

**Characterizing telomere-binding protein
Hmbox1a/b and Zbtb48 in zebrafish**

Dissertation

Zur Erlangung des Grades
„Doktor der Naturwissenschaften“

am Fachbereich Biologie
der Johannes Gutenberg-Universität Mainz

Goh Sho Yee, Carisa

geb. am 1. Juli 1995 in Singapur

Mainz, 2025

(Blank)

Table of Contents

Acknowledgement	1
Abstract	2
Zusammenfassung	3
Introduction.....	4
End replication problem	4
End protection problem.....	6
Telomere-binding proteins	7
HMBOX1	9
ZBTB48	10
Zebrafish as a model organism for telomere study.....	12
Aim & objectives	13
Result	14
Telomere pull-down in zebrafish cell line BRF41 and embryos.....	14
Telomere pull-down in BRF41 zebrafish cell line	14
Telomere pull-down in zebrafish embryo	17
Hmbox1	20
Zebrafish Hmbox1	20
Raising zebrafish Hmbox1a/b polyclonal antibodies.....	23
Antibody validation on PAC2 zebrafish cells.....	25
Creating CRISPR-Cas knockout PAC2 zebrafish cells.....	27
Creating CRISPR-Cas knockout zebrafish lines	30
Validating CRISPR-Cas knockout fish lines and antibodies using IP-MS and Western blot analysis	34
Immunofluorescence (IF) staining on 1 dpf embryos	38
IF staining on 4 hours post-fertilization (hpf) embryos	45
Spatiotemporal expression of Hmbox1a during embryo development.....	48
Spatiotemporal expression of Hmbox1b during embryo development.....	51
Distribution of <i>hmbox1a/b</i> genotype and survival	56
Telomere length of (<i>hmbox1a</i> ^{LD/LD} ; <i>hmbox1b</i> ^{LD/LD}) double mutants.....	59
Proteome analysis at 4 hpf.....	59
Zbtb48.....	62
Zebrafish Zbtb48	62
Creating and validating CRISPR-Cas knockout zebrafish lines.....	63
Generating and validating Zbtb48 antibodies on PAC2	64
Validating Zbtb48 antibodies on <i>zbtb48</i> ^{-/-} 5 dpf larvae	67

Distribution of <i>zbtb48</i> genotype and survival	68
Transcriptomic and proteomic profiling of 5 dpf	70
Validating the downregulation of <i>mtfp1</i> and its phenotype in <i>zbtb48</i> ^{-/-} embryos	73
Spatiotemporal expression of <i>zbtb48</i>	75
Validating Zbtb48 expression in embryo through IF.....	77
Western blot analysis on gonads	80
qRT-PCR analysis on gonads.....	81
Proteome analysis on gonads.....	82
Discussion	86
Overview.....	86
Hmbox1a/b paralogues.....	88
Zbtb48.....	92
Conclusion.....	96
Materials.....	97
Biological materials.....	97
Cell lines.....	97
<i>E. coli</i> strains.....	97
Zebrafish lines.....	97
Plasmids.....	98
Oligonucleotides.....	99
Antibodies	104
Enzymes and Polymerase	105
Chemical/ Reagent	106
Media.....	109
Buffers and Solutions.....	110
Commercial kits	114
Instruments	115
Softwares.....	115
Methods.....	116
References	132
Resume	141

Acknowledgement

(Blank)

Abstract

Telomeres are nucleoprotein structures at the ends of linear chromosomes, bound by various proteins. Some of these proteins protect the telomere ends, such as the shelterin complex, while others help regulate telomere length. Recently, our lab has characterized two telomere-binding proteins, HMBOX1 and ZBTB48, in mammalian cell lines. These proteins were found to be conserved across evolution and regulate telomere length. To understand their role at the organismal level, we have chosen to study these proteins in zebrafish, an emerging model organism for studying telomere biology. Due to a genome-wide duplication in teleost, *Hmbox1* exists as paralogues in zebrafish. This study aims to investigate whether the telomere length regulatory functions of these proteins can be recapitulated in zebrafish and how they influence the organism.

To facilitate the study, CRISPR-Cas9 knockout zebrafish lines and antibodies against each of these proteins were generated. However, the knockout fish lines did not exhibit any severe phenotypes, and there was no significant difference in the telomere length of the first generation of homozygous mutants compared to their wild-type counterparts. Conversely, this study has found more success in exploring the spatiotemporal expression of these proteins during early zebrafish development. Based on immunofluorescence staining, *Hmbox1a* was found to be expressed during the blastula and gastrulation stages (3 hpf to 10 hpf), while *Hmbox1b* is expressed in the notochord and pronephros of embryos from 9 hpf to 3 dpf. On the other hand, *Zbtb48* was found to be highly expressed in a small population of cells in the gonads, specifically the germ stem cells (undergoing mitosis) and gametocytes (undergoing meiosis), according to the published scRNA-seq data. Additionally, *Mtfp1* was consistently observed to be downregulated in the *zbtb48*^{-/-} mutant, which was previously reported as one of the genes dysregulated by ZBTB48 in the human cell line study.

Overall, although this study did not report any phenotypes related to the telomere function of these proteins, it has provided new information on their physiological spatiotemporal expression during development, which has not been described previously.

Zusammenfassung

Telomere sind Nukleoproteinstrukturen an den Enden von linearen Chromosomen, die durch verschiedene Proteine gebunden sind. Einige dieser Proteine schützen die Telomerenden, wie z. B. der Shelterin-Komplex, während andere zur Regulierung der Telomerlänge beitragen. Vor kurzem hat unser Labor zwei telomerbindende Proteine, HMBOX1 und ZBTB48, in Säugetierzelllinien charakterisiert. Es wurde festgestellt, dass diese Proteine in der Evolution konserviert sind und die Telomerlänge regulieren. Um ihre Rolle auf der Ebene des Organismus zu verstehen, haben wir uns entschieden, diese Proteine im Zebrafisch zu untersuchen, einem aufstrebenden Modellorganismus für die Untersuchung der Telomerbiologie. Aufgrund einer genomweiten Duplikation im Teleostier existiert Hmbox1 im Zebrafisch als Paralogon. In dieser Studie soll untersucht werden, ob die die Telomerlänge regulierenden Funktionen dieser Proteine im Zebrafisch rekapituliert werden können und wie sie den Organismus beeinflussen.

Um die Studie zu erleichtern, wurden CRISPR-Cas9-Knockout-Zebrafischlinien und Antikörper gegen jedes dieser Proteine erzeugt. Die Knockout-Fischlinien wiesen jedoch keine schwerwiegenden Phänotypen auf, und es gab keinen signifikanten Unterschied in der Telomerlänge der ersten Generation homozygoter Mutanten im Vergleich zu ihren Wildtyp-Pendants. Im Gegensatz dazu war diese Studie erfolgreicher bei der Untersuchung der räumlich-zeitlichen Expression dieser Proteine während der frühen Zebrafischentwicklung. Anhand von Immunfluoreszenzfärbungen wurde festgestellt, dass Hmbox1a während der Blastula- und Gastrulationsstadien (3 hpf bis 10 hpf) exprimiert wird, während Hmbox1b im Notochord und Pronephros von Embryonen von 9 hpf bis 3 dpf exprimiert wird. Andererseits wurde festgestellt, dass Zbtb48 in einer kleinen Zellpopulation in den Keimdrüsen stark exprimiert wird, insbesondere in den Keimstammzellen (die Mitose durchlaufen) und den Gametozyten (die Meiose durchlaufen), wie die veröffentlichten scRNA-seq-Daten zeigen. Darüber hinaus wurde beobachtet, dass Mtfp1 in der *zbtb48*^{-/-} Mutante herunterreguliert wurde, was zuvor als eines der durch ZBTB48 dysregulierten Gene in der menschlichen Zelllinienstudie berichtet wurde.

Insgesamt wurden in dieser Studie zwar keine Phänotypen im Zusammenhang mit der Telomerfunktion dieser Proteine festgestellt, aber sie lieferte neue Informationen über ihre physiologische räumlich-zeitliche Expression während der Entwicklung, die bisher noch nicht beschrieben worden war.

Übersetzt mit DeepL.com (kostenlose Version)

Introduction

Linearization of chromosomes has provided significant advantages in eukaryotes for enabling meiosis and its key processes like independent assortment and meiotic recombination to increase genetic diversity and adaptability in offspring ¹. However, these structural changes introduced two major challenges: protecting chromosome ends from being mistaken for DNA breaks (the end-protection problem) and preventing progressive DNA shortening with each replication cycle (the end-replication problem).

These challenges are mitigated by a specialized nucleoprotein structure known as the telomere, located at the ends of eukaryotic linear chromosomes. In most cases, telomeres are composed of repetitive DNA sequences and associated proteins. These sequences typically form a guanine-rich (G-rich) leading strand and a complementary cytosine-rich (C-rich) lagging strand. Vertebrates share a common telomeric repeat (TTAGGG), which is also present in some invertebrates, such as corals and sponges ^{2,3}. This sequence is thought to be an ancestral telomeric sequence, with minor variations appearing across various species ³⁻⁵. For instance, (TTTAGGG) in *Arabidopsis thaliana* ⁶, (TTAGGC) in *Caenorhabditis elegans* ⁷, and (TTGGGG) in *Tetrahymena thermophila* ⁸. While the canonical telomeric sequence is present in some fungal species, telomeric sequence is highly diverse within the Ascomycota phylum, where species such as *Schizosaccharomyces pombe* (fission yeast) and *Saccharomyces cerevisiae* (budding yeast) display distinct telomeric sequences that drastically deviate from canonical telomeric sequence ⁵.

Despite the telomere sequence is highly conserved within mammals, great variation in telomere length can be observed across species. The length ranges from as long as 50 kb in *Panthera tigris corbetti* (tiger) to as short as 5 kb in *Zalophus californianus* (sea lion) ⁹. In humans, it typically spans around 5 to 15 kb ⁹. Moreover, telomere length can vary at the individual level, with differences observed across tissues ⁹⁻¹¹, among chromosome arms ¹² and even between homologous chromosomes ¹³. Thus, emphasizing the heterogeneity of telomere length and the complexity of regulating it.

End replication problem

During DNA replication, DNA polymerase synthesizes DNA in a unidirectional 5' to 3' manner. The leading strand is extended continuously in the 5' to 3' direction from the origin of replication using a single RNA primer, whereas the lagging strand requires multiple RNA primers to facilitate DNA synthesis in short fragments, known as Okazaki fragments. At the 5' end of the lagging strand, the removal of the terminal RNA primer leaves a gap, as there is no upstream DNA template for polymerase extension. This results in a progressive shortening of telomeres with each round of cell division ^{14,15}.

As telomeres shorten, they gradually lose the protection provided by the shelterin proteins. Eventually, they reach a critical length that activates the DNA damage response (DDR) pathway ^{16,17}. This response activates the transcription factor p53, which induces the expression of the cell cycle inhibitor p21, causing cell cycle arrest ^{18,19}. Consequently, the cell

either enters replicative senescence or undergoes apoptosis, restricting its ability to divide further. This limit on cellular replication is referred to as the Hayflick limit ²⁰.

Replicative senescence can act as a tumor suppressor by irreversibly arresting cell growth and preventing aberrant cellular proliferation ²¹. Otherwise, further erosion can lead to chromosome instability ²². However, replicative senescence is also characterized by a series of changes in cell morphology, gene expression, metabolism, and epigenetics, extending beyond just cell cycle arrest ²³. A key consequence of senescence is the senescence-associated secretory phenotype (SASP), which produces pro-inflammatory cytokines and growth factors that influence interactions between senescent and neighboring cells. In acute settings, senescence plays vital physiological roles, such as wound healing, injury repair, and embryonic development, by recruiting immune cells to clear senescent cells. However, chronic senescence is often considered detrimental, as it is commonly associated with aging, age-related diseases, and other pathological effects ²³⁻²⁶. Paradoxically, senescence can also promote tumorigenesis. Several SASP factors are linked to pro-tumorigenic processes, including mitogenic signaling and chronic inflammation, which alter the microenvironment and genetic makeup, creating conditions that allow cells to evade cell cycle arrest and become malignant ^{27,28}.

However, not all cells face the fate of replicative senescence. Certain cell types, such as stem cells and germ cells, can divide almost indefinitely due to the expression of a reverse transcriptase known as telomerase ²⁹. Telomerase comprises two key components: the RNA template (TER) and the catalytic enzyme (TERT) ³⁰. It extends the G-rich strand of telomeres with assistance from the shelterin subunit TPP1, facilitating its recruitment and enhancing processivity ^{31,32}. Meanwhile, the C-strand is synthesized by DNA polymerase α /primase, through its association with the CTC1–STN1–TEN1 (CST) complex ³³.

High levels of telomerase activity are present during early human development but become minimal in adult somatic cells ²⁹. When telomerase expression is reactivated in somatic cells, it can extend telomeres, allowing these cells to bypass replicative senescence, maintain telomere length, and effectively reverse cellular aging ^{22,34}. However, this reactivation is a double-edged sword, as telomerase expression in somatic cells can also promote cancer development. Approximately 85% of cancer cells express telomerase to preserve telomere length ^{35,36}, with the remaining 15% utilize homologous recombination, known as the alternative lengthening of telomeres (ALT) pathway ^{37,38}.

Collectively, maintaining telomere length is important for cellular functions and organism survival. A study has shown a negative correlation between telomere length in blood and mortality rate in humans aged 60 years or older ³⁹. Interestingly, the proportion of short telomeres, rather than the average telomere length, has been identified as a more critical predictor of lifespan in mammals ⁴⁰. While it may seem intuitive that species with longer telomeres would have longer lifespans due to their increased capacity for cell division, this assumption has been refuted by a study. For instance, mice have a much longer initial telomere length (50 kb) compared to humans (15 kb), yet their average lifespan is only about 2 years, whereas humans typically live around 80 years ⁴¹. Notably, a key difference is that mice experience a much faster rate of telomere shortening, at 7000 bp per year, compared to 70 bp per year in humans. This study concluded that the rate of telomere shortening, rather than the

initial telomere length, is a stronger predictor of lifespan across species. Collectively, these studies highlight the complexity of factors governing lifespan, extending beyond just the initial or average telomere length.

End protection problem

Apart from the end-replication problem, telomeres also face the end-protection problem, which poses a significant threat to genome integrity. An unprotected chromosome end resembles DNA double-strand breaks and risks being recognized as damage by the cellular repair machinery. This misrecognition can lead to unnecessary repairs, which may cause chromosome fusion and apoptosis. In vertebrates, the shelterin complex binds to telomeres and prevents the activation of DDR by masking the chromosome ends. In humans, the shelterin complex comprises six core proteins—TRF1, TRF2, POT1, RAP1, TPP1, and TIN2—which collectively play crucial roles in telomere maintenance and genomic stability (Figure 1) ⁴².

TRF1 and TRF2 are homodimer proteins, and they bind directly to double-stranded telomeric DNA through their Myb DNA-binding domains ⁴³. TIN2 links TRF1 and TRF2, while POT1 binds directly to the single-stranded telomeric overhang through its oligonucleotide/oligosaccharide binding (OB) folds ⁴⁴ and associates with TIN2 via TPP1 ⁴⁵. TPP1, which also contains an OB fold, does not bind to DNA, but instead plays a critical role in telomerase recruitment and enhancing telomerase processivity ^{31,32,46}. While RAP1 in humans contains a Myb-like domain, it does not bind directly to DNA ⁴⁷. Rather, it interacts with TRF2 to enhance its binding affinity for telomeric sequences ⁴⁸.

The two main shelterin components that play a pivotal role in repressing two major DDR pathways are POT1 and TRF2 ⁴⁹. POT1 inhibits the activation of the ataxia telangiectasia and Rad3-related (ATR) kinase signaling pathway by preventing replication protein A (RPA), a sensor of single-stranded DNA damage, from binding to the telomeric overhang ⁵⁰. This process heavily relies on the association of POT1 with TPP1, and the rest of the shelterin, to outcompete RPA ⁵¹.

