., Acevedo, P.N.*, et al.* (2015). Pyruvate kinase isoform expression alters nucleotide synthesis to impact cell proliferation. Molecular cell *57*, 95-107.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell *103*, 1121-1131.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science (New York, NY) *293*, 1836-1839.

Lykke-Andersen, S., and Jensen, T.H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol *16*, 665-677.

References

Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372-373.

Ma, T., Yang, B., Gillespie, A., Carlson, E.J., Epstein, C.J., and Verkman, A.S. (1997). Generation and phenotype of a transgenic knockout mouse lacking the mercurial-insensitive water channel aquaporin-4. J Clin Invest *100*, 957-962.

Ma, Z., Zhu, P., Shi, H., Guo, L., Zhang, Q., Chen, Y., Chen, S., Zhang, Z., Peng, J., and Chen, J. (2019). PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. Nature *568*, 259-263.

MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J.K., Montgomery, S.B.*, et al.* (2012). A systematic survey of loss-of-function variants in human protein-coding genes. Science (New York, NY) *335*, 823-828.

Malabat, C., Feuerbach, F., Ma, L., Saveanu, C., and Jacquier, A. (2015). Quality control of transcription start site selection by nonsense-mediated-mRNA decay. eLife *4*.

Manley, G.T., Fujimura, M., Ma, T., Noshita, N., Filiz, F., Bollen, A.W., Chan, P., and Verkman, A.S. (2000). Aquaporin-4 deletion in mice reduces brain edema after acute water intoxication and ischemic stroke. Nature medicine *6*, 159-163.

Manolson, M.F., Wu, B., Proteau, D., Taillon, B.E., Roberts, B.T., Hoyt, M.A., and Jones, E.W. (1994). STV1 gene encodes functional homologue of 95-kDa yeast vacuolar H(+)-ATPase subunit Vph1p. The Journal of biological chemistry *269*, 14064-14074.

Martin, L., Grigoryan, A., Wang, D., Wang, J., Breda, L., Rivella, S., Cardozo, T., and Gardner, L.B. (2014). Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. Cancer Res *74*, 3104-3113.

Mather, K. (1953). Genetical control of stability in development. Heredity *7*, 297-336.

Matsui, M., Sakurai, F., Elbashir, S., Foster, D.J., Manoharan, M., and Corey, D.R. (2010). Activation of LDL receptor expression by small RNAs complementary to a

noncoding transcript that overlaps the LDLR promoter. Chemistry & biology *17*, 1344-1355.

McEwen, K.R., and Ferguson-Smith, A.C. (2009). Genomic Imprinting – A Model for Roles of HistoneModifications in Epigenetic Control. In Epigenomics, A.C. Ferguson-Smith, J.M. Greally, and R.A. Martienssen, eds. (Dordrecht: Springer Netherlands), pp. 235-258.

McJunkin, K., Mazurek, A., Premsrirut, P.K., Zuber, J., Dow, L.E., Simon, J., Stillman, B., and Lowe, S.W. (2011). Reversible suppression of an essential gene in adult mice using transgenic RNA interference. Proceedings of the National Academy of Sciences of the United States of America *108*, 7113-7118.

Meller, V.H., Wu, K.H., Roman, G., Kuroda, M.I., and Davis, R.L. (1997). roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system. Cell *88*, 445-457.

Mendell, J.T., ap Rhys, C.M., and Dietz, H.C. (2002). Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts. Science (New York, NY) *298*, 419-422.

Modarresi, F., Faghihi, M.A., Lopez-Toledano, M.A., Fatemi, R.P., Magistri, M., Brothers, S.P., van der Brug, M.P., and Wahlestedt, C. (2012). Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. Nature biotechnology *30*, 453-459.

Morgens, D.W., Deans, R.M., Li, A., and Bassik, M.C. (2016). Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. Nature biotechnology *34*, 634-636.

Morris, K.V., Santoso, S., Turner, A.M., Pastori, C., and Hawkins, P.G. (2008). Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. PLoS Genet *4*, e1000258.

Mueller, P.P., and Hinnebusch, A.G. (1986). Multiple upstream AUG codons mediate translational control of GCN4. Cell *45*, 201-207.

References

Muhlemann, O., Mock-Casagrande, C.S., Wang, J., Li, S., Custodio, N., Carmo-Fonseca, M., Wilkinson, M.F., and Moore, M.J. (2001). Precursor RNAs harboring nonsense codons accumulate near the site of transcription. Molecular cell *8*, 33-43.

Mukherjee, A.S., and Beermann, W. (1965). Synthesis of ribonucleic acid by the X-chromosomes of Drosophila melanogaster and the problem of dosage compensation. Nature *207*, 785-786.

Mukherjee, C., Patil, D.P., Kennedy, B.A., Bakthavachalu, B., Bundschuh, R., and Schoenberg, D.R. (2012). Identification of cytoplasmic capping targets reveals a role for cap homeostasis in translation and mRNA stability. Cell reports *2*, 674-684.

Muller, H. (1932). Further studies on the nature and causes of gene mutations.

Mulligan, G.J., Wong, J., and Jacks, T. (1998). p130 is dispensable in peripheral T lymphocytes: evidence for functional compensation by p107 and pRB. Molecular and cellular biology *18*, 206-220.

Murshed, M., Smyth, N., Miosge, N., Karolat, J., Krieg, T., Paulsson, M., and Nischt, R. (2000). The absence of nidogen 1 does not affect murine basement membrane formation. Molecular and cellular biology *20*, 7007-7012.

Myerowitz, R. (1997). Tay-Sachs disease-causing mutations and neutral polymorphisms in the Hex A gene. Human mutation *9*, 195-208.

Nadeau, J.H. (2001). Modifier genes in mice and humans. Nat Rev Genet *2*, 165-174.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'-->3' exoribonucleases: structure, mechanisms and functions. Biochimica et biophysica acta *1829*, 590-603.

Nagy, E., and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends in biochemical sciences *23*, 198-199.

Narasimhan, V.M., Hunt, K.A., Mason, D., Baker, C.L., Karczewski, K.J., Barnes, M.R., Barnett, A.H., Bates, C., Bellary, S., Bockett, N.A.*, et al.* (2016). Health and

population effects of rare gene knockouts in adult humans with related parents. Science (New York, NY) *352*, 474-477.

Ng, P.C., Levy, S., Huang, J., Stockwell, T.B., Walenz, B.P., Li, K., Axelrod, N., Busam, D.A., Strausberg, R.L., and Venter, J.C. (2008). Genetic variation in an individual human exome. PLoS Genet *4*, e1000160.

Nicchia, G.P., Srinivas, M., Li, W., Brosnan, C.F., Frigeri, A., and Spray, D.C. (2005). New possible roles for aquaporin-4 in astrocytes: cell cytoskeleton and functional relationship with connexin43. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *19*, 1674-1676.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P., and Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. Genes & development *16*, 479-489.

Nowak, K.J., Sewry, C.A., Navarro, C., Squier, W., Reina, C., Ricoy, J.R., Jayawant, S.S., Childs, A.M., Dobbie, J.A., Appleton, R.E.*, et al.* (2007). Nemaline myopathy caused by absence of alpha-skeletal muscle actin. Annals of neurology *61*, 175-184.

O'Leary, M.N., Schreiber, K.H., Zhang, Y., Duc, A.C., Rao, S., Hale, J.S., Academia, E.C., Shah, S.R., Morton, J.F., Holstein, C.A.*, et al.* (2013). The ribosomal protein Rpl22 controls ribosome composition by directly repressing expression of its own paralog, Rpl22l1. PLoS Genet *9*, e1003708.

Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H., and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell *87*, 953-959.

Ohno, S. (1970). Evolution by Gene Duplication. George Allen and Unwin, London.

Olejniczak, M., Galka, P., and Krzyzosiak, W.J. (2010). Sequence-non-specific effects of RNA interference triggers and microRNA regulators. Nucleic acids research *38*, 1-16.

Olejniczak, M., Urbanek, M.O., Jaworska, E., Witucki, L., Szczesniak, M.W., Makalowska, I., and Krzyzosiak, W.J. (2016). Sequence-non-specific effects

generated by various types of RNA interference triggers. Biochimica et biophysica acta *1859*, 306-314.

Orom, U.A., Nielsen, F.C., and Lund, A.H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Molecular cell *30*, 460-471.

Osato, N., Suzuki, Y., Ikeo, K., and Gojobori, T. (2007). Transcriptional interferences in cis natural antisense transcripts of humans and mice. Genetics *176*, 1299-1306.

Papadopoulos, M.C., Manley, G.T., Krishna, S., and Verkman, A.S. (2004). Aquaporin-4 facilitates reabsorption of excess fluid in vasogenic brain edema. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *18*, 1291-1293.

Parker, L.H., Schmidt, M., Jin, S.W., Gray, A.M., Beis, D., Pham, T., Frantz, G., Palmieri, S., Hillan, K., Stainier, D.Y.*, et al.* (2004). The endothelial-cell-derived secreted factor Egfl7 regulates vascular tube formation. Nature *428*, 754-758.

Pasquinelli, A.E. (2012). MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. Nat Rev Genet *13*, 271-282.

Patrinostro, X., O'Rourke, A.R., Chamberlain, C.M., Moriarity, B.S., Perrin, B.J., and Ervasti, J.M. (2017). Relative importance of betacyto- and gammacyto-actin in primary mouse embryonic fibroblasts. Molecular biology of the cell *28*, 771-782.

Pauli, A., Valen, E., Lin, M.F., Garber, M., Vastenhouw, N.L., Levin, J.Z., Fan, L., Sandelin, A., Rinn, J.L., Regev, A.*, et al.* (2012). Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome research *22*, 577-591.

Peach, S.E., York, K., and Hesselberth, J.R. (2015). Global analysis of RNA cleavage by 5'-hydroxyl RNA sequencing. Nucleic acids research *43*, e108.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2016). Genome-wide quantification of 5'-phosphorylated mRNA degradation intermediates for analysis of ribosome dynamics. Nature protocols *11*, 359-376.

References

Pinskaya, M., Gourvennec, S., and Morillon, A. (2009). H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation. The EMBO journal *28*, 1697-1707.

Place, R.F., Li, L.C., Pookot, D., Noonan, E.J., and Dahiya, R. (2008). MicroRNA-373 induces expression of genes with complementary promoter sequences. Proceedings of the National Academy of Sciences of the United States of America *105*, 1608-1613.

Portnoy, V., Lin, S.H., Li, K.H., Burlingame, A., Hu, Z.H., Li, H., and Li, L.C. (2016). saRNA-guided Ago2 targets the RITA complex to promoters to stimulate transcription. Cell research *26*, 320-335.

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R.*, et al.* (2015). Codon optimality is a major determinant of mRNA stability. Cell *160*, 1111-1124.

Protin, U., Schweighoffer, T., Jochum, W., and Hilberg, F. (1999). CD44-deficient mice develop normally with changes in subpopulations and recirculation of lymphocyte subsets. Journal of immunology (Baltimore, Md : 1950) *163*, 4917-4923.

Proulx, K., Lu, A., and Sumanas, S. (2010). Cranial vasculature in zebrafish forms by angioblast cluster-derived angiogenesis. Developmental biology *348*, 34-46.

Prykhozhij, S.V., Steele, S.L., Razaghi, B., and Berman, J.N. (2017). A rapid and effective method for screening, sequencing and reporter verification of engineered frameshift mutations in zebrafish. Disease models & mechanisms *10*, 811-822.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173-1183.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N.*, et al.* (2011). Metabolic labeling of RNA

uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat Biotech *29*, 436-442.

Radle, B., Rutkowski, A.J., Ruzsics, Z., Friedel, C.C., Koszinowski, U.H., and Dolken, L. (2013). Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. Journal of visualized experiments : JoVE.

Radulovich, N., Pham, N.-A., Strumpf, D., Leung, L., Xie, W., Jurisica, I., and Tsao, M.-S. (2010). Differential roles of cyclin D1 and D3 in pancreatic ductal adenocarcinoma. Molecular Cancer *9*, 1-15.