On the other hand, TRF2 represses the ataxia-telangiectasia mutated (ATM) kinase signaling pathway by facilitating the formation of a T-loop, allowing the 3' telomeric overhang to invade the double-stranded region of the telomere ⁵². This T-loop structure effectively conceals chromosome end, preventing it from being recognized by the MRE11-RAD50-NBS1 (MRN) complex as DNA breaks, thus blocking ATM kinase activation. Without the T-loop protection, ATM kinase activation could trigger cell cycle arrest and lead to senescence or apoptosis ⁵³.

Formation of T-loop also relies on the accessory protein SNM1B/Apollo exonuclease (also known as DCLRE1B), which interacts with TRF2. Apollo generates 3' single-stranded overhangs on newly replicated leading-strand telomeres, allowing it to invade the double-stranded region of the telomere to form a T-loop ⁵⁴. The absence of Apollo or an impaired of its interaction with TRF2 leads to telomere instability through the activation of ATM signaling and the non-homologous end joining (NHEJ) pathway ^{54–56}.

Apart from forming T-loop, TRF2 also prevents NHEJ by interacting with the Ku proteins, which are crucial for the initiation of repair pathway^{16,57}. TRF2/RAP1 also works together to prevent NHEJ and inhibit homology-directed repair by repressing the localization of PARP1 and SLX4 to the telomere^{58,59}. The loss of TRF2 results in a detrimental outcome of chromosome end-to-end fusions⁶⁰, emphasizing its importance in maintaining genome stability.

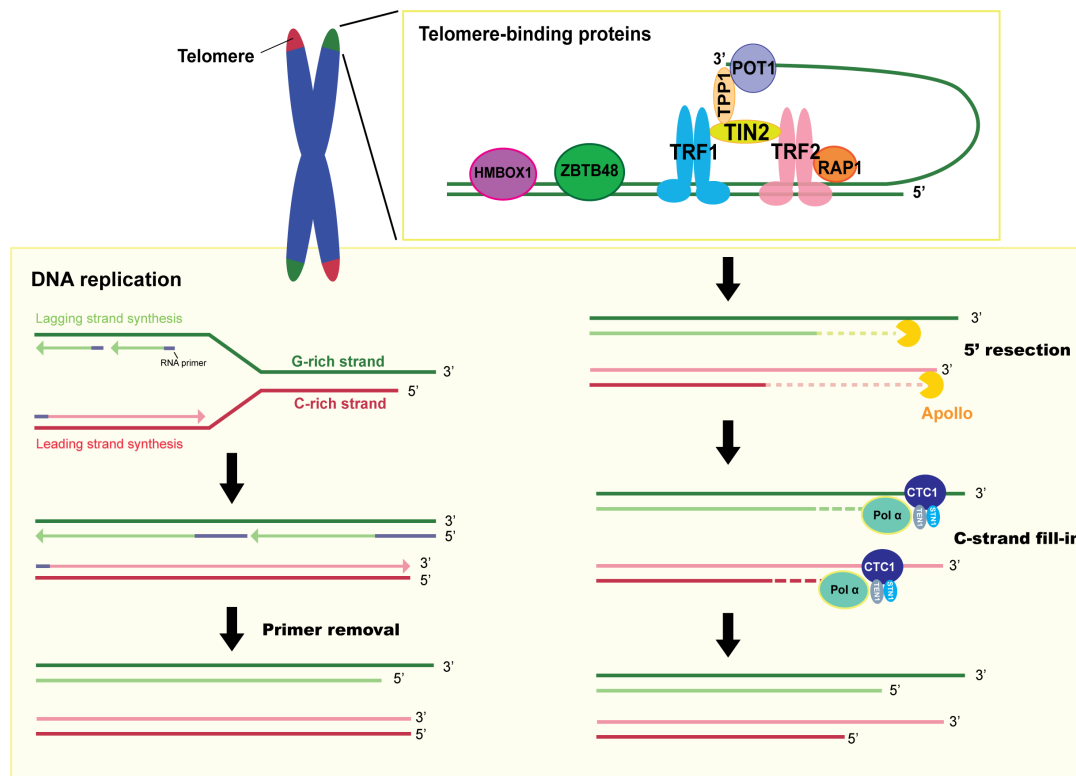


Figure I1. Overview of telomeres

A diagram illustrating telomere-binding proteins, including the six-subunit human shelterin complex and the proteins of interest in this study: HMBOX1 and ZBTB48 (top right). Below shows the telomere resection and C-strand fill-in during DNA replication. Images were adapted from de Lange (2018) and Maciejowski & de Lange (2019).

Telomere-binding proteins

In addition to the shelterin complex, the telomeres are also maintained by the trimetric CST complex. This complex is highly conserved and present across a wide range of species, including yeasts, plants, and mammals^{61,62}. Notably, its subunits are more closely related to RPA than to the shelterin proteins^{63,64}. The CST complex helps synthesize the lagging strand by binding to 3' G-rich telomeric overhangs and coordinating C-strand fill-in after Apollo exonuclease resection or telomerase extension (Figure I1)^{33,65,66}. This complex also maintains genomic stability by participating in various processes, such as DNA replication and DNA double-strand break repair throughout the human genome^{67–69}.

Over the decades, both the shelterin and CST complexes have been extensively studied by telomere biologists across various model organisms^{42,61,62,70}. Understandably so, as these complexes are highly conserved throughout evolution due to their critical roles in maintaining telomere integrity and overcoming the threats that telomeres face. In recent years, additional telomere-binding proteins, beyond the shelterin and CST subunits, have been identified and characterized, though most have only been briefly studied. Many of these proteins are also

conserved across vertebrates, presenting an opportunity to uncover new players that could play significant roles in telomere biology.

Among the identified telomere-binding proteins, several of them are zinc finger proteins. For examples, ZBTB10 ⁷¹, ZBTB40 ⁷², ZBTB48 ^{73,74}, and ZNF524 ⁷⁵. They all bind directly to telomeres via their BTB and/or ZNF domains. Some of these proteins reportedly help to maintain telomere stability or regulate telomere length, while the functions of one at the telomere remain unclear. Interestingly, many of these proteins appear to have a more significant impact on cells exhibiting ALT activity.

Both ZBTB40 and ZBTB48 have been reported as negative regulators of telomere length. The loss of ZBTB40 in ALT U2OS cells results in telomere elongation as well as telomere dysfunction ⁷². This is accompanied by an increased in ALT-associated PML bodies (APBs) accumulation, which APB serves as a site for ALT activity ⁷⁶. Thus, these findings indicate that ZBTB40 plays a role in both protecting telomeres and regulating their length.

In contrast, the loss of ZBTB48 leads to telomere elongation in telomerase-positive HeLa cells, but not in ALT U2OS cells ⁷³. However, overexpression of ZBTB48 in U2OS cells results in extrachromosomal accumulation and increased APB formation ⁷⁴, showing an opposite trend as compared to ZBTB40. While it remains unclear whether ZBTB48 provides telomere protection, it has been shown to regulate a small set of transcripts ⁷³.

The loss of ZNF524 in U2OS cells has shown to result in a slight reduction of the TRF2/RAP1 subcomplex at the telomere ⁷⁵. This leads to an increased DNA damage response and elevated recombination events, as indicated by telomeric sister chromatid exchanges (t-SCE), but not to the extent of telomere fusion observed in TRF2 knockout cells. Notably, the study did not find any evidence of physical interaction between ZNF524 and TRF2. Nevertheless, these observations suggest that ZNF524 plays a role in telomere protection.

On the other hand, ZBTB10 interacts with TRF2 through its N-terminal domain ⁷¹. However, its loss does not affect the abundance of TRF2 at the telomere, nor does it cause telomere dysfunction or activate ATM kinase. Unlike the other mentioned proteins, ZBTB10 also does not appear to play a role in regulating telomere length or transcription. It is likely that the primary function of ZBTB10 lies beyond the canonical telomere repeat sequence, as it shows preferential binding to the (TTGGGG) variant repeat sequence.

In addition to these zinc finger proteins, other proteins have also been identified as direct binders of telomeres. One such protein is the homeobox protein HMBOX1, which binds to telomeres through its homeobox domain ^{77,78}. It has been shown to act as a positive regulator of telomere length in HeLa cells ⁷⁷.

Collectively, these telomere-binding proteins exhibit varying roles at the telomere, with some offering intriguing insights that warrant further exploration. With the exception of ZBTB40 all the listed proteins are considered phylogenetically conserved telomere binders, as they have been found to be enriched in pull-down experiments with telomere probes in at least 5 out of the 16 vertebrate species screened ⁷⁹. Two proteins of particular interest for further study are

the homeobox protein HMBOX1 and the zinc finger protein ZBTB48. These proteins are highly conserved across vertebrates and they also regulate telomere length. Over the years, they have garnered attention from researchers investigating their potential roles in cancers and their physiological functions. The next section will delve deeper into our current understanding of these proteins.

HMBOX1

Homeobox containing 1 (HMBOX1), also known as homeobox telomere-binding protein 1 (HOT1) or telomere-associated homeobox-containing protein 1 (TAH1), is a homeobox protein belonging to the hepatocyte nuclear factor (HNF) gene class within the homeobox family. HMBOX1's function as a telomere binder is found in at least 5 out of the 16 vertebrate species screened ⁷⁹ and was first isolated from a cDNA library of the human pancreas ⁸⁰. Subsequent studies using immunostaining on tissue microarrays revealed its widespread expression in various human tissues, including the cerebrum, liver, testis, and renal tubules, among others ^{81,82}.

HMBOX1 binds directly to the telomeric sequences through its homeodomain, which is characterized by three alpha helices separated by a loop and a turn, resembling the Myb domain structure found in TRF1 and TRF2 ⁷⁷. A point mutation study identified several crucial residues in human HMBOX1, such as R271, K335, and R339, as essential for telomere binding. Among these, K335 was found to be the most critical for telomere sequence specificity, as the K335A mutation resulted in nonspecific binding.

The role of HMBOX1 in telomere function has been investigated by two independent groups using different cancer cell lines with distinct telomere lengthening mechanisms. While HMBOX1 was reported to act as a positive regulator of telomere length in telomerase-positive cells ⁷⁷, such an effect was not observed in the telomerase-negative ALT cancer cells ⁷⁸. Despite these differences, both studies consistently found HMBOX1 co-localizing with telomeres, regardless of the cells' telomerase activity.

Kappei et al. ⁷⁷ reported the knockdown of HMBOX1 in telomerase-positive HeLa cells resulted in telomere shortening, while its overexpression led to telomere lengthening. They proposed that HMBOX1 regulates telomere length by facilitating the recruitment of telomerase to telomeres, a hypothesis supported by several observations.

HMBOX1 was found to localize at the periphery of Cajal bodies, sub-nuclear structures where telomerase RNA accumulates and telomerase assembly occurs ^{83,84}. This localization pattern in Cajal bodies mirrors previous findings that associate telomerase RNA with telomeres ⁸⁵. Furthermore, immunoprecipitation experiments showed that HMBOX1 enriched for coilin, a structural component of Cajal bodies, along with the active telomerase complex components, including DKC1, GAR1, NHP2, and NOP10 ⁸⁶. Thus, suggesting that HMBOX1 helps recruit HMBOX1-bound telomeres to Cajal bodies, enabling telomerase activity.

Supporting this hypothesis, HMBOX1 knockout in mouse embryonic fibroblasts resulted in the loss of TERT from chromatin. Similarly, HMBOX1 knockdown in prostate adenocarcinoma

cells (LNCaP) has been shown to disrupt the interaction between TPP1 and TERT in another study ⁸⁷. Moreover, HMBOX1 showed higher colocalization with telomeres in cells with elevated telomerase activity, such as mouse spermatocytes ⁸⁸, where it localized to 87% of chromosome ends, compared to HeLa cells, which exhibited an average of 4.5 HMBOX1-telomeric foci per cell. These findings underscore HMBOX1's preferential localization to telomeres in cells with higher telomerase activity, supporting its role in telomerase recruitment.

On the other hand, Feng et al. ⁷⁸ observed that ALT U2OS cancer cells (70%) had a significantly higher proportion of HMBOX1 co-localized with telomeres compared to telomerase-positive HCT75 cancer cells (10%). In these ALT cells, HMBOX1 was also found to co-localize with APBs. While HMBOX1 knockdown led to a reduction in APBs, it did not significantly alter telomere length. However, it halved the levels of C-circles—key markers of ALT activity ⁸⁹—and increased telomere-associated DNA damage response (TIF) to levels similar to those observed with TRF2 knockdown. Thus, this finding indicates that HMBOX1 also plays a role in regulating ALT activity.

Taken together, HMBOX1 appears to positively regulate telomere length in both telomerase-positive and ALT cells. Its involvement in telomere length regulation has drawn attention to its potential link with cancer. However, tissue microarray studies have shown mixed results regarding HMBOX1 expression in cancerous and adjacent normal tissues, suggesting dysregulation may be tissue-specific ⁸¹. For example, HMBOX1 is decreased in liver cancer compared to normal tissue, but its expression is elevated in cancerous renal tubule cells. In contrast, HMBOX1 levels remain high in both normal and cancerous pancreatic tissues ⁸¹.

Several studies have explored the role of HMBOX1 in cancer progression by manipulating its expression levels. For instance, in gastric cancer, where HMBOX1 is upregulated, further overexpressing HMBOX1 correlates with poor prognosis as it promotes cell proliferation and migration ⁹⁰. In contrast, in high-grade serous ovarian cancer and other ovarian cancer cell lines (HO8910 and A2780), where HMBOX1 is downregulated, overexpression of HMBOX1 reduces cell proliferation, increases apoptosis, and elevates p53 expression ⁹¹. In both studies, HMBOX1 knockdown yielded the opposite effects. These findings suggest that the role of HMBOX1 in cancer progression is not universal but depends on the cancer type and its initial expression levels. Therefore, HMBOX1's involvement in cancer progression extends beyond telomere length regulation, and its detailed mechanisms remain to be explored.

Beyond its role in cancer, HMBOX1 also plays key physiological functions, such as promoting the differentiation of mouse embryonic stem cells and rat bone marrow stromal stem cells (BMSCs) into vascular endothelial cells (VECs) ^{92,93}. Additionally, HMBOX1 enhances autophagy and inhibits apoptosis in human umbilical vascular endothelial cells (HUVECs) ⁹⁴. HMBOX1 has also been shown to participate in immune regulation and inflammation ⁹⁵.

ZBTB48

Zinc finger and BTB domain containing 48 (ZBTB48), also known as telomeric zinc finger-associated protein (TZAP) and Human Krüppel-related 3 (HRK3), is a conserved telomere-binding protein across at least 13 of 16 vertebrate species screened ⁷⁹. Structurally, ZBTB48

contains an N-terminal BTB/POZ domain and 11 adjacent C2H2-type zinc fingers at the C-terminus. The telomere-binding function of ZBTB48 is mediated by several amino acid residues in its eleventh zinc finger and adjacent C-terminal arm, which directly interacts with TTAGGG telomere repeats⁹⁶. Telomere sequence recognition is facilitated by a recurring RxxHxxR motif shared among telomeric zinc finger proteins, including ZBTB10, ZNF524, and ZBTB48^{75,97}. Notably, ZBTB48 preferentially binds telomeres in an open chromatin state induced by the absence of the ATRX/DAXX protein complex⁹⁸.

The role of ZBTB48 in telomere length regulation has also been investigated by two independent groups, both of which identified it as a negative regulator of telomere length in both telomerase-positive and ALT cell lines. In telomerase-positive cells, such as HeLa cancer cells and mouse embryonic stem cells, ZBTB48 knockout led to telomere elongation^{73,74}. While ZBTB48 knockout in ALT U2OS cells did not result in significant telomere length changes⁷³, its overexpression in U2OS cells caused telomere shortening⁷⁴. This telomere shortening was accompanied by the accumulation of extrachromosomal DNA and the induction of APB formation. Furthermore, an inverse relationship between telomere length and ZBTB48 colocalization was observed. Based on these findings, Li et al.⁷⁴ proposed that the saturation of ZBTB48 at the telomere regulates telomere length by imposing an upper limit through telomere trimming.

Similar to HMBOX1, investigations into the correlation between ZBTB48 expression levels or mutations and survival outcomes in various cancers have yielded mixed results. These investigations range from breast cancer⁹⁹, cervical cancer¹⁰⁰, colorectal cancer¹⁰¹, lung adenocarcinoma and squamous cell carcinoma¹⁰², to hepatocellular carcinoma¹⁰³. In agreement with individual investigations, a meta-analysis of data from The Cancer Genome Atlas revealed that the prognostic impact varies depending on tumor type, and ZBTB48 alterations are present in only 5% of tumors¹⁰⁴.

The link between ZBTB48 dysregulation and cancer may be attributed to its roles beyond telomere regulation. ZBTB48 has been reported to bind directly to the promoter regions of certain genes and act as a transcriptional activator. Notably, it activates the tumor suppressor gene *Alternate Reading Frame (ARF)* in both immortalized human embryonic kidney cells (HEK293) and normal human dermal fibroblast (HDF)¹⁰⁵. As its name implies, ARF is a product of an alternative reading frame within the *INK4a/ARF* locus (*CDKN2A*), which the locus also encodes for p15^{INK4b} and p16^{INK4a}¹⁰⁶. ARF plays a critical role in maintaining p53 stability, thereby functioning as a tumor suppressor¹⁰⁷. However, this locus has been reported to be frequently mutated in melanoma¹⁰⁸.

ZBTB48 was also found to regulate *mitochondrial fission process 1 (MTFP1)*⁷³. The knockout of ZBTB48 in HeLa and U2OS cancer cells significantly downregulated *MTFP1* at both transcript and protein levels, leading to a mitochondrial matrix reorganization phenotype similar to that observed in MTFP1 depletion. Additionally, ZBTB48 has been recently reported as a regulator of B-cell-specific *C/ITA* expression in mice¹⁰⁹.

The role of ZBTB48 in stem cells has also been investigated, particularly in primary porcine mesenchymal stem cells (pMSCs). ZBTB48 levels were found to increase with pMSC passages, accompanied by the upregulation of cellular senescence markers such as p16^{INK4a}

and p21, indicative of cellular aging¹¹⁰. Overexpression of ZBTB48 in early-passage pMSCs led to phenotypes similar to those observed in late-passage cells, including reduced cell proliferation, diminished differentiation potential, and an increase in senescence markers. Consistent with ZBTB48's role in regulating ARF, this overexpression also elevated ARF, p53, and p21 expression levels. Conversely, ZBTB48 knockout yielded the opposite effects, even reversing the aging phenotype in late-passage pMSCs. Collectively, these findings highlight the critical role of ZBTB48 in maintaining stem cell proliferation capacity and differentiation potential through the p53 pathway.

In summary, previous studies have demonstrated that HMBOX1 and ZBTB48 play opposing roles in telomere length regulation. While their involvement in cancer has been widely studied, their expression levels and prognostic value remain inconsistent. Their roles in stem cell differentiation and proliferation have also been explored, though less thoroughly investigated compared to cancer. Preliminary findings indicate that HMBOX1 and ZBTB48 have contrasting effects on stem cells: HMBOX1 promotes differentiation, whereas ZBTB48 diminishes it. Overall, these findings highlight the complexity of their functions, demonstrating that there is no one-size-fits-all approach. A fine balance of their expression is required to regulate both physiological and pathological effects.

So far, much of our understanding of HMBOX1 and ZBTB48 has been limited to cell culture models, and their roles at the organismal level remain largely unexplored. These proteins have primarily been studied in isolation, within separate systems such as physiological cells (e.g., stem cells) and pathological cells (e.g., cancer cells). However, organisms are comprised of complex systems, and cell culture studies alone do not provide a complete understanding. To address this knowledge gap, we have selected zebrafish as a model organism to investigate their functions in a whole-organism context.

Zebrafish as a model organism for telomere study

Zebrafish (*Danio rerio*), a teleost fish native to freshwater habitats in South Asia, has an average lifespan of 3 to 5 years. Despite diverging from a common ancestor with humans about 4.5 million years ago, zebrafish share approximately 70% of their genes with humans, and around 80% of genes related to human diseases have at least an orthologue in zebrafish¹¹¹. However, due to a genome-wide duplication in the teleost lineage, some genes exist as paralogues in zebrafish. While Hmbox1 exists as paralogues in zebrafish, no paralogue of Zbtb48 has been identified.

Zebrafish are a popular model organism due to their high embryo count, rapid reproduction rate, transparent embryos, and ease of maintenance^{112,113}. Their transparent embryos, combined with ex vivo fertilization, make them especially valuable for developmental biology, enabling live imaging and in vivo studies¹¹⁴. They are also widely used in various research fields, including cancer^{115,116}, cardiovascular biology^{117–119}, and bone development¹²⁰, and have recently gained attention as a model organism for telomere studies¹²¹.

Zebrafish telomeres share several key characteristics with those of humans, including the presence of TTAGGG tandem repeats, which typically range from 5 to 15 kb in length¹²².

Despite the expression of telomerase in somatic tissue, telomeres in zebrafish shorten during aging ¹²³. In contrast to the laboratory inbred mouse strains, zebrafish exhibit telomere deficiency phenotypes already in the first generation of *tert*^{-/-} mutants correlated with significantly shorter telomeres compared to their wild-type counterparts ¹²².

The lifespan and telomere length of first-generation *tert*^{-/-} zebrafish mutants are reduced and they experience premature aging with a rapid decline in fertility ^{122,124}. Critical short telomeres accumulate in the gut and muscle ¹²⁵ and tissue-specific rescue of telomerase in the gut restores lifespan to some extent ¹²⁶. Second generation *tert*^{-/-} mutants do not survive to adulthood as the critical short telomeres activate p53-induced cellular apoptosis at the embryonic stage ¹²⁴. Consequently, *tp53*^{M214K} mutation in *tert*^{-/-} fish also results in partial lifespan extension ¹²⁷. In zebrafish, telomerase expression is further required for regeneration of the heart ¹²⁸ and for developmental hematopoiesis ¹²⁹. Notably, *tert*^{-/-} zebrafish mutants exhibit phenotypes comparable to those of dyskeratosis congenita (DC), a bone marrow failure disorder that affects multiple parts of the human body ¹²¹. Overall, these findings make zebrafish a suitable model for studying telomere biology.

Zebrafish possess all human shelterin orthologs, and their initial functional characterization has already been conducted. Loss of Terfa (the TRF2 ortholog) in zebrafish is embryonic lethal, and heterozygous adult fish exhibit accelerated aging ¹³⁰. In addition to the induction of DNA damage, *terfa*^{-/-} fish experience neurodevelopmental failure due to brain edema during embryo development based on transcriptional misregulation ¹³¹. Generated *pot1*^{-/-} and *acd*^{-/-} (the TPP1 ortholog) lines die at the larval stage, with a few *acd*^{-/-} mutant escapers showing premature aging ¹³². Furthermore, *acd* morphants exhibit neural death, heart defects, and extensive apoptosis during embryonic development ¹³³. In contrast, *tin2*^{-/-} (the TIN2 ortholog) and *terf1*^{-/-} mutants seemed unaffected and developed normally into adults ¹³². The *stn1* morphants, a component of the CST complex, resulted in a reduction in red blood cell count, an arrest in T cell progenitor development, and increased vascularity ¹³⁴. However, its impact on lifespan and development was not investigated. Furthermore, the effects of losing Terf2ip (the RAP1 ortholog) from the shelterin complex, as well as Ctc1 and Ten1 from the CST complex, remain unexplored in zebrafish. Similarly, the roles of Hmbox1 and Zbtb48 have not yet been investigated in zebrafish.

Aim & objectives

Given the promising role of HMBOX1 and ZBTB48 in regulating telomere length and cellular functions in mammalian cell lines, this study aimed to investigate their roles at an organismal level using zebrafish. To accomplish this, CRISPR-Cas9 knockout zebrafish lines and antibodies specific to these proteins were generated. The knockout lines were analyzed to assess their impact on telomere length and lifespan. Any abnormal phenotypes resulting from the mutations were documented. Using the generated antibodies, the spatiotemporal expression patterns of these proteins were examined. This study sought to determine whether the findings in zebrafish mirrored observations from human cell line studies and to explore any additional insights into how these proteins influence processes at the organismal level.

Result

Telomere pull-down in zebrafish cell line BRF41 and embryos

Telomere pull-down in BRF41 zebrafish cell line

A previous study has identified HMBOX1 and ZBTB48 as phylogenetically conserved telomere binders through pull-down using telomere probe ⁷⁹. They were considered phylogenetically conserved as they were enriched in at least 5 out of the 16 vertebrate species screened. However, only Zbtb48, but not Hmbox1a/b, was found significantly enriched in their BRF41 zebrafish cell line. To verify this result, the telomere pull-down experiment was repeated here with the procedures summarized in Figure 1a.

Telomeric TTAGGG oligonucleotide bait and a scrambled GTGAGT control sequence were lengthened through *in vitro* ligation and were biotinylated. They were immobilized on paramagnetic beads and incubated with nuclear-enriched protein lysate from the BRF41 zebrafish cell line. The experiments were performed in technical quadruplicate. The bound proteins were eluted, then measured on the mass spectrometer and analyzed by MaxQuant. Proteins enriched with telomere bait were compared to the scrambled control and the enrichment threshold was set at fold change of 2 and p-value of 0.05.

Out of the 1,269 proteins quantified in my BRF41 telomere pull-down, 46 were significantly enriched by the telomere oligonucleotide bait (Figure 1bi). This included all six orthologs of the core shelterin complex subunits: Terfa, Terf1, Pot1, Tinf2, Terf2ip, and Acd. Among the enriched proteins were also my proteins of interest: Hmbox1b and Zbtb48.

Interestingly, a large number of phylogenetically conserved TTAGGG binders were enriched in my BRF41 telomere pull-down experiments than previously reported ⁷⁹. Out of the 31 proteins that they listed as phylogenetically conserved TTAGGG binders, only 8 of them were enriched in their BRF41 experiment. In comparison, my pull-down experiment has not only successfully recapitulated the same 8 proteins but also enriched for 12 others phylogenetically conserved TTAGGG binders (Figure 2a).

The 8 phylogenetically conserved TTAGGG binders that were commonly enriched in both the BRF41 telomere pull-downs were the shelterin proteins Pot1, Terfa, and Terf2ip, nuclear receptor Nr2c2, transcription factor Runx2a, zinc finger protein Zbtb7a, and direct telomere binders Zbtb10 and Zbtb48. In addition to that, 3 other proteins that were not on the list of phylogenetically conserved TTAGGG binders were found to be enriched in both BRF41 pull-down studies. They were the catalytic proteins Accs and Parp1, as well as the zebrafish protein Zgc:171459 (Figure 2a).

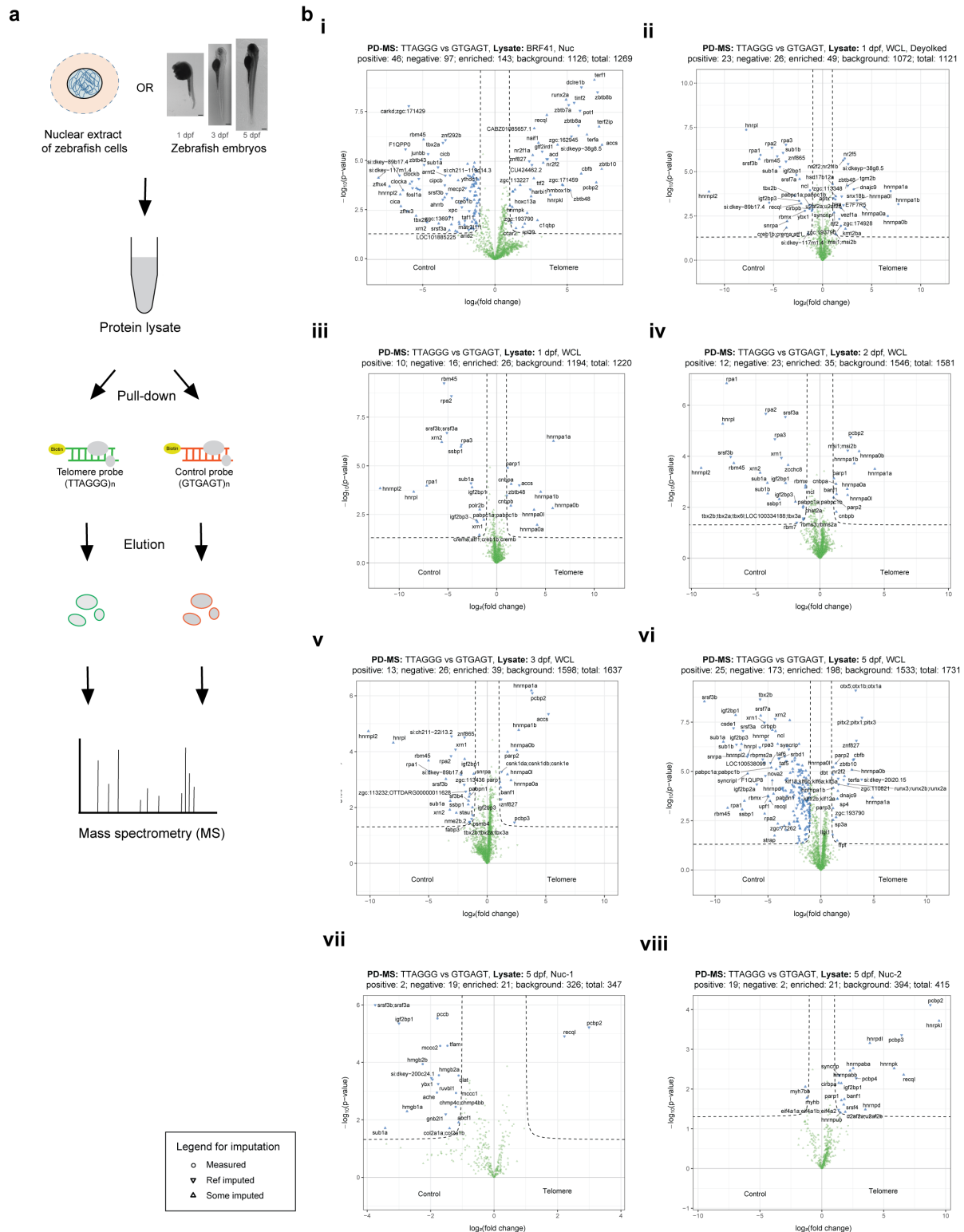


Figure 1. Telomere pull-down with the zebrafish cell line and embryos

a) A schematic diagram illustrating the telomere pull-down experiment using nuclear extracts from the zebrafish fin fibroblast BRF41 cell line and whole-cell lysates from zebrafish embryos. The images are embryos taken at 1 day post-fertilization (dpf), 3 dpf, and 5 dpf (scale bar = 200 μ m). The pull-down experiments were performed in quadruplicate using either concatenated telomeric TTAGGG sequences or a scrambled GTGAGT oligonucleotide as a control bait. The proteins eluted from the baits were analyzed using mass spectrometry.

b) Volcano plots visualizing the results from the label-free quantification of proteins bound in the telomere pull-down experiments conducted in this study. The experiments included telomere pull-downs performed on: **i**) nuclear extract of BRF41 cells, whole-cell lysate of **ii**) de-yolked and **iii**) non-de-yolked 1 dpf embryos, whole-cell lysate of non-de-yolked embryos at **iv**) 2 dpf, **v**) 3 dpf, **vi**) 5 dpf, and **vii**) & **viii**) nuclear extract of 5 dpf embryos. The results are plotted with $\log_2(\text{fold change})$ on the x-axis and $-\log_{10}(p\text{-value})$ on the y-axis. The enrichment threshold was set at a fold change > 2 and a p-value < 0.05 (Welch's t-test) with $c = 0.05$. Proteins are annotated with their gene names.