Raj, A., Rifkin, S.A., Andersen, E., and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. Nature *463*, 913-918.

Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. Cell *135*, 216-226.

Rajshankar, D., Downey, G.P., and McCulloch, C.A. (2012). IL-1β enhances cell adhesion to degraded fibronectin. The FASEB Journal *26*, 4429-4444.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research *42*, W187-191.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nature protocols *8*, 2281-2308.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative Genomics Viewer. Nature biotechnology *29*, 24-26.

Robu, M.E., Larson, J.D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S.A., and Ekker, S.C. (2007). p53 activation by knockdown technologies. PLoS Genet *3*, e78.

Rosenfeld, M.E., Prichard, L., Shiojiri, N., and Fausto, N. (2000). Prevention of hepatic apoptosis and embryonic lethality in RelA/TNFR-1 double knockout mice. The American journal of pathology *156*, 997-1007.

References

Rossi, A., Gauvrit, S., Marass, M., Pan, L., Moens, C.B., and Stainier, D.Y.R. (2016). Regulation of Vegf signaling by natural and synthetic ligands. Blood.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Holper, S., Kruger, M., and Stainier, D.Y. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature *524*, 230-233.

Ruthenburg, A.J., Wang, W., Graybosch, D.M., Li, H., Allis, C.D., Patel, D.J., and Verdine, G.L. (2006). Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. Nature structural & molecular biology *13*, 704-712.

Sage, J., Miller, A.L., Perez-Mancera, P.A., Wysocki, J.M., and Jacks, T. (2003). Acute mutation of retinoblastoma gene function is sufficient for cell cycle re-entry. Nature *424*, 223-228.

Santamaria, D., Barriere, C., Cerqueira, A., Hunt, S., Tardy, C., Newton, K., Caceres, J.F., Dubus, P., Malumbres, M., and Barbacid, M. (2007). Cdk1 is sufficient to drive the mammalian cell cycle. Nature *448*, 811-815.

Sato, H., and Maquat, L.E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta. Genes & development *23*, 2537-2550.

Schaeffer, D., and van Hoof, A. (2011). Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs. Proceedings of the National Academy of Sciences of the United States of America *108*, 2366-2371.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nature methods *9*, 676-682.

Schmid, M., and Jensen, T.H. (2008). The exosome: a multipurpose RNA-decay machine. Trends in biochemical sciences *33*, 501-510.

Schmidt, S.A., Foley, P.L., Jeong, D.H., Rymarquis, L.A., Doyle, F., Tenenbaum, S.A., Belasco, J.G., and Green, P.J. (2015). Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. Nucleic acids research *43*, 309-323.

References

Schmits, R., Filmus, J., Gerwin, N., Senaldi, G., Kiefer, F., Kundig, T., Wakeham, A., Shahinian, A., Catzavelos, C., Rak, J.*, et al.* (1997). CD44 regulates hematopoietic progenitor distribution, granuloma formation, and tumorigenicity. Blood *90*, 2217-2233.

Schuermann, A., Helker, C.S., and Herzog, W. (2015). Metallothionein 2 regulates endothelial cell migration through transcriptional regulation of vegfc expression. Angiogenesis *18*, 463-475.

Schuller, A.P., and Green, R. (2018). Roadblocks and resolutions in eukaryotic translation. Nat Rev Mol Cell Biol *19*, 526-541.

Schwartz, J.C., Younger, S.T., Nguyen, N.B., Hardy, D.B., Monia, B.P., Corey, D.R., and Janowski, B.A. (2008). Antisense transcripts are targets for activating small RNAs. Nature structural & molecular biology *15*, 842-848.

Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., Leon-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S.*, et al.* (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell *159*, 148-162.

Seiberlich, V., Bauer, N.G., Schwarz, L., Ffrench-Constant, C., Goldbaum, O., and Richter-Landsberg, C. (2015). Downregulation of the microtubule associated protein tau impairs process outgrowth and myelin basic protein mRNA transport in oligodendrocytes. Glia *63*, 1621-1635.

Serobyan, V., Kontarakis, Z., El-Brolosy, M.A., Welker, J.M., Tolstenkov, O., Saadeldein, A.M., Retzer, N., Gottschalk, A., Wehman, A.M., and Stainier, D.Y. (2020). Transcriptional adaptation in Caenorhabditis elegans. eLife *9*.

Shatirishvili, M., Burk, A.S., Franz, C.M., Pace, G., Kastilan, T., Breuhahn, K., Hinterseer, E., Dierich, A., Bakiri, L., Wagner, E.F.*, et al.* (2016). Epidermal-specific deletion of CD44 reveals a function in keratinocytes in response to mechanical stress. Cell death & disease *7*, e2461.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annual review of biochemistry *81*, 65-95.

References

Shmidt, T., Hampich, F., Ridders, M., Schultrich, S., Hans, V.H., Tenner, K., Vilianovich, L., Qadri, F., Alenina, N., Hartmann, E.*, et al.* (2007). Normal brain development in importin-alpha5 deficient-mice. Nature cell biology *9*, 1337-1338; author reply 1339.

Shum, E.Y., Jones, S.H., Shao, A., Dumdie, J., Krause, M.D., Chan, W.K., Lou, C.H., Espinoza, J.L., Song, H.W., Phan, M.H.*, et al.* (2016). The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. Cell *165*, 382-395.

Sicinska, E., Aifantis, I., Le Cam, L., Swat, W., Borowski, C., Yu, Q., Ferrando, A.A., Levin, S.D., Geng, Y., von Boehmer, H.*, et al.* (2003). Requirement for cyclin D3 in lymphocyte development and T cell leukemias. Cancer cell *4*, 451-461.

Sicinski, P., and Weinberg, R.A. (1997). A specific role for cyclin D1 in mammary gland development. Journal of mammary gland biology and neoplasia *2*, 335-342.

Silva, A.L., Ribeiro, P., Inacio, A., Liebhaber, S.A., and Romao, L. (2008). Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. Rna *14*, 563-576.

Sims, R.J., 3rd, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends in genetics : TIG *19*, 629-639.

Singh, P., James, R.S., Mee, C.J., and Morozov, I.Y. (2019). mRNA levels are buffered upon knockdown of RNA decay and translation factors via adjustment of transcription rates in human HepG2 cells. RNA biology *16*, 1147-1155.

Smith, J.E., and Baker, K.E. (2015). Nonsense-mediated RNA decay--a switch and dial for regulating gene expression. BioEssays : news and reviews in molecular, cellular and developmental biology *37*, 612-623.

Smits, A.H., Ziebell, F., Joberty, G., Zinn, N., Mueller, W.F., Clauder-Munster, S., Eberhard, D., Falth Savitski, M., Grandi, P., Jakob, P.*, et al.* (2019). Biological plasticity rescues target activity in CRISPR knock outs. Nature methods *16*, 1087-1093.

References

Speers, C., Zhao, S.G., Kothari, V., Santola, A., Liu, M., Wilder-Romans, K., Evans, J., Batra, N., Bartelink, H., Hayes, D.F.*, et al.* (2016). Maternal Embryonic Leucine Zipper Kinase (MELK) as a Novel Mediator and Biomarker of Radioresistance in Human Breast Cancer. Clin Cancer Res *22*, 5864-5875.

Spingola, M., Grate, L., Haussler, D., and Ares, M., Jr. (1999). Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. Rna *5*, 221-234.

Stuckenholz, C., Kageyama, Y., and Kuroda, M.I. (1999). Guilt by association: non-coding RNAs, chromosome-specific proteins and dosage compensation in Drosophila. Trends in genetics : TIG *15*, 454-458.

Su, Z., Wang, J., and Gu, X. (2014). Effect of duplicate genes on mouse genetic robustness: an update. BioMed research international *2014*, 758672.

Sulem, P., Helgason, H., Oddson, A., Stefansson, H., Gudjonsson, S.A., Zink, F., Hjartarson, E., Sigurdsson, G.T., Jonasdottir, A., Jonasdottir, A.*, et al.* (2015). Identification of a large set of rare complete human knockouts. Nature genetics *47*, 448-452.

Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Lariviere, L., Maier, K.C., Seizl, M., Tresch, A., and Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. Genome research *22*, 1350-1359.

Sun, W., and Chen, W. (2018). Metabolic Labeling of Newly Synthesized RNA with 4sU to in Parallel Assess RNA Transcription and Decay. Methods in molecular biology (Clifton, NJ) *1720*, 25-34.

Sztal, T.E., McKaige, E.A., Williams, C., Ruparelia, A.A., and Bryson-Richardson, R.J. (2018). Genetic compensation triggered by actin mutation prevents the muscle damage caused by loss of actin protein. PLoS Genet *14*, e1007212.

Takahashi, K., and Kaneko, I. (1985). Changes in nuclease sensitivity of mammalian cells after irradiation with 60Co gamma-rays. International journal of radiation biology and related studies in physics, chemistry, and medicine *48*, 389-395.

References

Tange, T.O., Nott, A., and Moore, M.J. (2004). The ever-increasing complexities of the exon junction complex. Current opinion in cell biology *16*, 279-284.

Tautz, D. (1992). Redundancies, development and the flow of information. BioEssays : news and reviews in molecular, cellular and developmental biology *14*, 263-266.

Teng, X., Dayhoff-Brannigan, M., Cheng, W.C., Gilbert, C.E., Sing, C.N., Diny, N.L., Wheelan, S.J., Dunham, M.J., Boeke, J.D., Pineda, F.J.*, et al.* (2013). Genome-wide consequences of deleting any single gene. Molecular cell *52*, 485-494.

Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nature methods *12*, 1143-1149.

Theodosiou, M., Widmaier, M., Bottcher, R.T., Rognoni, E., Veelders, M., Bharadwaj, M., Lambacher, A., Austen, K., Muller, D.J., Zent, R.*, et al.* (2016). Kindlin-2 cooperates with talin to activate integrins and induces cell spreading by directly binding paxillin. eLife *5*, e10130.

Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeier, C., Hentze, M.W., and Kulozik, A.E. (1998). Binary specification of nonsense codons by splicing and cytoplasmic translation. The EMBO journal *17*, 3484-3494.

Tondeleir, D., Lambrechts, A., Muller, M., Jonckheere, V., Doll, T., Vandamme, D., Bakkali, K., Waterschoot, D., Lemaistre, M., Debeir, O.*, et al.* (2012). Cells lacking beta-actin are genetically reprogrammed and maintain conditional migratory capacity. Molecular & cellular proteomics : MCP *11*, 255-271.

Townsley, F.M., Frigerio, G., and Pelham, H.R. (1994). Retrieval of HDEL proteins is required for growth of yeast cells. J Cell Biol *127*, 21-28.

Trieu, M., Ma, A., Eng, S.R., Fedtsova, N., and Turner, E.E. (2003). Direct autoregulation and gene dosage compensation by POU-domain transcription factor Brn3a. Development (Cambridge, England) *130*, 111-121.

References

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. Molecular cell *46*, 518-529.

Tufarelli, C., Stanley, J.A., Garrick, D., Sharpe, J.A., Ayyub, H., Wood, W.G., and Higgs, D.R. (2003). Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. Nature genetics *34*, 157-165.

Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., Barrett, Q., Herbert, J., Mathews, D.H., Kim, J.*, et al.* (2019). CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. Nature communications *10*, 4056.

Turunen, M.P., Lehtola, T., Heinonen, S.E., Assefa, G.S., Korpisalo, P., Girnary, R., Glass, C.K., Vaisanen, S., and Yla-Herttuala, S. (2009). Efficient regulation of VEGF expression by promoter-targeted lentiviral shRNAs based on epigenetic mechanism: a novel example of epigenetherapy. Circulation research *105*, 604-609.

Ueno, D., Yamasaki, S., Demura, T., and Kato, K. (2018). Comprehensive analysis of mRNA internal cleavage sites in Arabidopsis thaliana. Journal of bioscience and bioengineering *125*, 723-728.

Unterholzner, L., and Izaurralde, E. (2004). SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Molecular cell *16*, 587-596.