Due to the teleost lineage underwent a round of genome-wide duplication, genes like *hmbox1* exist as paralogues in zebrafish. Given the high similarity in their amino acid sequences, this raised the question whether any detection of Hmbox1a might have been mistakenly identified as Hmbox1b, especially since only the latter was detected in the pull-down experiment. To address this, the peptide sequences identified by the mass spectrometer and their corresponding identity assignments were inspected. However, it appears that although half of the 13 detected peptide sequences were common to both paralogues, the remaining sequences were unique to Hmbox1b (Figure 3). The absence of peptide sequences unique to Hmbox1a suggests Hmbox1b was likely the only detected paralogue.

Sequence detected as Hmbox1b in BRF41 IP-MS (Figure 1bi)	Position of amino acid (ref Hmbox1b)	
	Start	End
SHYTDPR	2	9
FTIEQIDLLQR	10	20
TGMTRPEIVHALDTLER	24	40
DALAATPNGK	128	137
YTVNSSVAVR	143	152
DSSLVKEEK	176	185
ISQAVVAQVTGISQSR	193	208
ISHWLLQGSSDLEQK	209	224
TTPGATLNMRTPTVSLSEMEWR	237	258
QTTPPISSAPGFSR	259	272
EEIANACNAVIQKPGK	305	320
VYNWFANR	333	340
PGGVASEIK	447	455

Figure 3. Identification of the Hmbox1 paralogue in the telomere pull-down from BRF41 cells

Peptide sequences detected by mass spectrometry that were identified as Hmbox1b in the BRF41 telomere pull-down (of Figure 1bi). Peptide sequences unique to Hmbox1b are highlighted in blue, and sequences shared between both paralogues are highlighted in purple.

Nevertheless, this pull-down experiment has confirmed the presence of Hmbox1b and *Zbtb48* in zebrafish and showed that they are able to recognize and bind to the telomere TTAGGG sequence. This experiment also identified additional TTAGGG binders in zebrafish, expanding the list of TTAGGG-binding proteins previously reported.

Telomere pull-down in zebrafish embryo

Proceeding with the success of telomere pull-down in the zebrafish cell line, I have extended the experiment to include embryos of various developmental stages to assess whether the results can be recapitulated in the organism. According to the RNA-seq data¹³⁵, *hmbox1a*, *hmbox1b*, and *zbtb48* transcripts, were highest during the early hours of development, especially before 1 day post-fertilization (dpf) (at prim-5 stage) (Figure 4). However, as embryos in the earlier stages have lower cell-to-yolk ratio, a large number of embryos would be required to obtain sufficient protein concentration for the pull-down experiments. As a result, for the ease of protein extraction, the earliest developmental stage used for pull-down experiments 1 dpf.

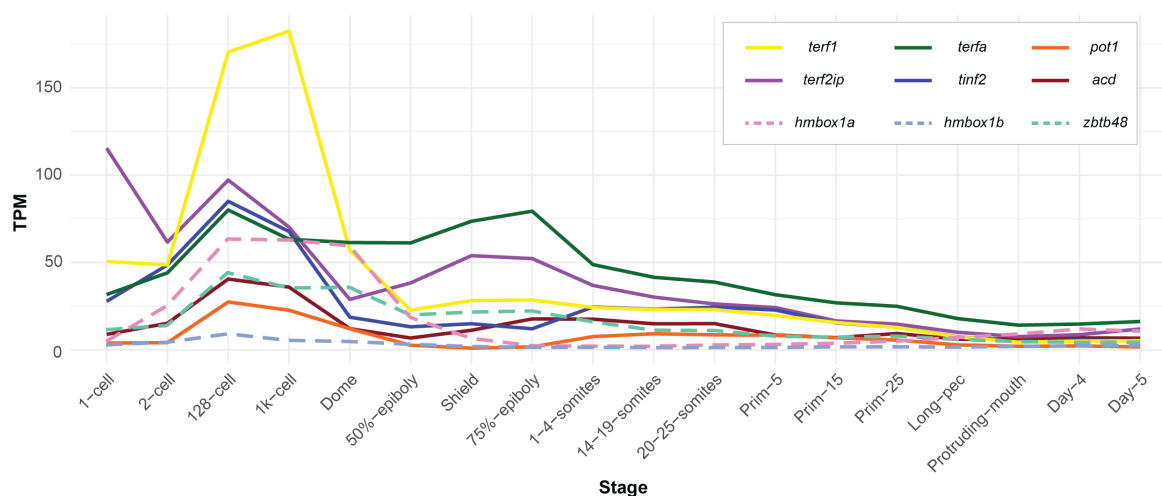


Figure 4. Developmental expression profile of shelterin subunits, *hmbox1* paralogues, and *zbtb48* mRNA across embryonic stages

Developmental mRNA expression level of shelterin subunits (plotted with solid lines), *hmbox1* paralogues and *zbtb48* (plotted with dashed lines) in whole embryos, replotted from previously published data (White et al., 2017). The ages of the embryos at the following developmental stages are 128-cell (2.25 hpf), shield (6 hpf), prim-5 (1 dpf), long-pec (2 dpf), and protruding-mouth (3 dpf).

At 1 dpf, the embryos have a notably large yolk, rich in nutrients and proteins essential for development, which could potentially compete with the proteins of interest during the pull-down process¹³⁶. To investigate whether the yolk material interferes with the pull-down results, I conducted the telomere pull-down using two different lysates: one subjected to de yolking to enrich embryonic cells and the other without de yolking. In these experiments, whole-cell protein lysates of the embryos were used instead of nuclear protein-enriched lysates.

The results revealed a comparable number of total quantified proteins between the two methods, with 1121 proteins identified in the de yolked sample and 1220 proteins in the non-de yolked sample. A slightly higher number of proteins were enriched with the telomere bait in the de yolked samples, with 23 proteins identified (Figure 1bii), compared to 10 proteins in the non-de yolked samples (Figure 1biii). However, only 6 of these proteins were commonly enriched in both 1 dpf pull-down experiments (Figure 2c). Among them were Zbtb48, which showed a 4.9-fold enrichment in the de yolked lysate and a 2-fold enrichment in the non-de yolked lysate, both with p-values below 0.0001 (Figure 2b). The remaining were RNA-binding proteins from the Hnrnpa family. Notably, unlike in previous pull-downs, Recql was found to be enriched with control bait in the de yolked samples.

When comparing the pull-down results from the 1 dpf embryo experiments to those from the BRF41 studies, Zbtb48 was the only phylogenetically conserved TTAGGG-binding protein consistently pulled down in both, while Vezf1a was enriched only in the de yolked lysates (Figure 2c). In addition to these 2 proteins, 3 other proteins (Si:dkeyp-38g8.5, Ttf2, and Zgc:193790) were commonly enriched between the BRF41 cells and the de yolked lysates. Conversely, 2 proteins (Accs and Parp1) that were enriched in both BRF41 studies were also commonly enriched in the 1 dpf non-de yolked lysates.

In conclusion, while the de yolking step increased the number of enriched proteins in the telomere bait, it did not improve the quality of the 1 dpf pull-down results. There was no increase in the enrichment of previously identified TTAGGG binders, nor significant enrichment of the shelterin proteins, and Hmbox1a/b paralogues were not detected.

To determine if the lack of detection of phylogenetically conserved TTAGGG binders persisted throughout development, I repeated the pull-down experiments with protein lysates from embryos at 2, 3, and 5 dpf (Figure 1biv to bvi). The de yolking procedure was omitted also due to the decreasing yolk size during development (see the example images in Figure 1a). Results showed a gradual increase in the total number of proteins quantified with development: 1581 proteins in 2 dpf, 1637 proteins in 3 dpf, and 1731 proteins in 5 dpf. Of which, 5 dpf exhibited the highest number of proteins enriched with telomere bait (25 proteins) compared to 2 dpf (12 proteins) and 3 dpf (13 proteins). However, the shelterin proteins remain largely absent in most samples, except for Terfa, which showed a 4.6-fold enrichment (p-value < 0.0001) at 5 dpf (Figure 2b).

Similar to the findings at 1 dpf, Hmbox1a/b paralogues remained undetected across all developmental stages. The enrichment of Zbtb48 at these stages was all below the defined threshold. Conversely, other direct telomere binders like Zbtb10 and Znf827 begin to show enrichment at later developmental stages, with Zbtb10 detected at 5 dpf and Znf827 after 3 dpf. Other phylogenetically conserved TTAGGG binders, such as Pcbp2, were consistently

detected across all stages but sometimes fell below the enrichment threshold. In contrast, Cbfb and Pcbp3 were only enriched in some of the samples. Interestingly, at 5 dpf, Hnrnpk and, once again, Recql were enriched by the control bait.

Given that direct telomere binders like Terfa, Zbtb10, and Znf827 were highly enriched at 5 dpf, this may be the optimal stage for enriching Hmbox1a/b paralogues and other shelterin proteins. To improve enrichment quality, the next pull-down experiment was performed using nuclear protein extracts instead of whole-cell lysates. However, this resulted in a decrease in the total number of quantified proteins to 347, with only Pcbp2 and Recql being significantly enriched with the telomere bait. Neither the shelterin proteins nor Zbtb48 were detected, despite their previous detection in whole-cell lysates at the same developmental stage (Figure 1bvii).

To address this issue, the experiment was repeated with double the amount of protein for the pull-down, additional washing steps were incorporated during nuclear protein extraction. Although these improvements increased the number of proteins enriched by the telomere bait to 19 out of 415 quantified proteins, shelterin proteins and other direct telomere binders, such as Hmbox1a/b, Zbtb10, Zbtb48, and Znf827, remained undetected (Figure 1bviii). Instead, Hnrnpk, Pcbp2, Pcbp3, and Recql were enriched in this pull-down. Interestingly, Hnrnpk and Recql were previously enriched with the control bait in whole-cell protein lysates, now exhibited higher enrichment with the telomere bait in nuclear protein-enriched extracts.

Unfortunately, the failure to detect direct telomere binders such as Terfa, Zbtb48, Zbtb10, and Znf827— which were present in the whole-cell lysates at 5 dpf (Figure 1bvi)— in any of the telomere pull-downs with nuclear protein-enriched lysates suggests that the nuclear protein extraction protocols used may have been ineffective.

Apart from the phylogenetically conserved TTAGGG binders, several proteins were observed to be consistently enriched in both zebrafish cell lines and embryos (Figure 2b, right). This suggests that they might be specific TTAGGG binders in zebrafish. These proteins include Accs and Parp1, previously identified in both the BRF41 studies (Figure 1bi and Kappei et al.⁷⁹), as well as DNA topoisomerase Top2a, nuclear receptor Nr2f2, RNA-binding proteins Hnrnpa0a, Hnrnpa0b, Hnrnpa0l, Hnrnpa1a, and Hnrnpa1b, and double-strand break repair proteins Parp2, and Parp3. Most of these proteins are involved in DNA or RNA surveillance. They may have been recruited because the baits used in the pull-down experiments mimic telomere ends with TTAGGG sequences, which resemble DNA breaks. Although the scrambled GTGAGT control sequence was designed to account for nonspecific binding, some proteins show a preference for either or both bait sequences. For instance, within the hnRNP family, Hnrnpk, identified as a phylogenetically conserved TTAGGG binder, was also found to be enriched with the scrambled sequence bait in the 5 dpf pull-down experiment with whole-cell lysate. In contrast, other hnRNP family members, such as the Hnrnpa proteins, were consistently enriched with the telomere bait, while Hnrnp1 proteins exhibited a stronger preference for the scrambled control over the telomere bait (Figure 2b, right).

In summary, these telomere pull-down experiments successfully enriched Zbtb48 in both the BRF41 zebrafish cell line and zebrafish embryos at various developmental stages. In contrast, only the Hmbox1b paralogue was detected in the BRF41 cell line and not in the embryos, indicating that further investigation into earlier stages may be necessary.

Hmbox1

Zebrafish Hmbox1

After the divergent from a common ancestor, the teleost lineage underwent a genome-wide duplication event resulting in genes like *hmbox1* to exist as paralogues in zebrafish. The zebrafish *hmbox1* paralogues are located on different chromosomes and span varying numbers of exons. *hmbox1a* is located on chromosome 17 and comprises 9 exons, while *hmbox1b* is on chromosome 20 and has 11 exons (Figure 5a). Based on gene synteny, zebrafish *hmbox1b* seems to be phylogenetically more related to the human *HMBOX1* than zebrafish *hmbox1a* (Figure 5b). All three *HMBOX1* homologs have an adjacent gene kinesin-like protein *KI13B*. However, the other neighbouring genes of human *HMBOX1* like the integrator complex gene *INT9* and the glycosyltransferase gene *EXTL3* are found in close proximity to the zebrafish *hmbox1b*, but not zebrafish *hmbox1a*.

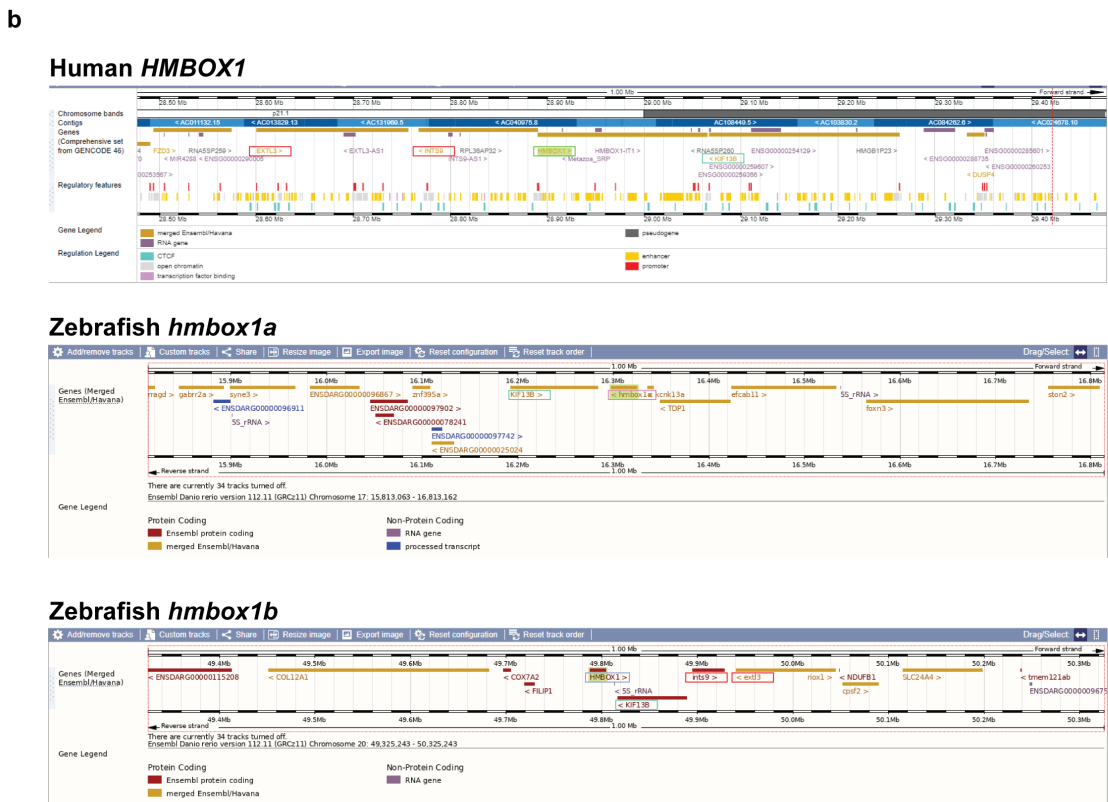
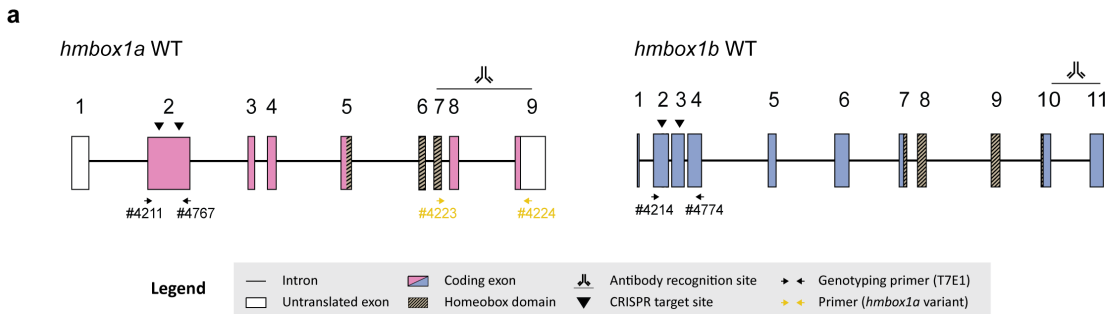


Figure 5. *HMBOX1* homologs

a) The zebrafish *hmbox1a* and *hmbox1b* genes with several key features highlighted CRISPR gRNA target sites, the antibody epitope site, the primer sites used for T7E1 assay and genotyping the *hmbox1a* variants.

b) The synteny of human *HMBOX1* (top), zebrafish *hmbox1a* (middle), and *hmbox1b* (bottom) are displayed. The screenshot was taken from the NCBI platform, with genes of the same identity highlighted using the same color boxes.

Despite millions of years of evolution, the amino acid sequences of zebrafish Hmbox1 paralogues remained highly similar (70% identity) to the human HMBOX1 (Figure 6a). The 72-amino-acid-long homeobox domain is the most conserved part of the protein, except for two residues. One difference is the asparagine in human HMBOX1 (N287) and zebrafish Hmbox1a (N299) is changed to another polar uncharged residue serine (S294) in zebrafish Hmbox1b. Another difference is that glutamic acid in human HMBOX1 (E288) is replaced by aspartic acid in both zebrafish paralogues (D300 in Hmbox1a and D295 in Hmbox1b), retaining the negative charge.

In the human study of HMBOX1, several amino acids within the homeobox domain have been identified as essential for telomere binding, and these are also conserved in zebrafish (highlighted in yellow in Figure 6a). A single mutation in any of these residues—R271, K325, K335, and R339—can completely abolish HMBOX1 telomere-binding activity, while a mutation in Y327 can weaken the binding to telomeres⁷⁷. Among these, K335 is critical for maintaining HMBOX1's specificity for telomeres; a mutation in this residue leads to non-specific binding to a scrambled sequence control (GTGAGT).

The high conservation of the telomere-binding homeobox domain across species, along with the successful enrichment of Hmbox1b in the telomere pull-down performed on BRF41 (Figure 1bi), strongly suggests that Hmbox1's ability to bind to the telomere sequence directly has been preserved throughout evolution.

The NCBI database has annotated and predicted multiple transcript isoforms for each paralogue; *hmbox1a* has 6 isoforms (one primary isoform (named as main) and isoform X1 to X5), while *hmbox1b* has 3 (isoform X1 to X3). These sequences diverged downstream of the homeobox domain, and the isoforms result from the inclusion of preceding intron sequences during splicing (annotated in lowercase in Figure 6b).

For *hmbox1a*, variations occur at two sites, resulting in 6 different isoforms. The first variation occurs immediately after the homeobox domain, where the addition of 3 to 6 intronic nucleotides before exon 8 resulted in 3 different isoforms (Figure 6bi). The second variation occurs 118 bp downstream that further divides the isoforms into two groups. The addition of 59 nucleotides from introns before exon 9 creates a frameshift mutation in the *hmbox1a* isoform X2, X3, and X4. Consequently, this led to *hmbox1a* X2, X3, and X4 isoforms encountering a premature stop codon 92 nucleotides earlier than the *hmbox1a* main and isoform X1 and X5. As a result, the proteins would be about 430 amino acids long, which are 10 amino acids shorter than the other group of isoforms (Figure 6c). On the other hand, *hmbox1b* isoform X1 has 15 nucleotides difference from isoform X2 and 18 nucleotides difference from isoform X3 (~460 amino acids) from the incorporation of intron sequences before exon 10 (Figure 6bii).

Despite the difference in *hmbox1a* and *hmbox1b* transcripts, all of them still retain at least 70% similarity at the amino acid level among the isoforms. Notably, the Hmbox1a main, X1, and X5 isoforms are more similar to Hmbox1b at the amino acid level than the Hmbox1a X2, X3, and X4 isoforms are to Hmbox1b when comparing the C-terminal sequence, where these two groups of Hmbox1a isoforms differ (Figure 6c).

a

Human_HMBOX1	MLSSFPVLETHSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	60
ZebraFISH_Hmbox1a	-----MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	48
ZebraFISH_Hmbox1b	-----MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	48
Human_HMBOX1	FGRRSSVYGGSSVGNSTN-----VPASSSTATASTQ--IQHSGMSPSPNS	104
ZebraFISH_Hmbox1a	FGRHAGGTSAINHSSNATSSNCTSSNMLTASSSTSTATQTFGRNGLSPPSPNS	108
ZebraFISH_Hmbox1b	FGRGSPDSIATNSITTSGLT-----PTSNASNATMATSSTSTATQTFPG-NLSPSPNS	104
Human_HMBOX1	YDTSPOCTN-----QNGREINERLSTSNKGMSPTRYHAN-SMGORSVFEASEE	154
ZebraFISH_Hmbox1a	YETSPPLAPTSMPVAVLPQNGP--DALASMSNGKLSPPRFVTS-NVSSRAFPEPPDE	165
ZebraFISH_Hmbox1b	YATSPPAVLPVSLV-ALAQGR--DALAATPNGLSPRNRYTVNSVAVRYTLEGGSE	161
Human_HMBOX1	DLDVDKVELMRDSSVIVKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLS	214
ZebraFISH_Hmbox1a	EMDIDVKEVELMRDSSVIVKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLS	225
ZebraFISH_Hmbox1b	DLDIDVKEVELMRDSSVIVKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLS	221
Human_HMBOX1	EKKQAFRYNQLEKTPGALSMRPAPIED-PEWRQTPPPVATSGTFLRRGSRFT	273
ZebraFISH_Hmbox1a	EKKQAFRYNQLEKTPGALSMRPAPIED-PEWRQTPPPVATSGTFLRRGSRFT	285
ZebraFISH_Hmbox1b	EKKQAFRYNQLEKTPGALSMRPAPIED-PEWRQTPPPVATSGTFLRRGSRFT	280
Human_HMBOX1	WRKECLAVMSEYFNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFR	333
ZebraFISH_Hmbox1a	WRKECLAVMSEYFNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFR	345
ZebraFISH_Hmbox1b	WRKECLAVMSEYFNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFR	340
Human_HMBOX1	RKEIKRRIANIE-----AILESHGIDVSPGHSNSDDVDGNDYSE	374
ZebraFISH_Hmbox1a	RKEIKRRIANIE-----AILESHGIDVSPGHSNSDDVDGNDYSE	392
ZebraFISH_Hmbox1b	RKEIKRRIANIE-----AILESHGIDVSPGHSNSDDVDGNDYSE	400
Human_HMBOX1	-----QDSTSHDHPQITSLAVEMAANHTILALARQGA--NEIKTEALD	418
ZebraFISH_Hmbox1a	-----GNDITGPNHQDPTALAVEMAAVNHSILALARQGSASDKTEVLDD	439
ZebraFISH_Hmbox1b	RFGYVRLPESPTDQDAGNHADQDPTSLAVEMAANHTILALARQGA--NEIKTEALD	460
Human_HMBOX1	DD 420	
ZebraFISH_Hmbox1a	DE 441	
ZebraFISH_Hmbox1b	DD 462	

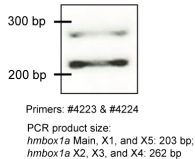
b

Exon 7		41223
Hmbox1a_Main	CTCTCTGAAATTTACACAGTGTTCGCACAGAGGAGAAATTAAGAGAGGCCCAACATAG	1894
Hmbox1a_X1	CTCTCTGAAATTTACACAGTGTTCGCACAGAGGAGAAATTAAGAGAGGCCCAACATAG	1897
Hmbox1a_X5	CTCTCTGAAATTTACACAGTGTTCGCACAGAGGAGAAATTAAGAGAGGCCCAACATAG	1880
Hmbox1a_X2	CTCTCTGAAATTTACACAGTGTTCGCACAGAGGAGAAATTAAGAGAGGCCCAACATAG	1880
Hmbox1a_X3	CTCTCTGAAATTTACACAGTGTTCGCACAGAGGAGAAATTAAGAGAGGCCCAACATAG	1897
Hmbox1a_X4	CTCTCTGAAATTTACACAGTGTTCGCACAGAGGAGAAATTAAGAGAGGCCCAACATAG	1894
Exon 8		41224
Hmbox1a_Main	CCAAACCCGGAGGATTTAAAGAGAGGAGATGATGGAGATGATGGAGATGATGGAGAT	1882
Hmbox1a_X1	CCAAACCCGGAGGATTTAAAGAGAGGAGATGATGGAGATGATGGAGATGATGGAGAT	1885
Hmbox1a_X5	CCAAACCCGGAGGATTTAAAGAGAGGAGATGATGGAGATGATGGAGATGATGGAGAT	1888
Hmbox1a_X2	CCAAACCCGGAGGATTTAAAGAGAGGAGATGATGGAGATGATGGAGATGATGGAGAT	1888
Hmbox1a_X3	CCAAACCCGGAGGATTTAAAGAGAGGAGATGATGGAGATGATGGAGATGATGGAGAT	1897
Hmbox1a_X4	CCAAACCCGGAGGATTTAAAGAGAGGAGATGATGGAGATGATGGAGATGATGGAGAT	1894
Exon 9		41224
Hmbox1a_Main	ATGTCCTCCCGTCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1326
Hmbox1a_X1	ATGTCCTCCCGTCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1329
Hmbox1a_X5	ATGTCCTCCCGTCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1332
Hmbox1a_X2	ATGTCCTCCCGTCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1296
Hmbox1a_X3	ATGTCCTCCCGTCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1293
Hmbox1a_X4	ATGTCCTCCCGTCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1293

ii

Exon 9		Exon 10	
Hmbox1a_X1	ACACATGTTTGGCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1340	1340
Hmbox1a_X2	ACACATGTTTGGCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1340	1340
Hmbox1a_X3	ACACATGTTTGGCCACACACACATTTGGCCGCGAGGAGGATGATGGAGATGATGGAGAT	1340	1340

d



c

Hmbox1a_Main	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	60
Hmbox1a_X1	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	60
Hmbox1a_X5	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	60
Hmbox1a_X2	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	60
Hmbox1a_X3	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	60
Hmbox1a_X4	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRHAGGTS	60
Hmbox1b_X2	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRGSDSIAT	60
Hmbox1b_X3	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRGSDSIAT	60
Hmbox1b_X1	MSHYDPRFTIEQIDLQLRLRRGTMTKHEILHALDLDLDRHEVKEFGRRGSDSIAT	60
Human_HMBOX1	INHSNATTSMSCTSSNMLTASSSTSTATQTFGRNGLSPPSPNSYETSPPTAPTS	120
Hmbox1a_X1	INHSNATTSMSCTSSNMLTASSSTSTATQTFGRNGLSPPSPNSYETSPPTAPTS	120
Hmbox1a_X5	INHSNATTSMSCTSSNMLTASSSTSTATQTFGRNGLSPPSPNSYETSPPTAPTS	120
Hmbox1a_X2	INHSNATTSMSCTSSNMLTASSSTSTATQTFGRNGLSPPSPNSYETSPPTAPTS	120
Hmbox1a_X3	INHSNATTSMSCTSSNMLTASSSTSTATQTFGRNGLSPPSPNSYETSPPTAPTS	120
Hmbox1a_X4	INHSNATTSMSCTSSNMLTASSSTSTATQTFGRNGLSPPSPNSYETSPPTAPTS	120
Hmbox1b_X2	NSTTSSTT--PTSNASNATMATSSTSTATQTFPGN--LSPSPNSYATSPPAVLPSP	116
Hmbox1b_X3	NSTTSSTT--PTSNASNATMATSSTSTATQTFPGN--LSPSPNSYATSPPAVLPSP	116
Hmbox1b_X1	NSTTSSTT--PTSNASNATMATSSTSTATQTFPGN--LSPSPNSYATSPPAVLPSP	116
Human_HMBOX1	MPVAVLPVQNGRDALASMSNGKLSPPRFVTS-NVSSRAFPEEPPDEDDDRVELMRM	179
Hmbox1a_X1	MPVAVLPVQNGRDALASMSNGKLSPPRFVTS-NVSSRAFPEEPPDEDDDRVELMRM	179
Hmbox1a_X5	MPVAVLPVQNGRDALASMSNGKLSPPRFVTS-NVSSRAFPEEPPDEDDDRVELMRM	179
Hmbox1a_X2	MPVAVLPVQNGRDALASMSNGKLSPPRFVTS-NVSSRAFPEEPPDEDDDRVELMRM	179
Hmbox1a_X3	MPVAVLPVQNGRDALASMSNGKLSPPRFVTS-NVSSRAFPEEPPDEDDDRVELMRM	179
Hmbox1a_X4	MPVAVLPVQNGRDALASMSNGKLSPPRFVTS-NVSSRAFPEEPPDEDDDRVELMRM	179
Hmbox1b_X2	VSLV-ALAQNGRDALATPNGLSPNRYTVNSVAVRYTLEGGSEDDDDVEELMR	175
Hmbox1b_X3	VSLV-ALAQNGRDALATPNGLSPNRYTVNSVAVRYTLEGGSEDDDDVEELMR	175
Hmbox1b_X1	VSLV-ALAQNGRDALATPNGLSPNRYTVNSVAVRYTLEGGSEDDDDVEELMR	175
Human_HMBOX1	DSAIKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	239
Hmbox1a_X1	DSAIKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	239
Hmbox1a_X5	DSAIKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	239
Hmbox1a_X2	DSAIKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	239
Hmbox1a_X3	DSAIKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	239
Hmbox1a_X4	DSAIKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	239
Hmbox1b_X2	DSSLVKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	235
Hmbox1b_X3	DSSLVKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	235
Hmbox1b_X1	DSSLVKEEKAFGLNRRISQAVAVQVIGISQSRISHMLLQGGSDLSQKQAFRYWQLE	235
Human_HMBOX1	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	299
Hmbox1a_X1	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	299
Hmbox1a_X5	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	299
Hmbox1a_X2	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	299
Hmbox1a_X3	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	299
Hmbox1a_X4	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	299
Hmbox1b_X2	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	294
Hmbox1b_X3	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	294
Hmbox1b_X1	KTTPGATLAMRPTPMALDIEVWRQTPPPISSTQGSFLRRGSRFTMRKECLAVMSEYFN	294
Human_HMBOX1	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANI	355
Hmbox1a_X1	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANI	355
Hmbox1a_X5	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANI	355
Hmbox1a_X2	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANI	355
Hmbox1a_X3	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANI	355
Hmbox1a_X4	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANI	355
Hmbox1b_X2	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANISGF	354
Hmbox1b_X3	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANISGF	354
Hmbox1b_X1	DNQYDPEAKREIETANACNAIQKPKKLSDLERVTSKLYVMFAFRKEKRRANISGF	354
Human_HMBOX1	-----AILESHGIDVSPGHSNSDDVDGNDYSE	393
Hmbox1a_X1	-----AILESHGIDVSPGHSNSDDVDGNDYSE	394
Hmbox1a_X5	-----EAAILESHGIDVSPGHSNSDDVDGNDYSE	395
Hmbox1a_X2	-----EAAILESHGIDVSPGHSNSDDVDGNDYSE	395
Hmbox1a_X3	-----AILESHGIDVSPGHSNSDDVDGNDYSE	394
Hmbox1a_X4	-----AILESHGIDVSPGHSNSDDVDGNDYSE	393
Hmbox1b_X2	-----EATILESHGIDVSPGHSNSDDVDGNDYSE	409
Hmbox1b_X3	-----EATILESHGIDVSPGHSNSDDVDGNDYSE	408
Hmbox1b_X1	QEATILESHGIDVSPGHSNSDDVDGNDYSE	419
Human_HMBOX1	MDNTGPNHQDPTALAVEMAAVNHSILALARQGSASDKTEVLDD	441
Hmbox1a_X1	MDNTGPNHQDPTALAVEMAAVNHSILALARQGSASDKTEVLDD	442
Hmbox1a_X5	MDNTGPNHQDPTALAVEMAAVNHSILALARQGSASDKTEVLDD	443
Hmbox1a_X2	MVCEAKAEGR-PISSLL--PYR--TTTLARMSRTRPSLWQ	432
Hmbox1a_X3	MVCEAKAEGR-PISSLL--PYR--TTTLARMSRTRPSLWQ	431
Hmbox1a_X4	MVCEAKAEGR-PISSLL--PYR--TTTLARMSRTRPSLWQ	430
Hmbox1b_X2	QDAGNHADQDPTSLAVEMAANHTILALARQGSASDKTEATD	457
Hmbox1b_X3	QDAGNHADQDPTSLAVEMAANHTILALARQGSASDKTEATD	456
Hmbox1b_X1	QDAGNHADQDPTSLAVEMAANHTILALARQGSASDKTEATD	462