Valen, E., Preker, P., Andersen, P.R., Zhao, X., Chen, Y., Ender, C., Dueck, A., Meister, G., Sandelin, A., and Jensen, T.H. (2011). Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes. Nature structural & molecular biology *18*, 1075-1082.

van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science (New York, NY) *295*, 2262-2264.

Vance, K.W., and Ponting, C.P. (2014). Transcriptional regulatory functions of nuclear long noncoding RNAs. Trends in genetics : TIG *30*, 348-355.

References

Vejnar, C.E., Moreno-Mateos, M.A., Cifuentes, D., Bazzini, A.A., and Giraldez, A.J. (2016). Optimized CRISPR-Cas9 System for Genome Editing in Zebrafish. Cold Spring Harbor protocols *2016*, pdb.prot086850.

von Koch, C.S., Zheng, H., Chen, H., Trumbauer, M., Thinakaran, G., van der Ploeg, L.H., Price, D.L., and Sisodia, S.S. (1997). Generation of APLP2 KO mice and early postnatal lethality in APLP2/APP double KO mice. Neurobiology of aging *18*, 661-669.

Waddington, C.H. (1959). Canalization of Development and Genetic Assimilation of Acquired Characters. Nature *183*, 1654-1655.

Wagner, K.D., Wagner, N., Ghanbarian, H., Grandjean, V., Gounon, P., Cuzin, F., and Rassoulzadegan, M. (2008). RNA induction and inheritance of epigenetic cardiac hypertrophy in the mouse. Developmental cell *14*, 962-969.

Wall, J.D., Stawiski, E.W., Ratan, A., Kim, H.L., Kim, C., Gupta, R., Suryamohan, K., Gusareva, E.S., Purbojati, R.W., Bhangale, T.*, et al.* (2019). The GenomeAsia 100K Project enables genetic discoveries across Asia. Nature *576*, 106-111.

Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Borchers, C., Tempst, P., and Zhang, Y. (2001). Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. Molecular cell *8*, 1207-1217.

Wang, J., Hamilton, J.I., Carter, M.S., Li, S., and Wilkinson, M.F. (2002). Alternatively spliced TCR mRNA induced by disruption of reading frame. Science (New York, NY) *297*, 108-110.

Wang, J., Wang, Q., Cui, Y., Liu, Z.Y., Zhao, W., Wang, C.L., Dong, Y., Hou, L., Hu, G., Luo, C.*, et al.* (2012). Knockdown of cyclin D1 inhibits proliferation, induces apoptosis, and attenuates the invasive capacity of human glioblastoma cells. Journal of neuro-oncology *106*, 473-484.

Wang, R.H., Li, C., and Deng, C.X. (2010). Liver steatosis and increased ChREBP expression in mice carrying a liver specific SIRT1 null mutation under a normal feeding condition. International journal of biological sciences *6*, 682-690.

References

Wang, Y., Lee, Y.M., Baitsch, L., Huang, A., Xiang, Y., Tong, H., Lako, A., Von, T., Choi, C., Lim, E.*, et al.* (2014). MELK is an oncogenic kinase essential for mitotic progression in basal-like breast cancer cells. eLife *3*, e01763.

Wang, Y., Schnegelsberg, P.N., Dausman, J., and Jaenisch, R. (1996). Functional redundancy of the muscle-specific transcription factors Myf5 and myogenin. Nature *379*, 823-825.

Weake, V.M., and Workman, J.L. (2008). Histone ubiquitination: triggering gene activity. Molecular cell *29*, 653-663.

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.G., and Qi, Y. (2012). A role for small RNAs in DNA double-strand break repair. Cell *149*, 101-112.

Wei, Z., Huang, D., Gao, F., Chang, W.H., An, W., Coetzee, G.A., Wang, K., and Lu, W. (2013). Biological implications and regulatory mechanisms of long-range chromosomal interactions. The Journal of biological chemistry *288*, 22369-22377.

Wery, M., Gautier, C., Descrimes, M., Yoda, M., Vennin-Rendos, H., Migeot, V., Gautheret, D., Hermand, D., and Morillon, A. (2018). Native elongating transcript sequencing reveals global anti-correlation between sense and antisense nascent transcription in fission yeast. Rna *24*, 196-208.

Westerfield, M. (2000). The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio), 4th Edition, 4 edn (University of Oregon Press, Eugene).

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M.*, et al.* (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467-481.

White, J.K., Gerdin, A.K., Karp, N.A., Ryder, E., Buljan, M., Bussell, J.N., Salisbury, J., Clare, S., Ingham, N.J., Podrini, C.*, et al.* (2013). Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell *154*, 452-464.

Williams, J.A., Ni, H.M., Haynes, A., Manley, S., Li, Y., Jaeschke, H., and Ding, W.X. (2015). Chronic Deletion and Acute Knockdown of Parkin Have Differential

References

Responses to Acetaminophen-induced Mitophagy and Liver Injury in Mice. The Journal of biological chemistry *290*, 10934-10946.

Wintersinger, J.A., and Wasmuth, J.D. (2015). Kablammo: an interactive, web-based BLAST results visualizer. Bioinformatics *31*, 1305-1306.

Wolin, S.L., and Maquat, L.E. (2019). Cellular RNA surveillance in health and disease. Science (New York, NY) *366*, 822-827.

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet *18*, 517-534.

Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell *121*, 859-872.

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C.Y., and Wei, L. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic acids research *39*, W316-322.

Yamamoto, S., Jaiswal, M., Charng, W.L., Gambin, T., Karaca, E., Mirzaa, G., Wiszniewski, W., Sandoval, H., Haelterman, N.A., Xiong, B.*, et al.* (2014). A drosophila genetic resource of mutants to study mechanisms underlying human genetic diseases. Cell *159*, 200-214.

Yamashita, A., Chang, T.C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.Y., and Shyu, A.B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nature structural & molecular biology *12*, 1054-1063.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T.*, et al.* (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes & development *23*, 1091-1105.

References

Yang, Y., Li, T., Vu, T.H., Ulaner, G.A., Hu, J.F., and Hoffman, A.R. (2003). The histone code regulating expression of the imprinted mouse Igf2r gene. Endocrinology *144*, 5658-5670.

Yasuhara, N., Shibazaki, N., Tanaka, S., Nagai, M., Kamikawa, Y., Oe, S., Asally, M., Kamachi, Y., Kondoh, H., and Yoneda, Y. (2007). Triggering neural differentiation of ES cells by subtype switching of importin-alpha. Nature cell biology *9*, 72-79.

Ye, D., Wang, X., Wei, C., He, M., Wang, H., Wang, Y., Zhu, Z., and Sun, Y. (2019). Marcksb plays a key role in the secretory pathway of zebrafish Bmp2b. PLoS Genet *15*, e1008306.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R.*, et al.* (2003). Widespread occurrence of antisense transcription in the human genome. Nature biotechnology *21*, 379-386.

Young, R., Passet, B., Vilotte, M., Cribiu, E.P., Beringue, V., Le Provost, F., Laude, H., and Vilotte, J.L. (2009). The prion or the related Shadoo protein is required for early mouse embryogenesis. FEBS letters *583*, 3296-3300.

Yu, W., Gius, D., Onyango, P., Muldoon-Jacobs, K., Karp, J., Feinberg, A.P., and Cui, H. (2008). Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature *451*, 202-206.

Zhang, X., Li, H., Burnett, J.C., and Rossi, J.J. (2014). The role of antisense long noncoding RNA in small RNA-triggered gene activation. RNA *20*, 1916-1928.

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. Journal of computational biology : a journal of computational molecular cell biology *7*, 203-214.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.B. (2015a). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature *526*, 591-594.

Zhou, Q., Yang, D., Ombrello, A.K., Zavialov, A.V., Toro, C., Zavialov, A.V., Stone, D.L., Chae, J.J., Rosenzweig, S.D., Bishop, K.*, et al.* (2014). Early-onset stroke and

vasculopathy associated with mutations in ADA2. The New England journal of medicine *370*, 911-920.

Zhou, S., Xie, Y., Tang, J., Huang, J., Huang, Q., Xu, W., Wang, Z., Luo, F., Wang, Q., Chen, H.*, et al.* (2015b). FGFR3 Deficiency Causes Multiple Chondroma-like Lesions by Upregulating Hedgehog Signaling. PLoS Genet *11*, e1005214.

Zhu, P., Ma, Z., Guo, L., Zhang, W., Zhang, Q., Zhao, T., Jiang, K., Peng, J., and Chen, J. (2017). Short body length phenotype is compensated by the upregulation of nidogen family members in a deleterious nid1a mutation of zebrafish. Journal of genetics and genomics = Yi chuan xue bao *44*, 553-556.

Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D.C., Lukas, J., Bekker-Jensen, S., Bartek, J., and Shiloh, Y. (2006). Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. Nature cell biology *8*, 870-876.

11. Acknowledgement

During my Ph.D. studies, I was influenced by several wonderful people who had a great impact on my scientific and personal development throughout those years.

Foremost, I would like to thank my supervisor, Prof. Dr. Didier Stainier, for being the best mentor a graduate student can wish for. I thank Didier for putting his trust on me to pursue this project and providing me with the perfect balance of guidance and freedom. Didier has supported me in several ways to become a better scientist; he always provided me with some of his time whenever I needed guidance, he allowed me to attend several scientific conferences to gain more knowledge and establish connections for the future, he always provided me with feedback to improve my scientific writing and presentation skills, gave me the opportunity to supervise students in the lab to develop my management skills and most importantly, taught me how to properly think and build-up a scientific project. I thank him as well for always being very caring about my future career. The support I received/continue to receive from him is tremendous; something I will be forever grateful for.

I also thank my thesis advisory committee members, Prof Dr. Dr. Thomas Braun, Prof. Dr. Ingo Ebersberger and Prof. Dr. Erin Schuman for all of their very useful input and advice throughout my Ph.D. and for taking some of their precious time to meet me annually. I also thank Prof. Dr. Erin Schuman for reviewing my thesis and Kenny Mattonet for translating the thesis summary (Zusammenfassung).

I would like to express my gratitude to the entire transcriptional adaptation team of the Stainier lab, especially for all of the amazing discussions we had at our monthly brainstorming meetings. I foremost thank Andrea Rossi and Zacharias Kontarakis for the great collaboration we had, their initial publication on genetic compensation has been the inspiration of my studies and working with them had been very motivating and intellectually stimulating. I also thank Vahan Serobyan, Jane Jiang, and other members of the team who joined later: Gabrielius Jakutis, Jordan Welker, Guilherme Valente and Preethi Krishnaraj for all of their useful input throughout my Ph.D..

I would also like to thank all collaborators I worked with during my studies. From the Max Planck institute for heart and lung research, I thank Carsten Kuenne for all of his help with the bioinformatic analyses, Stefan Günther for his help with next generation sequencing, Ann Atzberger and Khrievono Kikhi for their help with cell sorting, Nana

Acknowledgement

Fukuda, Giulia Boezio, Ben Lai, Ryuichi Fukuda, Jenny Pestel and Claudia Gerri for helping me with some experiments and providing me with some mutant zebrafish. I am also grateful to Dr. Carter Takacs and Prof. Dr. Antonio Giraldez (Yale University, New Haven, CT, USA), Reinhard Fässler (Max Planck institute for biochemistry, Münster, Germany) and Prof. Dr. Alexander Hoffmann (UCLA, Los Angeles, CA, USA) for providing me with mutant zebrafish and mouse cell lines that were useful for my work.

I am very fortunate to have received tremendous help from a great support team at the Stainier lab. I start with Sharon Meaney-Gardian who has been supporting me since my first day in Bad Nauheim, I am very grateful for her help in getting me settled in Bad Nauheim, her enormous support to the entire lab not only scientifically but also with all of the bureaucratic procedure, and ensuring we have an optimal environment to focus on our research, in addition to her personal support for my career and future. I thank Sabine Fischer and Simon Perathoner for their help with animal proposals and documentations. I am also very grateful for the entire technician team at the lab, Carmen Büttner, Carmen Kremser, Marianne Ploch, Petra Neeb, Hans-Martin Maischein, Simon Howard, Beate Grohmann, Khrievono Kikhi, Rebecca Lee, Monika Endl, Arleta Karczewska and Nana Fukuda for all of their support. I also thank Rita Retzloff, Nouha Ritschel, Jennifer Behm, Monika Mueller-Boche, Martin Laszczyk and the entire zebrafish and mouse facilities for their help maintaining our animals in good health.