Figure 6. Sequence alignment of HMBOX1 homologs

a) Amino acid alignment of human HMBOX1 with zebrafish Hmbox1a and Hmbox1b paralogues. The homeobox domain is boxed in brown. Amino acid residues critical for telomere binding, as identified in human studies (Kappei et al., 2013), are highlighted in yellow. Sequences written in green and highlighted in pink or blue represent the epitope sequences used to raise antibodies against Hmbox1a and Hmbox1b, respectively.

b) cDNA sequence alignment showing the site of variations among the different i) *hmbox1a* and ii) *hmbox1b* isoforms. Exons are color-coded with different shades and labeled with exon numbers. Exon sequences are written in uppercase, while intron sequences are in lowercase, variations are boxed in orange, and stop codons are written in red. An annealing site for primer pair #4223 and #4224 to determine the presence of the two group of isoforms used in (d) are highlighted in yellow.

c) Amino acid alignment of all zebrafish Hmbox1a and Hmbox1b isoforms. The homeobox domain is boxed in brown. Sequences written in green and highlighted in pink or blue represent the epitope sequences used to raise antibodies against Hmbox1a and Hmbox1b, respectively. Sequences in green that are not highlighted represent common epitope sequences in both paralogues.

d) PCR gel results from the amplification of *hmbox1a* cDNA obtained from wild-type BRF41 using primer pair #4223 and #4224.

To assess the existence of the two groups of *hmbox1a* isoforms in zebrafish, the RNA was extracted from BRF41 and reverse transcribed into cDNA. The region spanning the 59-nucleotide insertion site using a primer pair was amplified as indicated in Figure 5a. The PCR results confirmed the presence of both groups of *hmbox1a* isoforms, as evidenced by the appearance of two distinct bands of correct size (Figure 6d).

Next, to check for the presence of all the transcript isoforms, the full-length transcripts were amplified using respectively primers and cloned into a pCR8/TOPO/GW vector and sequenced. In each set of cloning, 6 colonies were sequenced, and the results confirmed the existence of all *hmbox1a* and *hmbox1b* transcript isoforms except for *hmbox1a* isoform X5, which was not detected.

Raising zebrafish Hmbox1a/b polyclonal antibodies

Since antibodies against zebrafish Hmbox1a and Hmbox1b were not commercially available, these antibodies were raised in this study. The process of antibody production is summarized in Figure 7ai and ii for the respective antibodies. To ensure these antibodies are paralogue-specific, the epitope sequences were designed to be as unique as possible by targeting the C-terminus region of Hmbox1a/b, where the variation between paralogues primarily occurs. Additionally, these antibodies will help validate whether the CRISPR-induced mutation results in an in-frame functional protein, as they target a region downstream of the gRNA's target site (refer to Figure 5a).

The epitope sequences were amplified from a holding vector using primers with overhangs for subcloning into the pCoofy4 expression vector, which tags the epitope sequence with His₆-MBP at the N-terminus. The cloned pCoofy4 plasmids were transformed into BL21 *E. coli* cells, which were then induced with IPTG to overexpress the respective fusion proteins. The His₆-MBP-Hmbox1a epitope (48 kDa) was successfully overexpressed, as indicated by a prominent band observed between 42 and 57 kDa (Figure 7b). However, the His₆-MBP-Hmbox1b epitope (49 kDa) did not show a comparable band. To improve expression, the LB culture of BL21 cells was supplemented with 3% ethanol¹³⁷, which led to a noticeable improvement in expression. These epitope proteins were then purified and sent to an external company for immunization.

Three rabbits were immunized for each epitope, and sera were collected 61 days after immunization. Unfortunately, one of the rabbits (T2) immunized with the His₆-MBP-Hmbox1a epitope died during the process, so sera were only purified from the remaining two rabbits. For the Hmbox1b epitope, only the serum from rabbit T2 was purified. Each serum was purified using the corresponding epitope, and the polyclonal antibodies were collected across 10 elution fractions.

During the epitope design stage, the chosen epitope sequences were BLASTed on NCBI and only the paralogue itself was returned as a hit. However, after the antibodies were raised it was discovered that parts of the sequence are common between the paralogues (Figure 6d, common sequence written in green). The first 15 amino acids of the 47 amino acids long Hmbox1a epitope sequence is shared between both paralogues, which would generate a pan-Hmbox1a/b antibody. Conversely, the 57 amino acid long epitope of Hmbox1b has the last 15 amino acids shared with the main, X1, and X5 isoforms of Hmbox1a. Despite the potential for generating pan-Hmbox1a/b antibodies from both epitopes, paralogue-specific antibodies are still likely to be developed since the remaining two-thirds of the epitope sequence are unique to each paralogue. However, rigorous testing is required to check for their specificity.

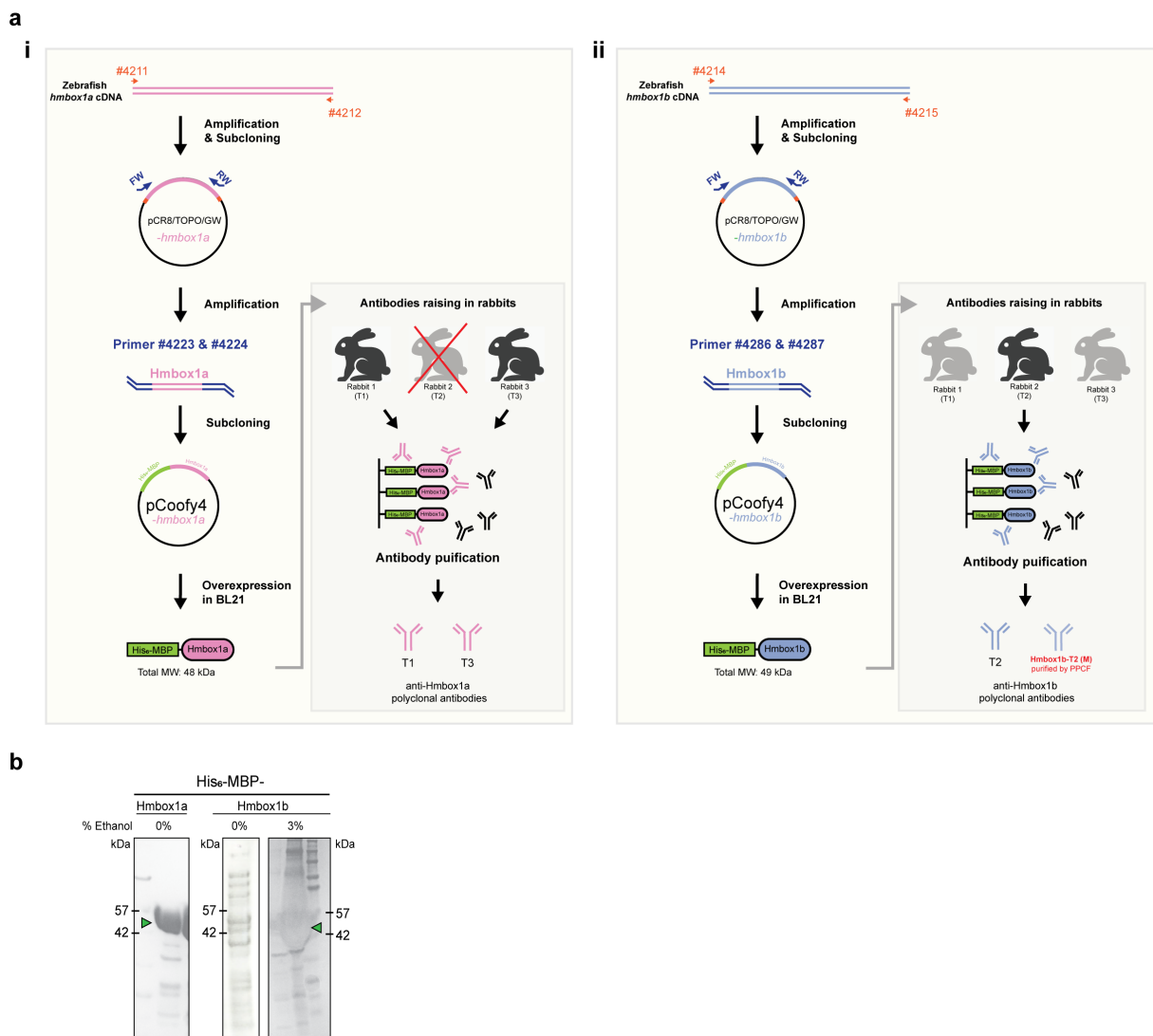


Figure 7. Raising zebrafish Hmbox1a/b polyclonal antibodies in rabbits

a) Schematic diagram illustrating the procedures for raising polyclonal antibodies against i) Hmbox1a and ii) Hmbox1b. The epitope sequence was amplified from the respective vector and subcloned into a pCoofy4 expression vector, which tags the synthesized epitope sequence with His₆-MBP at the N-terminus. The tagged epitopes were then overexpressed in BL21 *E. coli*. Three rabbits were immunized per epitope, and the sera were purified against the epitope proteins. One of the rabbits injected with the Hmbox1a epitope died during the immunization (marked with a cross).

b) Gel images showing the overexpression of epitope proteins in BL21 after an overnight IPTG induction. The recombinantly expressed His₆-MBP-Hmbox1a epitope is 48 kDa, while His₆-MBP-Hmbox1b is 49 kDa (marked with green arrowhead).

Antibody validation on PAC2 zebrafish cells

Antibodies purified from the first five elutions were tested in Western blot against whole cell lysates from the PAC2 cell line (Figure 8a). PAC2, an embryonic fibroblast cell line derived from 1 dpf zebrafish embryos, was used in place of BR41 due to its extensive literature^{138,139} on transfection applications for subsequent experiments. However, no bands between the Hmbox1a molecular weight of 48 and 49 kDa were observed with Hmbox1a-T1 and Hmbox1a-T3 antibodies. Instead, bands appeared between the 53 and 70 kDa markers and below the 170 kDa marker for Hmbox1a-T1 antibody (Figure 8ai), while only a band was observed around the 130 kDa marker for Hmbox1a-T3 antibody (Figure 8aii). On the other hand, for the Hmbox1b-T2 antibody, multiple bands were observed around the 53 kDa marker, which is close to the predicted molecular weight of Hmbox1b (51 kDa), and another band was also observed around the 130 kDa marker (Figure 8aiii). However, assessing the success of antibody production based solely on the Western blot results is challenging, as multiple bands were observed, making it unclear whether any of them correspond to the Hmbox1a/b protein.

To identify the targets of these antibodies, immunoprecipitation followed by mass spectrometry (IP-MS) analysis was deployed. The experiments were performed on the nuclear extract of PAC2. To prevent false positive hits, the epitope sequences of both Hmbox1a and Hmbox1b along with the MBP sequences were also added to the MS search library. Interestingly, the results revealed an enrichment of Hmbox1b and MBP proteins in both the Hmbox1a-T3 and Hmbox1b-T2 antibody IP experiments, while showing negligible enrichment of these proteins with the Hmbox1a-T1 antibody (Figure 8bi to biii). As MBP is not an endogenous zebrafish protein, its presence in the IP-MS results suggests that this epitope protein was not thoroughly removed during the antibody purification process.

To verify the detection of the endogenous Hmbox1a/b, the peptide sequences identified by mass spectrometry were examined. The sequences were categorized into four groups: Hmbox1a paralogue-specific sequences (pink), Hmbox1b paralogue-specific sequences (blue), sequences shared between both paralogues (purple), and epitope sequences used to raise the antibodies (green) (Figure 8c). Both Hmbox1a-T3 and Hmbox1b-T2 antibodies were able to detect all four categories of Hmbox1a/b sequences. This finding confirmed the enrichment of Hmbox1b, as both antibodies identified several unique Hmbox1b sequences in addition to the epitope sequences. However, their detection of unique Hmbox1a sequences was limited and primarily through matching across different replicates. The extensive coverage of Hmbox1b sequences by Hmbox1a-T3 antibody indicates that it functions as a pan-Hmbox1a/b antibody. On the contrary, the Hmbox1a-T1 antibody showed minimal detection of Hmbox1a/b sequences and failed to identify any unique Hmbox1a peptide sequences.

Proteins enriched by each antibody were analyzed to identify common hits. Most of the identified proteins were structural proteins, including actin, myosin, tubulin, and collagen (Figure 8d). Five proteins (Col5a1, Myh10, Myh9b, Myo1b, and Myo5a) were commonly enriched across all antibodies. While this observation could suggest that these proteins are binding partners of Hmbox1a/b, the absence of Hmbox1a/b detection with the Hmbox1a-T1 antibody indicates otherwise. It is more plausible that these proteins represent non-Hmbox1a/b recognition by other antibodies present in rabbit sera, which were either not removed or bound to the His₆-MBP tag during the purification process.

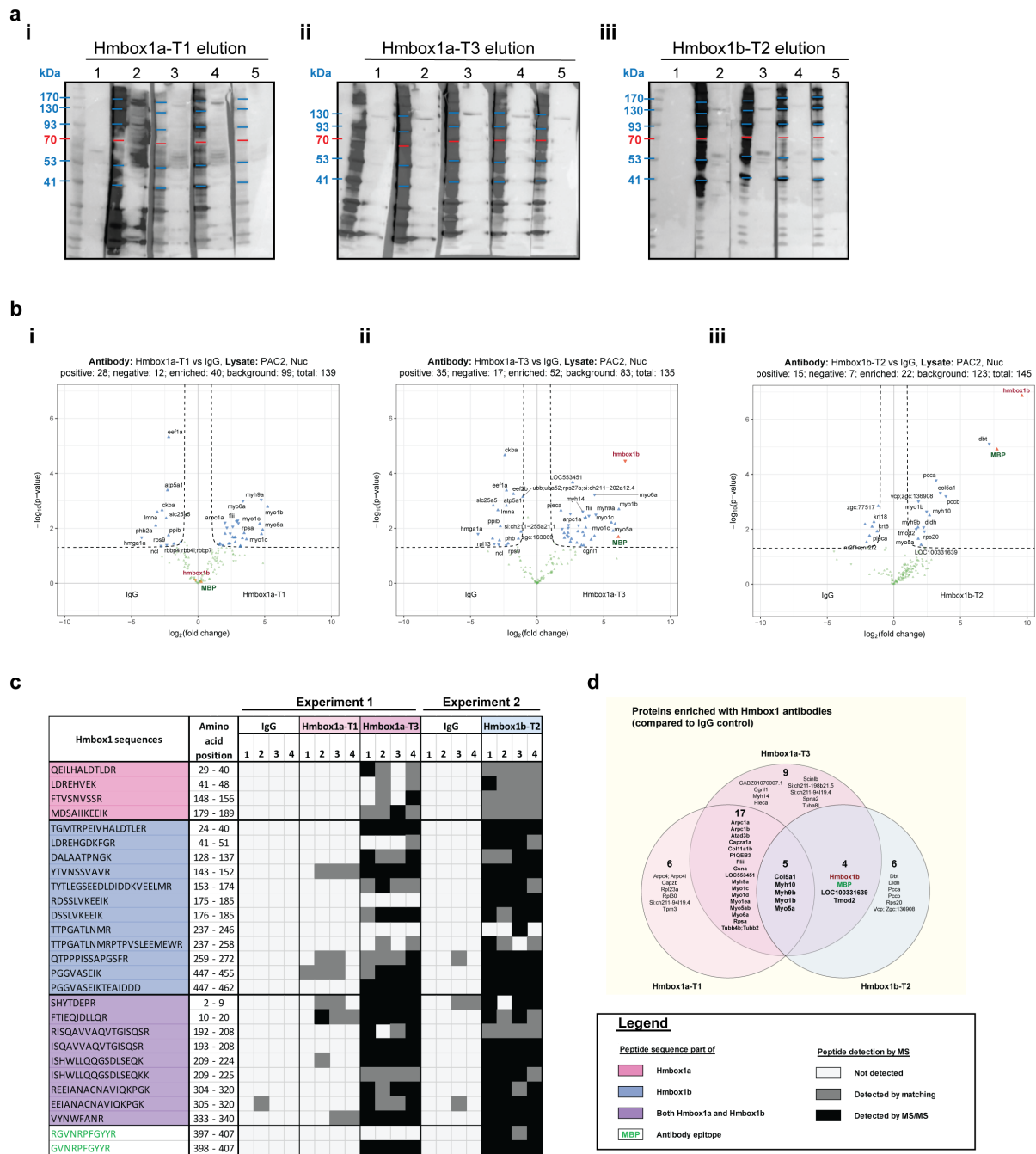


Figure 8. Validating antibodies on PAC2 zebrafish cells

a) Western blot analysis showing the first five elutions for each of the following purified antibodies: **i)** Hmbox1a-T1, **ii)** Hmbox1a-T3, and **iii)** Hmbox1b-T2. These antibodies were used as primary antibodies against whole cell lysates from the PAC2 cell line. The expected molecular weights are 48-49 kDa for Hmbox1a and 51 kDa for Hmbox1b.

b) Volcano plots visualizing the results from the label-free quantification of proteins bound in the immunoprecipitation (IP) experiments using **i)** Hmbox1a-T1, **ii)** Hmbox1a-T3, and **iii)** Hmbox1b-T2 antibodies against IgG control, on nuclear extract of the PAC2 zebrafish cell line. The results are plotted with $\log_2(\text{fold change})$ on the x-axis and $-\log_{10}(\text{p-value})$ on the y-axis. The enrichment threshold was set at a fold change > 2 and a p-value < 0.05 (Welch's t-test) with $c = 0.05$. Proteins are annotated with their gene names. MBP protein, part of the epitope sequence, is labeled in green.

c) Hmbox1 peptide sequences detected by mass spectrometry in the IP from (b). The sequences were categorized into four groups: Hmbox1a paralogue-specific sequences (pink), Hmbox1b paralogue-specific sequences (blue), sequences shared between both paralogues (purple), and sequences from the Hmbox1 epitope used in antibody design (green).

d) Venn diagram showing the common proteins enriched by the antibodies compared to the IgG control.

In addition to Hmbox1b and MBP, the Hmbox1a-T3 and Hmbox1b-T2 antibodies also commonly enriched for myosin heavy chain (LOC100331639) and tropomyosin Tmod2 (Figure 8c). Interestingly, despite the low detection of Hmbox1a/b by the Hmbox1a-T1 antibody, it shared as many as 17 commonly enriched proteins with the Hmbox1a-T3 antibody, but none with Hmbox1b-T2 antibody.

In contrast to the IP-MS analysis performed by Kappei et al. ⁷⁷, which effectively enriched telomere-associated proteins such as the four core components of the H/ACA snoRNPs (DKC1, GAR1, NHP2, and NOP10) and the Ku70-Ku80 heterodimer proteins using the HMBOX1 antibody in human studies, none of these proteins were enriched by any of my antibodies. This discrepancy raises questions about whether zebrafish orthologs have the same binding partners and mechanisms for maintaining telomere homeostasis. It may be necessary to perform IP experiments on samples with high Hmbox1a/b expression to draw a definitive conclusion.

Nevertheless, this IP-MS result confirms that both Hmbox1a-T3 and Hmbox1b-T2 antibodies recognize Hmbox1b. However, it remains uncertain whether any of the antibodies specifically target the Hmbox1a. Further validation with a knockout model is needed to address this issue.

Creating CRISPR-Cas knockout PAC2 zebrafish cells

To validate the generated antibodies and investigate the role of Hmbox1a/b, CRISPR-Cas knockout cell lines and fish lines were developed concurrently. While generating CRISPR-Cas knockout fish lines takes approximately a year, creating knockout cell lines generally requires less time. The knockout cell lines can offer preliminary insights before the knockout fish are ready for experimental use.

Before transfecting the cells with a CRISPR-Cas9 plasmid, the transfection methods were tested. In the first transfection experiment, a GFP-expressing plasmid (pCS2-eGFP) was transfected into PAC2 cells using FuGENE® HD Transfection Reagent (Promega) at a 3:1 ratio of reagent to DNA, following a previously published protocol ¹³⁸. However, no GFP expression was observed under a fluorescence microscope at 24 hours post-transfection (hpt). The protocol was subsequently repeated with adjustments to optimize the DNA-to-transfection reagent ratios, but GFP expression still could not be detected. This lack of success was likely due to the low transfection efficiency rate of 5 to 10%, as reported by the authors ¹³⁸.

Apart from FuGENE® HD transfection, another study reported that Nucleofector solution (Lonza) achieves a higher transfection efficiency rate of 40 to 50% for PAC2 cells ¹³⁹. In this study, two different nucleofection settings were tested with two types of empty plasmids (pCS2-eGFP and PX459, the latter being a gRNA-Cas9 plasmid): T027, optimized for PAC2 ¹³⁹, and X001, a classic setting. Similar to the results with FuGENE® HD transfection, no visible GFP expression was detected under a fluorescence microscope at 24 hpt in cells transfected with the pCS2-eGFP plasmid using either nucleofection setting. However, when Western blot analysis was conducted on these transfected cells, very weak signals for both GFP and Cas9 were detected in cells from both nucleofection conditions (Figure 9a). This suggests that while

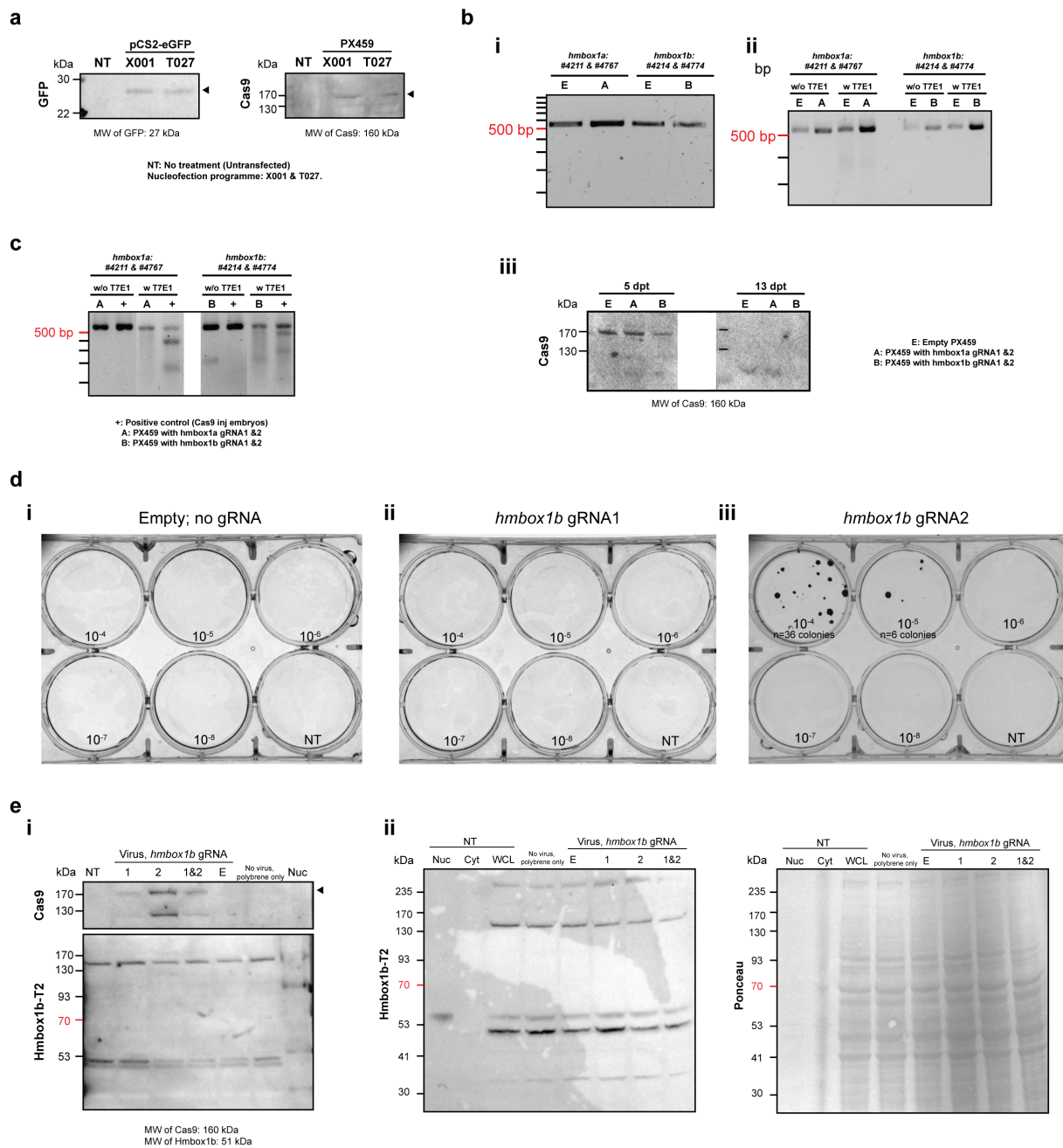


Figure 9. Creating CRISPR-Cas knockout PAC2 cells

a) Western blot analysis performed on PAC2 cells transfected with empty vector of either pCS2-eGFP (top) or PX459 (bottom) using Nucleofector solution (Lonza) with two different nucleofection settings (T07 and X001). Cells were harvested at 1 day post transfection (dpt) and the whole-cell lysates were probed against GFP and Cas9, respectively.

bi) Genotyping result of PAC2 cells co-transfected with two PX459 vectors (gRNA 1 and 2) for each gene target (*hmbbox1a* and *hmbbox1b*). Cells were harvested at 5 dpt and genotyped using primer pairs denoted in Figure 2.1a, ii) shows the bands after T7E1 treatment. **iii)** Western blot analysis performed on cells harvested at 5 dpt and 13 dpt, and probed against Cas9.

c) T7E1 assay performed on puromycin-selected PAC2 cells that were transfected with the gRNA-PX459 vectors. The positive controls were embryos injected with the same gRNA.

d) Plaque assay performed on HEK293T cells transduced with **i**) no gRNA (control), **ii**) *hmbbox1b* gRNA 1, and **iii**) *hmbbox1b* gRNA 2 lentivirus after puromycin selection. The virus dilutions were written at the bottom of each well. NT: No treatment.

e) Western blot analysis performed on PAC2 cells transduced with lentivirus carrying *hmbbox1b* gRNA(s), empty vector (E), or no treatment (NT). **i**) Cells were harvested at 2 dpt and the whole-cell lysates were probed against Cas9 (top) and with anti-Hmbox1b-T2 antibody (bottom). Nuclear extract (Nuc) from non-treated PAC2 was loaded for comparison. **ii**) Cells were again harvested at 12 dpt, without puromycin selection. The gel was also loaded with nuclear extract (Nuc), cytoplasmic extract (Cyt) and whole-cell lysate (WCL) from a non-treated PAC2 cells for comparison. (Left blot) Probed with anti-Hmbox1b-T2 antibody. (Right blot) Same blot with ponceau staining, serves as protein loading control.

the transfection was successful, the efficiency was low, which explains the poor visibility of GFP in fluorescence microscopy.

Next, PAC2 cells were co-transfected with a pair of PX459 plasmids, each carrying gRNA to target specific genes, using the T027 nucleofection setting. The gRNAs were designed to target the most upstream coding exon(s) of each gene. At 5 dpt, the cells were harvested and genotyped using primers flanking the deletion sites (see Figure 5a). However, no band shift was observed in the PCR analysis of the gRNA-PX459-transfected cells (Figure 9bi).

Subsequently, a T7 endonuclease I (T7E1) assay was performed to detect any small indels. The amplicons were melted and slowly re-hybridized before treatment with T7E1, but no band cleavage was observed (Figure 9bii). To confirm the success of the transfection, Western blot analysis was conducted, revealing Cas9 expression at 5 dpt, which was lost by 13 dpt (Figure 9biii). These results indicated successful nucleofection but also highlighted the transient nature of plasmid expression. Despite the presence of Cas9 expression, gene editing appeared to occur at an undetectable rate, as no band shifts were observed in the genotyping and T7E1 assays.

To address the low transfection efficiency issue, the experiment was repeated with puromycin to select for cells that were successfully transfected. Similar to previous experiments, the T7E1 assay yielded negative results. No lower bands were observed in the T7E1 assay performed on gRNA-PX459-transfected cells, despite successful cleavage being detected in embryos injected with the same gRNA sequences (Figure 9c). It appears that this transfection method was not efficient for generating a knockout cell line.

Next, a lentiviral transduction method was explored. Due to the large quantity of materials required for virus production, the experiment was focused on *Hmbox1b*, as it had been successfully detected in previous telomere pull-down and IP-MS experiments. Additionally, the Hmbox1b-T2 antibody appears to be more promising and could be useful for cross-validation.

In this experiment, lentiviruses carrying *hmbox1b* gRNA 1 and 2, as well as an empty vector, were produced. However, the virus titers were very low based on the plaque assay performed on transduced HEK293T cells. None of the virus-infected cells survived puromycin selection, except for the two wells transduced with 10^{-4} and 10^{-5} dilutions of the *hmbox1b* gRNA 2 virus, which showed 36 and 6 colonies, respectively (Figure 9d). The calculated virus titer for *hmbox1b* gRNA 2 was approximately 4×10^5 Infectious Units/mL.