I am also grateful to many other current and former Ph.D. students and postdocs of the lab. I thank Oliver Stone and Sven Reischauer for having taught me handling zebrafish when I first joined the lab as an undergrad intern, for being great friends and for constantly providing me with advice. I also thank Aly Villasenor, Sebastian Guavrit, Anna Sokol, Jason Lai, Claudia Gerri, Michele Marass, Michelle Collins, Arica Beisaw, Anabela Bensimon Brito, Hyouk-Bum Kwon, Hyun-Taek Kim, Paolo Panza, Rashmi Priya, Felix Gunawan, Brian Njaine, Ayele Taddese-Tsedeke, Veronica Uribe Sokolov, Ryota Matsouka and Radhan Ramadass for all of the intellectual discussions, support and being great lab partners. I also enjoyed supervising Mahmoud El-Sakka during his summer internship in 2017 and would like to thank him for his help. I thank my friends and lab partners Javad Rasouli, Sri Teja Mullapudi, Giulia Boezio, Hadil El-Sammak, Claudia Carlantoni, Alessandra Gentile, Gabrielius Jakutis, Srinath Ramkumar, Kenny Mattonet and Srinivas Allanki for

making being at the lab always feel like home, for truly being brothers and sisters and sharing a lot of fun moments together inside and outside of the lab.

I also thank my undergrad mentor Prof. Dr. Ahmed Ihab Abdelaziz for his continuous support and advice. I thank my friends from home: Omar Mossad, Mohamed El-Maraghi, Khaled Hany, Reem Atef, Khaled Osama, Mahmoud Khalifa, Amr AbdelMoniem, Ziad Amr and Ali Seliet for their support and bringing the home feeling whenever we talked or met. I also thank Marie van Vliet for having been a very supportive partner during my Ph.D., her support, motivation, and love have eased many of the challenging moments during my studies.

I am also thankful to the Boehringer Ingelheim Fonds (BIF) for supporting my research and all of the BIF fellows who provided me with great feedback and with whom I shared several intellectual discussions during the summer retreats in Hirschegg, Austria. I also thank Bilge Reischauer for coordinating our graduate school and providing us with a lot of support throughout our studies.

My Karate team, KD Dietzenbach, has also been an important part of my Ph.D., a place where I released some pressure and enjoyed doing something I care about besides science. I am thankful to my coaches and all of the team members for the inspiration, fun and glory moments. I am also thankful to members of the 2018 Max Planck PhDnet in general, and the steering group members in particular, with whom I enjoyed working on several science policy issues that interest me.

Finally, a special thank you goes to my family, to whom I dedicate this dissertation. Despite the distance, they have been always an integral part of my journey. I am deeply thankful to my parents, Hoda El-Shaarawy and Ahmed El-Brolosy, for boosting my love to science while raising me, for the sacrifices they made to support my studies, even if it meant that I will be no longer living close to them, and for being a constant source of inspiration, unconditional love and encouragement. I am also very lucky to have amazing siblings: Mostafa, Ghada and Yousef. Their support, understanding and love mean a lot to me and fill my heart with gratitude.

12. Appendix

12.1. Sequence alignment of the coding sequences of zebrafish *actb1* and mouse *Actb*.

Sequence alignment of the coding sequences of zebrafish *actb1* mRNA (ENSDART00000054987; Query) and mouse gene *Actb* mRNA (ENSMUST00000100497; Subject) using MUSCLE (Edgar, 2004). Shared sequence identity counts to 88% over 1128 nucleotides.

```
CLUSTAL multiple sequence alignment by MUSCLE (3.8)
Query      ATGGATGAGGAAATCGCTGCCCTGGTCGTTGACAACGGCTCCGGTATGTGCAAAGCCGGT
Subject    ATGGATGACGATATCGCTGCGCTGGTCGTCGACAACGGCTCCGGCATGTGCAAAGCCGGC
           ******** ** ******** ******** ************** **************

Query      TTTGCTGGAGATGATGCCCCTCGTGCTGTTTTCCCCTCCATTGTTGGACGACCCAGACAT
Subject    TTCGCGGGCGACGATGCTCCCCGGGCTGTATTCCCCTCCATCGTGGGCCGCCCTAGGCAC
           ** ** ** ** ***** ** ** ***** *********** ** ** ** ** ** **

Query      CAGGGAGTGATGGTTGGCATGGGACAGAAAGACTCCTATGTGGGAGATGAGGCCCAGAGC
Subject    CAGGGTGTGATGGTGGGAATGGGTCAGAAGGACTCCTATGTGGGTGACGAGGCCCAGAGC
           ***** ******** ** ***** ***** ************** ** ************

Query      AAGAGAGGTATCCTGACCCTCAAATACCCCATTGAGCACGGTATTGTGACTAACTGGGAT
Subject    AAGAGAGGTATCCTGACCCTGAAGTACCCCATTGAACATGGCATTGTTACCAACTGGGAC
           ******************** ** *********** ** ** ***** ** ********

Query      GACATGGAGAAGATCTGGCATCACACCTTCTACAATGAGCTCCGTGTTGCCCCTGAGGAG
Subject    GACATGGAGAAGATCTGGCACCACACCTTCTACAATGAGCTGCGTGTGGCCCCTGAGGAG
           ******************** ******************** ***** ************

Query      CACCCTGTGCTGCTCACTGAGGCTCCCCTGAATCCCAAAGCCAACAGAGAGAAGATGACA
Subject    CACCCTGTGCTGCTCACCGAGGCCCCCCTGAACCCTAAGGCCAACCGTGAAAAGATGACC
           ****************** ***** ******** ** ** ****** * ** ********

Query      CAGATCATGTTCGAGACCTTCAACACCCCTGCCATGTATGTGGCCATCCAGGCTGTGCTC
Subject    CAGATCATGTTTGAGACCTTCAACACCCCAGCCATGTACGTAGCCATCCAGGCTGTGCTG
           *********** ***************** ******** ** ******************

Query      TCTCTGTACGCTTCTGGTCGTACTACTGGTATTGTGATGGACTCTGGTGATGGTGTGACC
Subject    TCCCTGTATGCCTCTGGTCGTACCACAGGCATTGTGATGGACTCCGGAGACGGGGTCACC
           ** ***** ** *********** ** ** ************** ** ** ** ** ***

Query      CACACCGTGCCCATCTATGAGGGTTACGCTCTTCCCCATGCCATCCTGCGTCTGGATCTG
Subject    CACACTGTGCCCATCTACGAGGGCTATGCTCTCCCTCACGCCATCCTGCGTCTGGACCTG
           ***** *********** ***** ** ***** ** ** ***************** ***

Query      GCTGGTCGTGACCTGACAGACTACCTGATGAAGATCCTGACCGAGCGTGGCTACAGCTTC
Subject    GCTGGCCGGGACCTGACAGACTACCTCATGAAGATCCTGACCGAGCGTGGCTACAGCTTC
           ***** ** ***************** **********************************

Query      ACCACCACGGCCGAAAGAGAAATTGTCCGTGACATCAAGGAGAAGCTGTGCTACGTGGCC
Subject    ACCACCACAGCTGAGAGGGAAATCGTGCGTGACATCAAAGAGAAGCTGTGCTATGTTGCT
           ******** ** ** ** ***** ** *********** ************** ** **

Query      CTGGACTTTGAGCAGGAGATGGGAACCGCTGCCTCCTCTTCCTCCCTGGAGAAGAGCTAT
Subject    CTAGACTTCGAGCAGGAGATGGCCACTGCCGCATCCTCTTCCTCCCTGGAGAAGAGCTAT
           ** ***** *************  ** ** ** ****************************

Query      GAGCTGCCTGACGGTCAGGTCATCACCATTGGCAATGAGCGTTTCCGTTGCCCCGAGGCT
Subject    GAGCTGCCTGACGGCCAGGTCATCACTATTGGCAACGAGCGGTTCCGATGCCCTGAGGCT
           ************** *********** ******** ***** ***** ***** ******

Query      CTCTTCCAGCCTTCCTTCCTGGGTATGGAATCTTGCGGTATCCATGAGACCACCTTCAAC
Subject    CTTTTCCAGCCTTCCTTCTTGGGTATGGAATCCTGTGGCATCCATGAAACTACATTCAAT
           ** *************** ************* ** ** ******** ** ** *****

Query      TCCATCATGAAGTGCGACGTGGACATCCGTAAGGACCTGTATGCCAACACAGTGCTGTCT
Subject    TCCATCATGAAGTGTGACGTTGACATCCGTAAAGACCTCTATGCCAACACAGTGCTGTCT
           ************** ***** ************ ***** **********************
```

```
Query      GGAGGTACCACCATGTACCCTGGCATTGCTGACCGTATGCAGAAGGAGATCACCTCTCTT
Subject    GGTGGTACCACCATGTACCCAGGCATTGCTGACAGGATGCAGAAGGAGATTACTGCTCTG
           ** ***************** ************ * ************** **  ****

Query      GCTCCTTCCACCATGAAGATCAAGATCATTGCTCCCCCTGAGCGTAAATACTCCGTCTGG
Subject    GCTCCTAGCACCATGAAGATCAAGATCATTGCTCCTCCTGAGCGCAAGTACTCTGTGTGG
           ******  ************************** ******** ** ***** ** ***

Query      ATCGGTGGCTCCATCTTGGCCTCCCTGTCCACCTTCCAGCAGATGTGGATCAGCAAGCAG
Subject    ATCGGTGGCTCCATCCTGGCCTCACTGTCCACCTTCCAGCAGATGTGGATCAGCAAGCAG
           *************** ******* ************************************

Query      GAGTACGATGAGTCTGGCCCATCCATCGTCCACAGGAAGTGCTTCTAA
Subject    GAGTACGATGAGTCCGGCCCCTCCATCGTGCACCGCAAGTGCTTCTAG
           ************** ***** ******** *** * ***********
```

12.2. Sequence alignment of *hif1ab* mRNA and the synthetic transcript containing only sequences that exhibit similarity with *epas1a* locus.

Alignment of the synthetic *hif1ab* transcript composed solely of sequences exhibiting sequence similarity to *epas1a* genomic locus (Query) and *hif1ab* mRNA (ENSDART00000018500; Subject) using MUSCLE (Edgar, 2004).

```
CLUSTAL multiple sequence alignment by MUSCLE (3.8)

Query      ------------------------------------------------------------
Subject    GGACTGGTGTGCCATCAGTTTTGTGATGGGTTTCGAAGTGCGACACCGCTGTCTGAATGC


Query      -------------AGAGCCTCGTGAGCGCG------------------------------
Subject    GCTCGCTCCCGTCAGAGCCTCGTGAGCGCGCGAGGCACAGTGAGAGTAGTCAGAGCGCAG
                      *****************

Query      ------------------------------------------------------CAAATGATT
Subject    GCACGGTTTAACTCTTTGTACCGACCAGAAGCTGCACCGACCGGTGCCAAACAAATGATT
                                                                *********

Query      TGACCTT-------------------TCTTTACCATTCTGA--------------------
Subject    TGACCTTCAGCGAACCGGAGGGTTGATCTTTACCATTCTGACACTTTCAACATACCACAC
           *******                   ***************

Query      ------------------------------------------------------TTTTTA
Subject    GAGACCTGACCCGGATTATTAGTTCGCGTCTGGCCGTTTTCTTTAATAAACGCGTTTTTA
                                                                  ******

Query      TTTGTT-------------------------------------------------------
Subject    TTTGTTAGCAAAAGGTGGCCACGTCCTTGATTTTCTGGCCTCGCGGTTCGGAGAAAACCT
           ******

Query      ----------------------------------------TTGTCACTGAAAAGAAAGG
Subject    AACACATACTGAGTGGTTTCACCCAGGAATGGATACTGGAGTTGTCACTGAAAAGAAAAG
                                                  ***************** *

Query      GTGAGCTCGGAGCGCAGGAAGGAGAAGTCCAGGGATGCAGCGCGATCTCGCAGGGGAAAG
Subject    GTGAGCTCGGAGCGCAGGAAGGAGAAGTCCAGGGATGCAGCGCGATCTCGCAGGGGAAAG
           ************************************************************

Query      GAGTCTGAGGTGTTCTACGAGTTAGCACACCAGCTCCCCCTGCCACACAATGTCACGTCT
Subject    GAGTCTGAGGTGTTCTACGAGTTAGCACACCAGCTCCCCCTGCCACACAATGTCACGTCT
           ************************************************************

Query      CACCTGGACAAAGCCTCCATTATGAGGCTCACCATCAGCTACCTGCGC------------
Subject    CACCTGGACAAAGCCTCCATTATGAGGCTCACCATCAGCTACCTGCGCATGAGGAAGCTG
           *********************************************

Query      ------------------------AGGAGGAGAATGAGC----------------------
Subject    CTCAATTCCATGAAAAAGAGGAGAAGGAGGAGAATGAGCTGGAAAGTCAGCTGAATGGCT
                                  ***************

Query      ------------------------------------------------------------
Subject    TTTATCTGAAGGCCCTTGAGGGTTTCCTTATGGTCCTGTCTGAGGATGGAGACATGGTTT
```

Appendix

```
Query      ------------------------------------------GATCTGACCGGTCACAG
Subject    ATCTCTCTGAGAATGTCAGCAAGAGCATGGGCCTCACACATTTGATCTGACCGGTCACAG
                                                    *****************

Query      CATCTTTGAATTTTCACACCCATGTGACCATGA----------------------------
Subject    CATCTTTGAATTTTCACACCCATGTGACCATGAAGAGTTGAGAGAGATGCTCGTCCACAG
           *******************************

Query      ------------------------------------------------------------
Subject    AACAGATCCAAAAAGACCAAGGAACAAAACACAGAGCGTAGCTTCTTCCTGCGGATGAAG


Query      ------------------------------------------------------------
Subject    TGCACACTCACTAGCAGAGGACGCACCGTCAATATCAAGTCTGCGACGTGGAAGTTCTTC


Query      ----------------------------------------CTTCAGAAGACTCTG------
Subject    ACTGCGCTGGTCATGTTCGTGTGCACGAGGGCAGCGAGGCTTCAGAAGACTCTGGCTTTA
                                                   ***************

Query      ---------------------------TCATCTGTGAGCCCAT------------------
Subject    AAGAGCCCCCTGTCACTTACCTTGTGCTCATCTGTGAGCCCATTCCTCATCCCTCGAACA
                                     ****************

Query      ----------CATTGGACAGCAAGACCTTCCTCAGCCGCCACACTCTGGACATGAAGTTCT
Subject    TCGAGGTGCCATTGGACAGCAAGACCTTCCTCAGCCGCCACACTCTGGACATGAAGTTCT
             ************************************************

Query      CATACTGTGAT--------------------------------CAGATGACCTC--------
Subject    CATACTGTGATGAAAGATCACTGAGCTGATGGGATATGAGCCAGATGACCTCTTGAACAG
           ***********                                ***********

Query      -------------------------------TTCAGATCACCTCACCAAG-----------
Subject    ATCAGTCTACGAGTACTATCACGCCCTTGATTCAGATCACCTCACCAAGACACATCACAA
                                          ******************

Query      --------------------CACCACAGGCCAGTACCGCATGCTGGCTAAGAAAGGTGGT
Subject    CTGTTTGCAAAGGGCCAGGCCACCACAGGCCAGTACCGCATGCTGGCTAAGAAAGGTGGT
                              ***************************************

Query      TTTGTGTGGGTTGAGACTCAGGCCACTGTAATCTACAACCCCAAGAATTCTCAGCCGCAA
Subject    TTTGTGTGGGTTGAGACTCAGGCCACTGTAATCTACAACCCCAAGAATTCTCAGCCGCAA
           ************************************************************

Query      TGCATTGTGTGCGTCAACTACGTTCTCA--------------------------------
Subject    TGCATTGTGTGCGTCAACTACGTTCTCATGGCATTGTAGAGGGGGATGTAGTCCTGTCCT
           ****************************

Query      ------------------------------------------------------------
Subject    TGCAGCAGACCGTGACTGAGCCCAAGGCTGTTGAGAAAGAAAGTGAGGAGACTGAGGAAA


Query      ------------------------------------------------------------
Subject    AGACCTCTGAATTGGATATTCTCAAGCTCTTCAAGCCAGAAAGCCTCAATTGCTCATTGG


Query      -----TCTACACTTTATAA-------------------------------------------
Subject    AAAGCTCTACACTTTATAATAAGCTGAAAGAGGAGCCAGAGGCCCTCACTGTGTTGGCAC
                ***************

Query      ------------------------------------------------------------
Subject    CTGCAGCAGGCGACGCCATTATCTCTCTGGACTTCAACAACTCAATTCTGACATACAGCT


Query      -------------------------------TCATGCTGCCTTC-----------------
Subject    GCTGAAGGAGGTGCCCCTCTACAATGATGTCATGCTGCCTTCCAGCAGTGAGAAGCTGCC
                                           *************

Query      ------------------------------------------------------------
Subject    ACTCAGCCTATCTCCTCTCACACCCAGCGACTCCATCCCAGCTCTGACCAAACTAGAGAC


Query      ------------------------GCTCTGCCTCTGATCG----------------------
Subject    TGGAGGAGAGGACTTCCCTTTCAGCTCTGCCTCTGATCGTGTGCCAGACCCCACAAACAC
                                  ****************

Query      ----------------------------------------CATGGATTACGG---CCCA
```

Appendix

```
Subject    ACCCTCCACATCTGGACTCGGCTCTTCGGGCCCAACAGCCCCATGGATTACGGTTTCCCA
                                                ************   ****

Query      G-------------------------------------------------------------
Subject    GTGGAACCAGACATCAGTTCTGAATTTAAACTCGACCTGGTTGAGAAACTGTTTGCTATT
           *

Query      ---------------------------------CCCATGGAGGATCTCGACCTAGAGATGC
Subject    GATACCGAAGCAAAGACACCTTTTTCCACCCACCCATGGAGGATCTCGACCTAGAGATGC
                                          ****************************

Query      TGGCTCCTTACATCCCAATGGATGACGACTTCCAGCTG----------------------
Subject    TGGCTCCTTACATCCCAATGGATGACGACTTCCAGCTGCGCATCCCATCTCCACTGGATC
           ************************************

Query      -------------------------------------------------------------
Subject    CGCTCCCATCTGCCACTCACTCTGTGTCAGCCATGAGCTCTTTATTCCAACCCTTACCCT


Query      ---------------CAGCCTCATCTAC--------------------------------
Subject    CCTCGCCAGCATCTCCAGCCTCATCTACCAGCAGCACAGTGAAGCAGGAGGCGTCATCCC
                       *************

Query      ----------------------GCTGCAGGAG--GTGCAGTGCACCTGTCTCGC------
Subject    GGGCCCCTTCACCCCTACACCTGCTGCAGGAGGTGTGCAGTGCACCTGTCTCGCCCTTCA
                               **********  ********************

Query      -------------------------------------------------------------
Subject    GTGGCAGTCGGGATGCCTCACCTGTTCGATCCAGCACCCCACAGAGCAGCAGTCAGCTCA


Query      -------------------AAGATGTTAG--------------------------------
Subject    ACAACAAGAAATGTCTCCAAAGATGTTAGCCTTCCAAAATATCCAGCGTAAGAGGAAGCT
                            **********

Query      -------------------------------------------------------------
Subject    AAACGAAGTGACGTCGCTTTCTGAAGCTGTTGGACTGGGGCTTTGCTTCACAGTGTGGAC


Query      -----------------------------------------------------CCAGTGTGCTC
Subject    AGTGCTATAGACCCTGGAAAGAGAGCGAAGGTTTTAGAGGTGAAAGGGTCCAGTGTGCTC
                                                              ***********

Query      GGG----------------------------------------------------------
Subject    GGGGGAAACAAAACAATTCTCATACTGCCCTCTATGTGGCCAGTCGTCTGTTGAGCAGTT
           ***

Query      -------------------------------------------------------------
Subject    CTCTAGAGGGCAGCGGCGGCCTCCCTCAGCTCACACGCTACGACTGCGAAGTCAACGCTC


Query      -------------------------------------------------------------
Subject    CCGTGCAGGACCGCCACCATCTGCTCCAGGGAGAGGAGCTGCTGCGTGCTCTGGACCAAG


Query      ----------------------TAGCAACTGCGGACACTC----------------------
Subject    TCAACTGAGCTCTGCTGCTTTTAGCAACTGCGGACACTCTTCCCCCCTTCTCCCGAACCC
                              ******************

Query      ---------------------------------------------TATTCTCTAAAAACCCC
Subject    CTGAAATCTGCCTCCACTTGTCTCCTTTTATCCCAAGCCCTAATATTCTCTAAAAACCCC
                                                        *****************

Query      AG------------------ACTAGACCTGCA-----------------------------
Subject    AGTGTTTTAAAACGAAATGTACTAGACCTGCACCTTCCATTGATGTGAACAGGAGCCCAG
           **                  ************

Query      -------------------------------------------------------------
Subject    AGGGAGTCAGTCCATCCCATCCCCAGTGTGTGCAGGGCCTGCAGACACAGGAACGTACCG


Query      ---------------------AGTTAAGAGCCT---------------------------
Subject    ATACTCAACAGTGGCATCAAGAGTTAAGAGCCTTAATGTGAAATGCACAACTCGCCCTGA
                              ************

Query      ----CAGATGCAGATGCGTACAATC-------------------------------------
Subject    CACACAGATGCAGATGCGTACAATCCACCCCCACCACCCAAAAACTCCCTCTGGATCTCC
               *********************
```

Appendix

```
Query      ------------------------------------------------------------
Subject    GACTTTGACCTAAAAAGCCTCCCGCTGTGTTTTATCGCCCCCACTCCCAAAGCGTGTGTA


Query      ------CTACAGTCGCACAAT-----------------------------------------
Subject    TTGTAGCTACAGTCGCACAATAATCTATTTTCTTAAGACAGAATACCAGCAGTCCATGCA
                 ***************


Query      ----------AACCATTTTTACGGA----------TAATGAAATATCGAA--------TT
Subject    ATATATACGAAACCATTTTTACGGATGTGTACTTTTAATGAAATATCGAACTGTTTATTT
                     ***************          ***************        **


Query      TTTTCCCCCCTCCCCTCTC-------------------------------------------
Subject    TTTTCCCCCCTCCCCTCTCTCCTCTCTTTCATTGTGTAGCGGTTATGTACGAGTTCTGTG
           ******************


Query      --------------CTCAATGGTTAAAAA--------------------------TTTTC
Subject    TTTGCACTCAAAGGCTCAATGGTTAAAAATGCTGTTGGATGTTTGTGAGTGTTGCTTTTC
                       ***************                            *****


Query      ATCGCATCGATATTTTCACTTTC---------------------------------------
Subject    ATCGCATCGATATTTTCACTTTCAGTTTTCTCTTGAACGGCACAGACGAATTGTAAAAGA
           ***********************


Query      ------------------AGAATGTGCAAAAA-----------------------------
Subject    GCCATTGAAAGTGATGGAGAATGTGCAAAAAAAAAAAAAGGGAATTCAGAATCATTCTACG
                         **************


Query      ---------------GCTTTTTTTGTC----AGAGGTTGAAACAGA----TAGTTTAATT
Subject    GTGTTCTCATTGATGGCTTTTTTTGTCTCAAAGAGGTTGAAACAGATTTCTAGTTTAATT
                       ************    ***************    **********


Query      ATGCAATAGATTTTTATATTTTTAGTGGTCTGTT--------------------------
Subject    ATGCAATAGATTTTTATATTTTTAGTGGTCTGTTTTGTTTAGTTCTATGTGGTAGACGCT
           *********************************


Query      -----TAATTCACGCTTGTAAT---------------------------------------
Subject    TAAGGTAATTCACGCTTGTAATAAGTCATAGCTGGTAACCATTTCTCTGCGTTTTAGATC
                ******************


Query      ---------------ATCATTTATCCCAGTTTTTCC------------------------
Subject    ATGCAAAGCTGAAGCATCATTTATCCCAGTTTTTCCAAATATTTATAATGGCATTAGAAC
                       **********************


Query      ------GCTTAGTTTAAGACAG--------TGTTTTCTTCAT-------------------
Subject    AGAAAAGCTTAGTTTAAGACAGACAACTTTTGTTTTCTTCATGCTTTCAGATGAATGTCG
                 ****************        ************


Query      ----------------------AGTATTTTATGACCT----TCTTGTTTTACGATAGAAAA
Subject    GCCTGTTGACAATGTCCTGCGAGTATTTTATGACCTTTTATCTTGTTTTACGATAGAAAA
                            ***************    ********************


Query      AGAAA------------------TATTGTAATCTTC-------------------------
Subject    AGAAATACATTTTTGTTTGTTGTTATTGTAATCTTCTGTTCATTTATCTCTACAGTCTTC
           *****                   *************


Query      -----------------------------------TTGTTTATTGCATAATCAGTAAAAGG
Subject    AGTCTTCTATTCAACTCTCTTAACACTAAATGTATTGTTTATTGCATAATCAGTAAAAGG
                                             **************************


Query      --ATTGTGGACCAGACATTTGTGATGTATGTTTCA--------------------------
Subject    CAATTGTGGACCAGACATTTGTGATGTATGTTTCAATTTTCATATTGTTTTTTTTATTTT
             *********************************


Query      -------------ATATCAAGTTATTCAAAGT--------------------------AGT
Subject    TATTTATATATATATATCAAGTTATTCAAAGTACACTGAAATGTAATTAGTTCATAAAGT
                       ******************                           ***


Query      CACTTGAATGCA------------------------TTCTGAATTATTACATA------
Subject    CACTTGAATGCACTTCTATAATGAGCTTTATTCCTATTTCTGAATTATTACATAAATGTC
           ************                         *****************


Query      ------TTTACTATCCTGGAAT-----------------------------
Subject    TTTGTTTTTACTATCCTGGAATAACAACTAAACATTAAACATCACTCGGTT
                 ****************
```

12.3. BLASTn alignment of *epas1a* genomic locus and the synthetic transcript composed of sequences of *hif1ab* exhibiting similarity to *epas1a*.

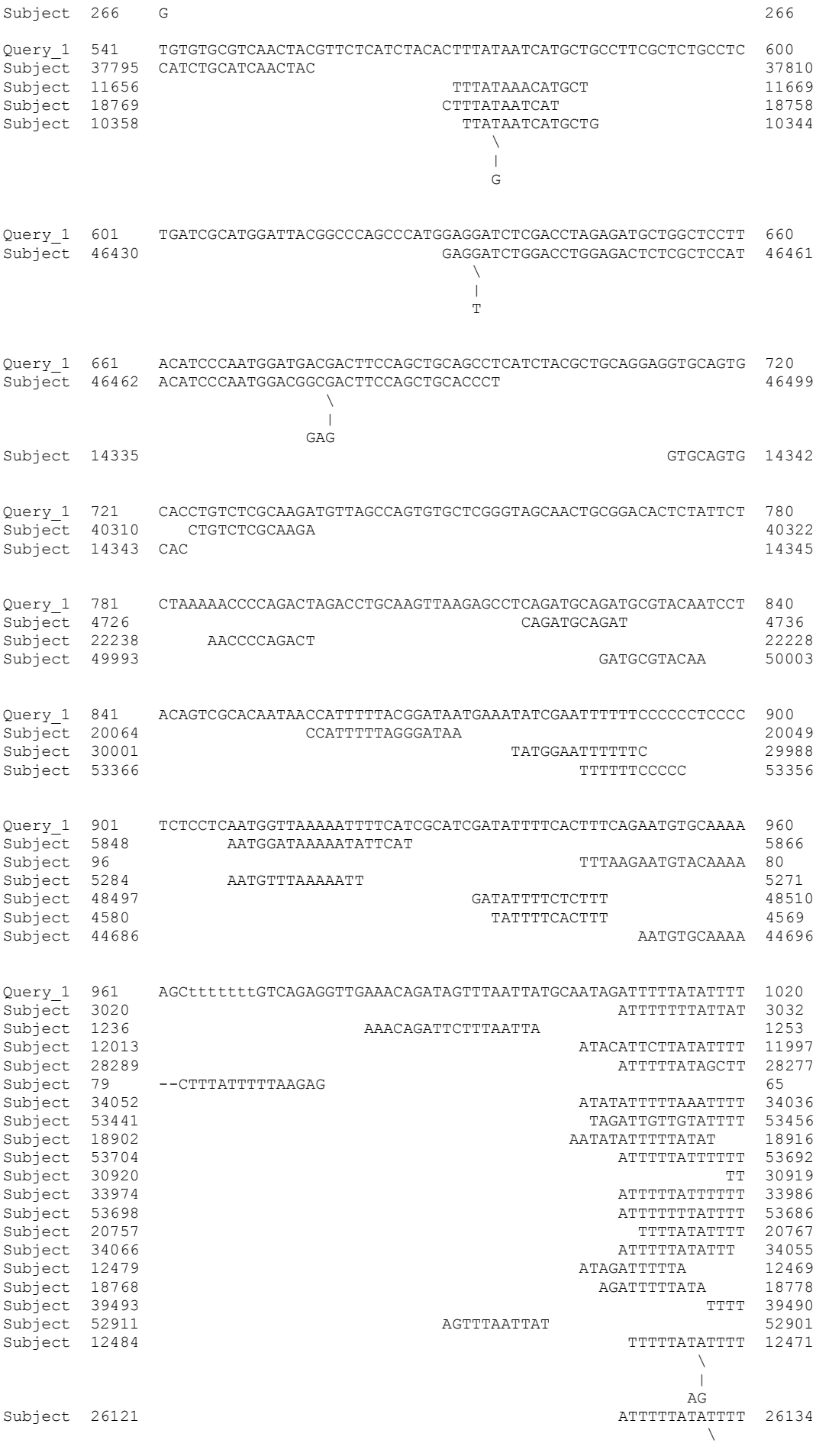
Alignment of the synthetic transcript composed solely of sequences of *hif1ab* mRNA exhibiting sequence similarity to *epas1a* genomic locus (Query) and the *epas1a* genomic locus (gene body +2kb upstream of transcription start site; Subject) using BLASTn (Zhang et al., 2000). To allow for better visibility, the maximum *E* value was set to 25 for this alignment.

```
Query= hif1ab_synthetic_transcript

Length=1277
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

  epas1a genomic inclusive promoter tss2000                          69.4    9e-14


Query_1  1      AGAGCCTCGTGAGCGCGCAAATGATTTGACCTTTCTTTACCATTCTGATTTTTATTTGTT  60
Subject  10906                                                    TTTTGTTTGTT  10916
Subject  52397                                                TCTGATTGTTACTTGTT  52413
Subject  20007                   CAAATGATTTAACCTTT                            20023
Subject  48643                              TTTCTTTCCCATTCTGA                 48659
Subject  33967                                       CATTTTTATTTTTATTT-TT  33985
Subject  488                                             TTTTTTTTTTTT  477
Subject  56443                                       CATACT-ATATTTATTTATT  56461
Subject  18314                         ATTTGACTTTTCTTTA                       18329
Subject  4904                                          ATTTTTATTTGTT  4916
Subject  5109                                          ATTTTAATTTGTT  5121
Subject  17118                                          TTTTTTTTTGTT  17107
Subject  27962                                          TTTTTATTTTTT  27973
Subject  53695                                          TTTTTATTTTTT  53684
Subject  1914   AGA-CCTCGTGAGCGCG                                        1929
Subject  10527                                         ATTTTTATTTGTT  10515
Subject  24063                                           TTTTATTTGTT  24053
Subject  35641                                           TTTTATTTGTT  35651
Subject  56342                                           TTTTATTTGTT  56352
Subject  54130                                             TTATTTGTT  54122
Subject  2062    GAG-CTCGTGAGCGC                                         2049
Subject  8579                                       TTCTGATTTTT          8589
Subject  11099                                            TTATTTGTT  11107
Subject  13252                    AATGATTTGAC                           13242
Subject  17948                                            TTATTTGTT  17956
Subject  44297                           TGACCTTTCTT                    44287
Subject  35614                                          TTTT-TTTGTT  35623
Subject  43794                                         TTTTTATTTGTT  43782
                                                           \
                                                           |
                                                           C
Subject  53704                                         ATTTTTATTT-TT  53693
Subject  32125                                          TTTTTATTTGTT  32137
                                                             \
                                                             |
                                                             A


Query_1  61     TTGTCACTGAAAAGAAAGGGTGAGCTCGGAGCGCAGGAAGGAGAAGTCCAGGGATGCAGC  120
Subject  11287                                  AAGGAGAAGTCTCGTGATGCGGC  11309
Subject  10917  TGGTTACTG                                                10925
Subject  52414  T                                                       52414
Subject  33986  TTG                                                     33988
Subject  476    TTGTC                                                   472
Subject  56462  TTGT                                                    56465
Subject  41624  TTGTCACTTAAAAGAA                                        41639
Subject  4917   T                                                       4917
Subject  5122   T                                                       5122
Subject  17106  TT                                                      17105
Subject  27974  TT                                                      27975
```

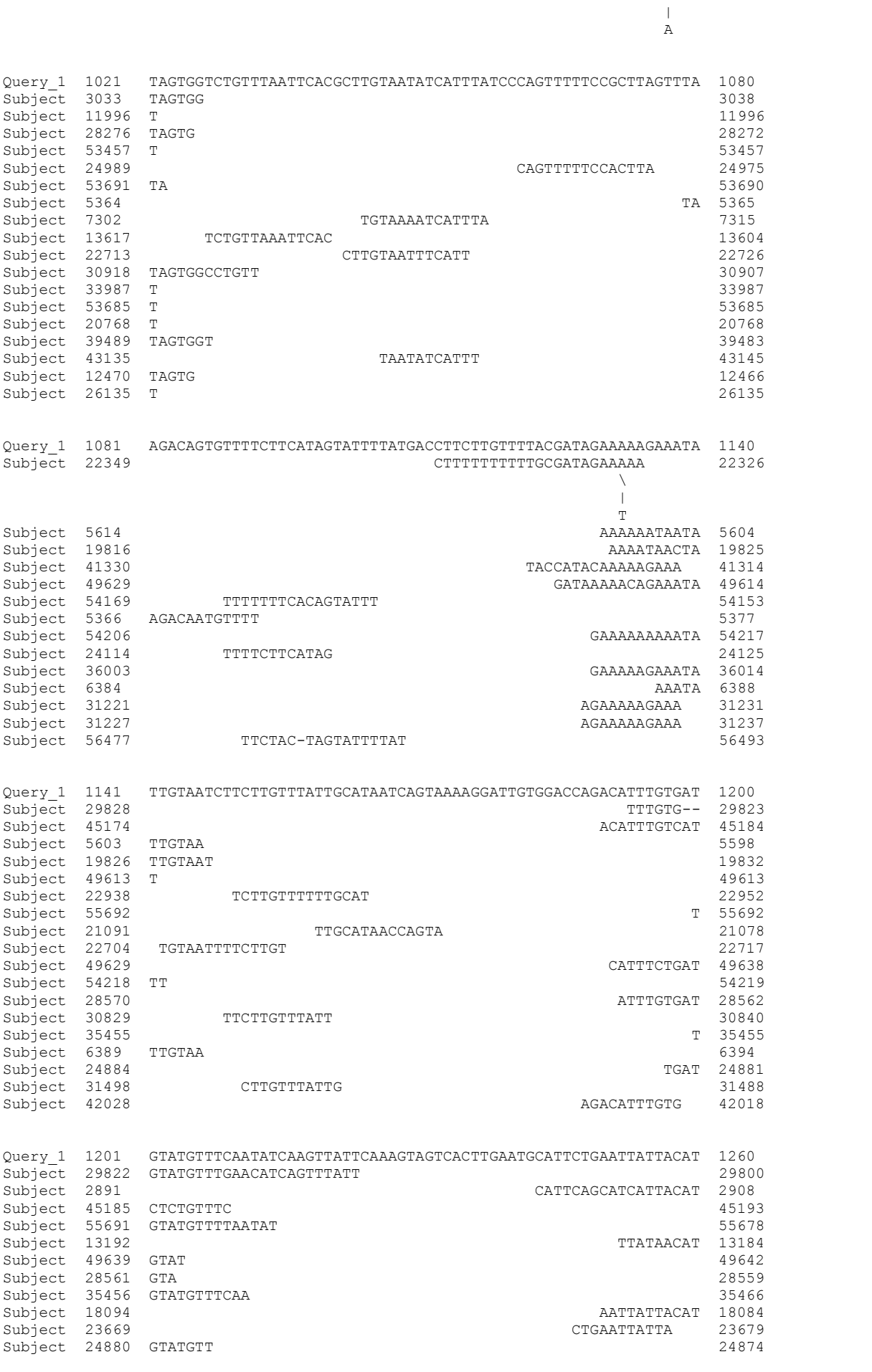
Appendix

```
Subject  53683  TT                                                          53682
Subject  24052  TT                                                          24051
Subject  35652  TT                                                          35653
Subject  56353  TT                                                          56354
Subject  21962             AGAAAGGGTGAG                                     21973
Subject  54121  TTG                                                         54119
Subject  11108  TT                                                          11109
Subject  17957  TT                                                          17958
Subject  47207      CACTGAAAAGA                                             47197
Subject  35624  TTGT                                                        35627
Subject  43781  TT                                                          43780
Subject  53692  TT                                                          53691
Subject  32138  TT                                                          32139


Query_1  121    GCGATCTCGCAGGGGAAAGGAGTCTGAGGTGTTCTACGAGTTAGCACACCAGCTCCCCCT  180
Subject  11310  GCG--CT-GCAGGGGTAAAGAGACAGAGGTGTTTTATGAGCTGGCCCATCAGCTACCATT  11369
                         \\
                         ||
                         C|
                          |
                         CA
Subject  11748                                          CACCCCAGCTCCACCT  11733
Subject  45475                                               AGCTCCCCCT  45466


Query_1  181    GCCACACAATGTCACGTCTCACCTGGACAAAGCCTCCATTATGAGGCTCACCATCAGCTA  240
Subject  11370  ACCCCACAGCATCAGCTCACACCTGGACAAAGCCTCTATCATGAGACTGGCTATCAGCTT  11429
Subject  11732  G                                                           11732
Subject  11355                                                  CCATCAGCTA  11364
Subject  45465  G                                                           45465


Query_1  241    CCTGCGCAGGAGGAGAATGAGCGATCTGACCGGTCACAGCATCTTTGAATTTTCACACCC  300
Subject  11430  CCTGCGCA                                                    11437
Subject  20420                          CTGACAGGCCACAGCATCTTCGACTTCACACATCC  20454
Subject  27006                               CGGTAACTGCATCTTTG               27022
Subject  3629                                          CATCTTTGATTTTT        3642
Subject  11365  CC                                     CATCTTTGAAT           26766
Subject  26756                                 GACCGGTCACA                  43315
Subject  43325                                           TTGAATTTTCA        44326


Query_1  301    ATGTGACCATGACTTCAGAAGACTCTGTCATCTGTGAGCCCATCATTGGACAGCAAGACC  360
Subject  24489                                                GACAGCAAGACG  24500
Subject  20455  TTGCGATCATGA                                                 20466
Subject  46716                  GAAGACTCTGTC                                46705
Subject  40122                                                 AGCAAGACC  40130
Subject  49980               CAGAAGACTCT                                    49970


Query_1  361    TTCCTCAGCCGCCACACTCTGGACATGAAGTTCTCATACTGTGATCAGATGACCTCTTCA  420
Subject  24501  TTCATGAGCAGACATAGTATGGATATGAAGTTCATTTACTGTGATGAGA             24549
Subject  22882                    TGGACATAAAGTTCT                           22868
Subject  46751                              GTTTTCATACTGTGA                  46737
Subject  14726                  TCTGAACATGAAGT                              14713
Subject  25678                  TCTGAACATGAAGT                              25665
Subject  45084                                                          A   45084
Subject  40131  TT                                                          40132


Query_1  421    GATCACCTCACCAAGCACCACAGGCCAGTACCGCATGCTGGCTAAGAAAGGTGGTTTTGT  480
Subject  37697                         GGCCAGTACAGAATGCTTGCCAAAAATGGAGGCTACGT  37734
Subject  20790                                              AAAGATGGTTTTGT  20803
Subject  45083  GATCACCTCATCAAGC                                             45067
                          \
                          |
                          T
Subject  256                                                    GTGGTTTTGT  265
Subject  12965                                       TGGCTAAGAAA            12955
Subject  21629                                              AAAGGTGGTTT     21639
Subject  39746     CACCTCACCAA                                              39756


Query_1  481    GTGGGTTGAGACTCAGGCCACTGTAATCTACAACCCCAAGAATTCTCAGCCGCAATGCAT  540
Subject  37735  TTGGGTGGAGACTCAAGCAACTGTCATCTACAACAACCGCAACTCCCAACCGCAGTGCAT  37794
Subject  20804  G                                                           20804
```

Appendix


Subject 266 G 266
Query_1 541 TGTGTGCGTCAACTACGTTCTCATCTACACTTTATAATCATGCTGCCTTCGCTCTGCCTC 600
Subject 37795 CATCTGCATCAACTAC 37810
Subject 11656 TTTATAAACATGCT 11669
Subject 18769 CTTTATAATCAT 18758
Subject 10358 TTATAATCATGCTG 10344
G
Query_1 601 TGATCGCATGGATTACGGCCCAGCCCATGGAGGATCTCGACCTAGAGATGCTGGCTCCTT 660
Subject 46430 GAGGATCTGGACCTGGAGACTCTCGCTCCAT 46461
T
Query_1 661 ACATCCCAATGGATGACGACTTCCAGCTGCAGCCTCATCTACGCTGCAGGAGGTGCAGTG 720
Subject 46462 ACATCCCAATGGACGGCGACTTCCAGCTGCACCCT 46499
GAG
Subject 14335 GTGCAGTG 14342
Query_1 721 CACCTGTCTCGCAAGATGTTAGCCAGTGTGCTCGGGTAGCAACTGCGGACACTCTATTCT 780
Subject 40310 CTGTCTCGCAAGA 40322
Subject 14343 CAC 14345
Query_1 781 CTAAAAACCCCAGACTAGACCTGCAAGTTAAGAGCCTCAGATGCAGATGCGTACAATCCT 840
Subject 4726 CAGATGCAGAT 4736
Subject 22238 AACCCCAGACT 22228
Subject 49993 GATGCGTACAA 50003
Query_1 841 ACAGTCGCACAATAACCATTTTTACGGATAATGAAATATCGAATTTTTTCCCCCCTCCCC 900
Subject 20064 CCATTTTTAGGGATAA 20049
Subject 30001 TATGGAATTTTTTC 29988
Subject 53366 TTTTTTCCCCC 53356
Query_1 901 TCTCCTCAATGGTTAAAAATTTTCATCGCATCGATATTTTCACTTTCAGAATGTGCAAAA 960
Subject 5848 AATGGATAAAAATATTCAT 5866
Subject 96 TTTAAGAATGTACAAAA 80
Subject 5284 AATGTTTAAAAATT 5271
Subject 48497 GATATTTTCTCTTT 48510
Subject 4580 TATTTTCACTTT 4569
Subject 44686 AATGTGCAAAA 44696
Query_1 961 AGCtttttttGTCAGAGGTTGAAACAGATAGTTTAATTATGCAATAGATTTTTATATTTT 1020
Subject 3020 ATTTTTTTATTAT 3032
Subject 1236 AAACAGATTCTTTAATTA 1253
Subject 12013 ATACATTCTTATATTTT 11997
Subject 28289 ATTTTTATAGCTT 28277
Subject 79 --CTTTATTTTTAAGAG 65
Subject 34052 ATATATTTTTAAATTTT 34036
Subject 53441 TAGATTGTTGTATTTT 53456
Subject 18902 AATATATTTTTATAT 18916
Subject 53704 ATTTTTATTTTTT 53692
Subject 30920 TT 30919
Subject 33974 ATTTTTATTTTTT 33986
Subject 53698 ATTTTTTTATTTT 53686
Subject 20757 TTTTATATTTT 20767
Subject 34066 ATTTTTATATTT 34055
Subject 12479 ATAGATTTTTA 12469
Subject 18768 AGATTTTTATA 18778
Subject 39493 TTTT 39490
Subject 52911 AGTTTAATTAT 52901
Subject 12484 TTTTTATATTTT 12471
AG
Subject 26121 ATTTTTATATTTT 26134

Appendix


|
A
Query_1 1021 TAGTGGTCTGTTTAATTCACGCTTGTAATATCATTTATCCCAGTTTTTCCGCTTAGTTTA 1080
Subject 3033 TAGTGG 3038
Subject 11996 T 11996
Subject 28276 TAGTG 28272
Subject 53457 T 53457
Subject 24989 CAGTTTTTCCACTTA 24975
Subject 53691 TA 53690
Subject 5364 TA 5365
Subject 7302 TGTAAAATCATTTA 7315
Subject 13617 TCTGTTAAATTCAC 13604
Subject 22713 CTTGTAATTTCATT 22726
Subject 30918 TAGTGGCCTGTT 30907
Subject 33987 T 33987
Subject 53685 T 53685
Subject 20768 T 20768
Subject 39489 TAGTGGT 39483
Subject 43135 TAATATCATTT 43145
Subject 12470 TAGTG 12466
Subject 26135 T 26135
Query_1 1081 AGACAGTGTTTTCTTCATAGTATTTTATGACCTTCTTGTTTTACGATAGAAAAAGAAATA 1140
Subject 22349 CTTTTTTTTTTGCGATAGAAAAA 22326
\
|
T
Subject 5614 AAAAAATAATA 5604
Subject 19816 AAAATAACTA 19825
Subject 41330 TACCATACAAAAAGAAA 41314
Subject 49629 GATAAAAACAGAAATA 49614
Subject 54169 TTTTTTTCACAGTATTT 54153
Subject 5366 AGACAATGTTTT 5377
Subject 54206 GAAAAAAAAATA 54217
Subject 24114 TTTTCTTCATAG 24125
Subject 36003 GAAAAAGAAATA 36014
Subject 6384 AAATA 6388
Subject 31221 AGAAAAAGAAA 31231
Subject 31227 AGAAAAAGAAA 31237
Subject 56477 TTCTAC-TAGTATTTTAT 56493
Query_1 1141 TTGTAATCTTCTTGTTTATTGCATAATCAGTAAAAGGATTGTGGACCAGACATTTGTGAT 1200
Subject 29828 TTTGTG-- 29823
Subject 45174 ACATTTGTCAT 45184
Subject 5603 TTGTAA 5598
Subject 19826 TTGTAAT 19832
Subject 49613 T 49613
Subject 22938 TCTTGTTTTTTGCAT 22952
Subject 55692 T 55692
Subject 21091 TTGCATAACCAGTA 21078
Subject 22704 TGTAATTTTCTTGT 22717
Subject 49629 CATTTCTGAT 49638
Subject 54218 TT 54219
Subject 28570 ATTTGTGAT 28562
Subject 30829 TTCTTGTTTATT 30840
Subject 35455 T 35455
Subject 6389 TTGTAA 6394
Subject 24884 TGAT 24881
Subject 31498 CTTGTTTATTG 31488
Subject 42028 AGACATTTGTG 42018
Query_1 1201 GTATGTTTCAATATCAAGTTATTCAAAGTAGTCACTTGAATGCATTCTGAATTATTACAT 1260
Subject 29822 GTATGTTTGAACATCAGTTTATT 29800
Subject 2891 CATTCAGCATCATTACAT 2908
Subject 45185 CTCTGTTTC 45193
Subject 55691 GTATGTTTTAATAT 55678
Subject 13192 TTATAACAT 13184
Subject 49639 GTAT 49642
Subject 28561 GTA 28559
Subject 35456 GTATGTTTCAA 35466
Subject 18094 AATTATTACAT 18084
Subject 23669 CTGAATTATTA 23679
Subject 24880 GTATGTT 24874

```
Subject  36036                                                  CTGAATTATTA     36026
Subject  38438                                                            ACAT  38435
Subject  44844                                                              AT  44843


Query_1  1261   ATTTACTAT  1269
Subject  2909   ATTTA      2914
                  \
                  |
                  G
Subject  13183  ATTTA      13179
Subject  38434  ATTTACT    38428
Subject  44842  ATTTACTAT  44834


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 71154903

  Database: epas1a_genomic_inc_promoter.fas
    Posted date:  Sep 18, 2018  1:36 PM
  Number of letters in database: 56,535
  Number of sequences in database:  1

Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 0, Extension: 2.5
```

Certain lines in this subsection have been quoted verbatim for the scientific accuracy of the terms used from (El-Brolosy et al., 2019).

13. Curriculum Vitae

Profile

Name	Mohamed Ahmed Abdallah Mohamed ElBrolosy
Sex	Male
Birth date	July 16th, 1992
Place of birth	Cairo, Egypt
Nationality	Egyptian
Address	Frankfurterstrasse 109, Bad Nauheim, Germany
Email	mohbrolosy@hotmail.com, mohamed.el-brolosy@mpi-bn.mpg.de
ORCiD	0000-0003-2433-1851
Languages	Arabic (mother-tongue), English (fluent) and German (very good)

Education

May 2016 – 2020	Ph.D. studies at the lab of Prof. Didier Stainier at the Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany	
2014 - 2016	M.Sc. in Molecular Biology at the International Max Planck Research School in Göttingen, Germany	GPA: 1.3 (Excellence)
2009 - 2014	B.Sc. in Pharmacy and Biotechnology at the German University in Cairo, Egypt	GPA: 0.7 (Excellence with highest honors). Ranked 1st on the class (composed of more than 400 students)
2008-2009	Egyptian high school diploma (Thanaweya Amma) at the Egyptian Language School in Cairo, Egypt	Grade: Excellence

Research Experience

05/2016 - Present	Ph.D. thesis entitled: "Investigating the molecular machinery underlying genetic compensation" **Thesis advisor**: Didier Stainier **Thesis committee**: Erin Schuman, Thomas Braun & Ingo Ebersberger
09/2015 - 03/2016	MSc. thesis entitled: "The role of Pyruvate Kinase-M2 in endothelial cell metabolism, angiogenesis and tumor progression", Prof. Didier Stainier's lab, Max Planck Institute for Heart & Lung Research, Bad Nauheim, Germany **Thesis advisors**: Didier Stainier & Herbert Jäckle
05/2015 - 06/2015	Rotation project entitled: "Germ Plasm Component z-vasa is Sufficient for Germ Cell Formation in Zebrafish", Dr. Roland Dosch's lab, Georg August University Medical Center, Göttingen, Germany
03/2015 - 04/2015	Rotation project entitled: "Overexpression analysis and CRISPR/Cas9-mediated endogenous protein tagging of prohibitins in human cells", Prof. Stefan Jakobs's lab, department of NanoBioPhotonics, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

01/2015 - 02/2015	Rotation project entitled: "Precise Levels of Dystroglycan are Required for Proper Gliogenesis and Blood Brain Barrier Integrity in Drosophila melanogaster", Dr. Halyna Shcherbata's lab, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany
10/2012 - 6/2014	Internship project entitled: "Investigating the miRNA 17-5p, E2F-1, c-MYC triad in Hepatocellular Carcinoma", the Molecular Pathology Research group (MPRG), Assoc. Prof. Ahmed Ihab's lab, the German University in Cairo, Cairo, Egypt
6/2013 - 8/2013	Internship project entitled: "Investigating the role of PKM isoforms on Vertebrate Vascular development and designing new cell ablation methods", Prof. Didier Stainier's lab, Max-Planck Institute for Heart & Lung Research, Bad Nauheim, Germany
8/2012 - 9/2012	Internship project entitled: "Investigating the mechanism of COX-II Inhibitors as an adjunct in pancreatic cancer therapy", Prof. Jörg Hoheisel's lab, German Cancer Research Center (DKFZ), Heidelberg, Germany
6/2012 - 8/2012	Internship project entitled: "Investigating the microtubule-association domain of the TRP ion channel; NOMPC", Prof. Howard's lab, Max Planck Institute of Molecular Cell Biology & Genetics, Dresden, Germany

Conferences attended and presentations

Oral presentations

11/2018	"Lindau alumni retreat", Frankfurt, Germany
10/2018	"Networks of cellular surveillance mechanisms", Heidelberg, Germany
07/2018	"Merck Curious 2018", Darmstadt, Germany
05/2018	"4th RMU RNA mini symposium - Molecular mechanisms of posttranscriptional regulation" - IMB, Mainz, Germany
04/2018	"3rd Regional fish meeting", COS, Heidelberg, Germany
09/2017	"International Max Planck Research school retreat", Ringberg, Germany

Poster presentations

10/2019	"International Max Planck Research school retreat", Hohenroda, Germany
06/2019	"2019 RNA society annual meeting", Krakow, Poland
10/2018	"Cold Spring Harbor Asia RNA biology meeting", Suzhou, China
10/2018	"Epigenetics: Basic principles and clinical applications", Cologne, Germany
09/2018	"International Max Planck Research school retreat", Weilheim, Germany
06/2018	"68th Lindau Nobel laureate meeting" - Lindau, Germany
06/2017	"Gene regulation by the numbers", IMB, Mainz, Germany
05/2017	"Microsymposium on small RNAs", Vienna Biocenter, Austria
10/2016	"Developmental epigenetics", IMB, Mainz, Germany

Key courses and workshops

03/2020	Programming for evolutionary biology course at the Freie Universität in Berlin, Germany.
2016 - 2019	Courses attended at the International Max Planck Research school for Heart and Lung Research in Bad Nauheim, Germany: • Image processing for microscopy • Bioinformatics: Analysis of Next-Generation-Sequencing data • Data analysis: from Excel to R • Statistics for biomedical researchers • Cell Sorting and flow cytometry • Adobe Illustrator and Photoshop
08/2013	Drug design intensive practical and theoretical summer school at the Eberhard Karls University of Tübingen, Germany.

09/2011	Training of tutors workshop at the German University in Cairo
2008 - 2011	Soft skill courses on topics like leadership and management, Cairo, Egypt.

Other work experience

2017	Supervision of a Bachelor student's research internship project at the Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany
2013	Community pharmacist at El-Brolosy pharmacies, Cairo, Egypt

Scholarships, Awards & Honors

2020	Forbes 30 under 30 list – Europe – Category of science and healthcare
03/2020	Finalist of the 2020 iBiology young scientist seminars competition.
10/2019	Poster prize at the 2019 International Max Planck Research school retreat
11/2019	Egyptian ministry of higher education medal
09/2019	International Max Birnstiel award
09/2019	Otto Bayer Scholarship from the Bayer foundation
10/2018	Best poster prize at the Cold Spring Harbor Asia 2018 RNA biology meeting
06/2018	Best poster prize at the 2018 Lindau Nobel laureate meeting
09/2017	Best student talk at the 2017 International Max Planck Research school retreat
06/2017	Best poster prize at the 2017 IMB conference "Gene regulation by the numbers"
2017 - 2019	Ph.D. fellowship from the Boehringer Ingelheim Fonds
2014 - 2015	Stipend by the International Max Planck Research School, Göttingen, Germany
2011, 2013	Top ranked full Scholarship at the German University in Cairo – Egypt
2013	DAAD scholarship for attending a drug design intensive practical and theoretical summer school at the Eberhard Karls University of Tübingen
2009	Scholarship for new entrants from the German University in Cairo (GUC) – Egypt
1997-2009	Several school honors school for outstanding performance

Extracurricular activities

2019	2019 Open Access week advisory committee
2018-2019	Board member of the N2 – Network of networks, representing doctoral researchers at non-university institutions in Germany
2018	General secretary and deputy spokesperson of Max Planck PhDnet
2018	Organizing committee of the 2018 OpenCon conference held in Toronto, Canada
2017 - 2018	Ph.D. student representative of the Max-Planck institute for Heart and Lung research
2015- Present	Member of Dietzenbach and Hessen Karate teams
2014	Vice President of the German University in Cairo Student Union
2002 - 2013	Member in Al-Ahly sporting club in Cairo karate team
2011 - 2012	Member of the German University in Cairo Curriculum Committee for developing the Pharmacy & Biotechnology curriculum

Publications

1. Serobyan V, Kontarakis Z, **El-Brolosy MA,** Welker JM, Tolstenkov O, Saadeldein AM, Retzer N, Wehman AM, Gottschalk A and Stainier DYR. Transcriptional adaptation in C. elegans. ***eLife***, 2020, Jan 17;9. pii: e50014
2. Fukuda R, Beisaw A, Marín-Juez R, El Sammak H, Aharonov A, Ong Y-T, Stone OA, **El-Brolosy MA**, Maischein H-M, Potente M, Tzahor E and Stainier DYR. Metabolic modulation by Erbb2 signaling regulates cardiac wall morphogenesis and regeneration. ***eLife***, 2019, Dec 23;8. pii: e50161

3. Lai JK, Gagalova K, Kuenne C, **El-Brolosy M**, Stainier DYR. Induction of Interferon-Stimulated Genes and Cellular Stress Pathways by Morpholinos. ***Developmental Biology***, 2019, 454, 21-28.
4. **El-Brolosy MA**, Kontarakis Z, Rossi A, Kuenne C, Günther S, Fukuda N, Khrievono K, Boezio G, Takacs CM, Lai S-L, Fukuda R, Gerri C, Giraldez AJ, Stainier DYR. Genetic compensation triggered by mutant mRNA degradation. ***Nature***, 2019, 568, 193-197
5. Rasouli SJ, **El-Brolosy M**, Tsedeke AT, Bensimon-Brito A, Ghanbari P, Maischein HM, Kuenne C, Stainier DY. The flow responsive transcription factor Klf2 is required for myocardial wall integrity by modulating Fgf signaling. ***eLife***. 2018 Dec 28;7. pii: e38889
6. Krishnakumar P, Riemer S, Perera R, Lingner T, Goloborodko A, Khalifa H, Bontems F, Kaufholz F, **El-Brolosy MA**, Dosch R. Functional equivalence of germ plasm organizers. ***PLOS genetics***. 2018 Nov 6;14(11):e1007696
7. Stone OA, **El-Brolosy M**, Wilhelm K, Liu X, Romão AM, Grillo E, Lai JKH, , Günther S, Jeratsch S, Kuenne C, Lee IC, Braun T, Santoro MM, Locasale JW, Potente M, Stainier DYR. Loss of pyruvate kinase M2 limits growth and triggers innate immune signaling in endothelial cells. ***Nature Communications***. 2018 Oct 9;9(1):4077
8. **El-Brolosy MA**, Stainier DYR. Genetic compensation: A phenomenon in search of mechanisms. ***PLOS Genetics***. 2017;13(7):e1006780
9. El Tayebi HM, Omar K, Hegy S, El Maghrabi M, **El Brolosy M**, Hosny KA, et al. Repression of miR-17-5p with elevated expression of E2F-1 and c-MYC in non-metastatic hepatocellular carcinoma and enhancement of cell growth upon reversing this expression pattern. ***Biochemical and Biophysical Research Communications***. 2013, 434 (3), 421-7