Nevertheless, transduction attempts were made using the produced lentivirus on PAC2 cells despite the low titers. To validate the success of the transduction, Western blot analysis was performed on cells 2 dpt to detect Cas9 (Figure 9ei). Cas9 expression was observed in cells infected with gRNA1, even though the plaque assay showed no colonies. In contrast, Cas9 was not detected in cells infected with the empty vector virus, which was consistent with the plaque assay results showing no colonies. The cells were allowed to recover from the transduction without puromycin treatment. When cells were harvested again at 12 dpt and probed with the Hmbox1b-T2 antibody, no significant differences in band patterns or Hmbox1b expression levels were observed between cells transduced with *hmbox1b* gRNAs and the non-

transduced control cells (Figure 9eii). This lack of difference may be attributed to the low virus titer.

Unfortunately, CRISPR-Cas9 knockout cell lines for *hmbox1a* and *hmbox1b* could not be established despite numerous attempts. Although PAC2 cells are well-documented for transfection, they still require extensive optimization. Due to the lengthy and challenging nature of this process, the study focus shifted to fish knockout lines, which have shown greater success.

Creating CRISPR-Cas knockout zebrafish lines

To create a CRISPR knockout zebrafish line, the same gRNA sequences used in the cell culture work were employed, but with redesigned overhangs. New gRNA oligos were cloned into the pDR274 plasmid and *in vitro* transcribed into the gRNA scaffolds (Figure 10). The embryos were co-injected with a pair of gRNAs and Cas9 protein at 1-cell stage, which were then grown into adults. These adults were out-crossed with wild-type zebrafish, and their offspring (F1) were screened for deletions in the respective gene locus.

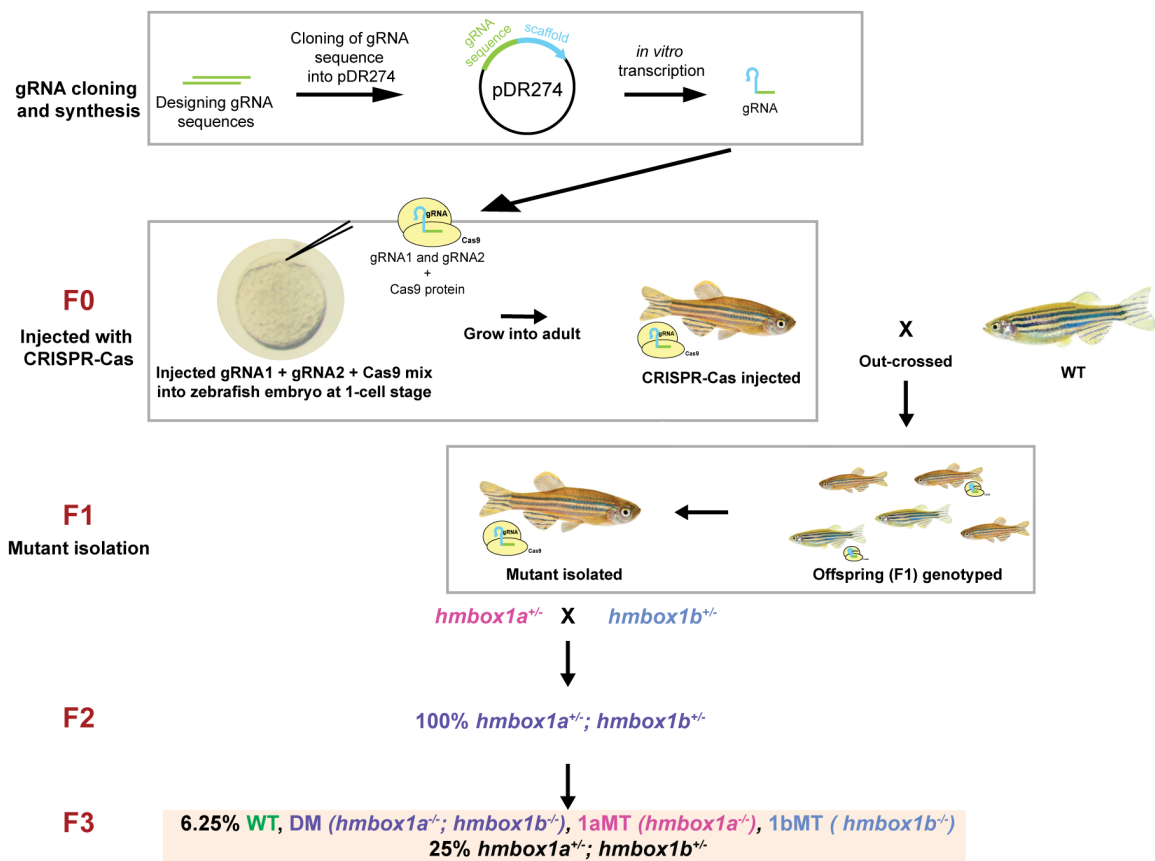


Figure 10. Generating CRISPR-Cas knockout fish line

Guide RNAs (gRNAs) were cloned into the pDR274 vector and *in vitro* transcribed along with the gRNA scaffold. The synthesized gRNAs and Cas9 protein were co-injected into 1-cell stage embryos. These embryos were raised to adulthood and out-crossed with wild-type zebrafish. Their offspring were screened for deletions, and individuals with the desired mutations were isolated. The F1 $hmbox1a^{+/}$ fish were then crossed with $hmbox1b^{+/}$ fish to produce F2 progeny with the $hmbox1a^{+/}; hmbox1b^{+/}$ genotype. These F2 progeny were in-crossed to generate homozygous mutants.

For the *hmbox1a* gene, a pair of gRNAs was designed to target its second exon. Out of 104 injected embryos, 52 survived to adulthood. Of these, 3 founders were identified with successful germline transmission. Ultimately, only two mutant lines were established, primarily due to the ease of genotyping and as their mutation resulted in a premature stop codon. The mutants were named as either large deletion or small deletion based on the deletion size.

The *hmbox1a* large deletion (*hmbox1a^{LD}*) mutant has a 780 bp deletion in the genome (Figure 11ai). Due to this large deletion, a three-primer system was adopted for a more accurate genotyping, with one primer annealing to the sequence deleted in the mutants. This produced a PCR product of 656 bp for the wild-type allele, while the mutant allele resulted in a 234 bp product (Figure 11bi). The deletion occurs between the intron preceding exon 2 and the gRNA 2 target site (Figure 11ci), leading to the loss of exon 2 and the start codon in the mRNA transcript. Theoretically, this mutation should prevent the translation of the Hmbox1a protein. However, there is an ATG sequence at the start of exon 3 that could seamlessly replace the original start codon since the sequence before the deleted start codon remained unchanged in the mutant mRNA, and presumably the promoter region too. This may serve as an alternative start codon, potentially leading to the production of an in-frame, N-terminally truncated Hmbox1a mutant protein. If this occurs, the resulting mutant Hmbox1a protein would still possess the homeobox domain and the C-terminus, thereby retaining telomere-binding activity and remaining detectable by the Hmbox1a antibody.

As a backup, another mutant line, *hmbox1a^{SD}*, was established. This mutant features a smaller 214 bp deletion and was genotyped using a different pair of primers (Figure 11aii). The wild-type allele produced a PCR product of 522 bp, while the mutant allele produced a 300 bp product (Figure 11bii). This deletion occurs between the two gRNA target sites and results in a frameshift mutation. However, a stop codon is introduced only 28 amino acids downstream of the deletion site, leading to the production of a truncated Hmbox1a mutant protein that is 78 amino acids long (9.2 kDa) (Figure 11cii), that would not be detectable by the Hmbox1a antibody.

For *hmbox1b*, a pair of gRNAs was designed to target the second and third exons of the gene. Out of the 125 injected embryos, 76 survived to adulthood. Of these, 7 founders were identified with successful germline transmission. Similar to *hmbox1a*, only two mutant lines (large deletion and small deletion mutants) were established.

The *hmbox1b* large deletion (*hmbox1b^{LD}*) mutant features an 869 bp deletion in the genome (Figure 12ai). This deletion was genotyped using a three-primer system, which generated a 546 bp PCR product for the wild-type allele and a 181 bp PCR product for the mutant allele (Figure 12bi). The deletion begins at the gRNA1 target site and extends through the introns following exon 4 (Figure 12ci). This results in a 448 bp deletion from the mRNA transcript, effectively removing portions of exons 2 through 4. Consequently, the frameshift mutation leads to the translation of a truncated Hmbox1b mutant protein that is 30 amino acids long (3.6 kDa).

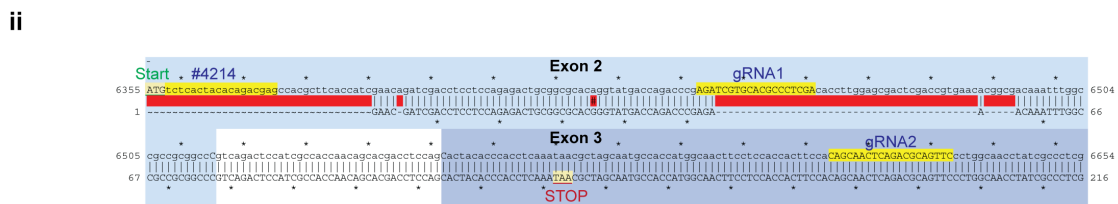
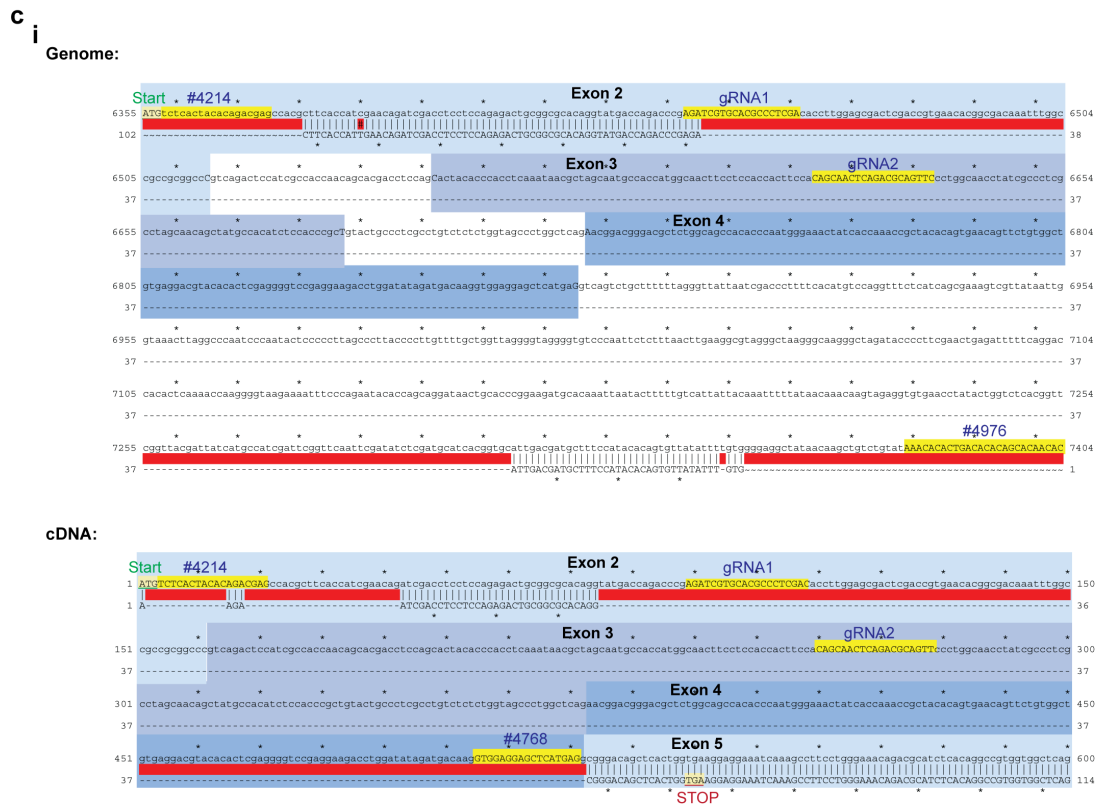
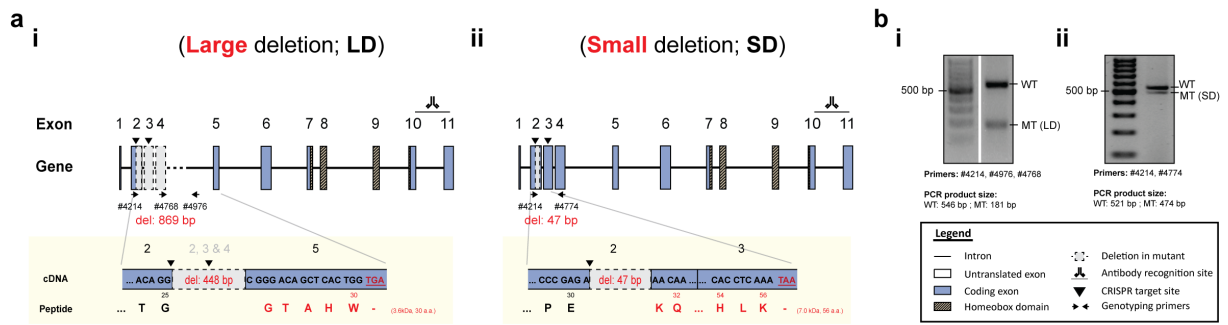


Figure 12. Creating *hmbox1b* CRISPR-Cas knockout fish line

Two *hmbox1b* mutant lines were created, they are named i) large deletion and ii) small deletion according to the deletion size.

a) Schematic diagram of the *hmbox1b* gene showing the deletion site at the genome level and the genotyping strategy. The resultant mutant cDNA sequence and its corresponding amino acid sequence was written below.

b) Genotyping result using the respective primers mentioned in (a). WT: Wild-type; MT: Mutant.

c) Sequencing result of the mutant aligned to the wild-type sequence. Features marked on the sequences include the exon numbers, gRNA target sites, primer annealing sites, start and stop sites. For the large deletion (i), both the genome (top) and cDNA (bottom) were sequenced, while for the small deletion (ii), only the genome was sequenced.

On the other hand, for *hmbox1b* small deletion (*hmbox1b^{SD}*), 47 bp was deleted between the gRNA target sites (Figure 12aⁱⁱ). Genotyping with a pair of primers resulted in a 521 bp PCR product for the wild-type allele, while the mutant allele produced a 474 bp product (Figure 12bⁱⁱⁱ). This deletion leads to a frameshift mutation that encounters the stop codon 26 amino acids after the deletion site, similar to *hmbox1a^{SD}*. Consequently, this would give rise to a 56 amino acid long Hmbox1b mutant protein (7.0 kDa) (Figure 12cⁱⁱ).

The individual F1 *hmbox1a^{+/-}* and *hmbox1b^{+/-}* mutants were crossed with each other to generate F2 (*hmbox1a^{+/-}; hmbox1b^{+/-}*) heterozygotes (Figure 10). These F2 heterozygotes were then incrossed to generate offspring with a 1/16 probability for each genotype of interest: wild-type (*hmbox1a^{+/+}; hmbox1b^{+/+}*), double mutant (*hmbox1a^{-/-}; hmbox1b^{-/-}*), *hmbox1a* mutant (*hmbox1a^{-/-}; hmbox1b^{+/+}*), and *hmbox1b* mutant (*hmbox1a^{+/+}; hmbox1b^{-/-}*). The study primarily focuses on large deletion mutants, with the smaller deletion mutants bred and maintained as backups.

Validating CRISPR-Cas knockout fish lines and antibodies using IP-MS and Western blot analysis

To check whether Hmbox1a and Hmbox1b are detectable in zebrafish embryos, the IP experiment was first tested on the nuclear extract of wild-type 2 dpf larvae using Hmbox1a-T3 and a newly purified Hmbox1b-T2 polyclonal antibody from the Protein Production Core Facility (PPCF), which will hereafter be referred to as Hmbox1b-T2 (M). In this experiment, the Hmbox1a-T1 antibody was not tested because it had shown inferior performance compared to the Hmbox1a-T3 antibody in previous IP-MS experiments conducted on PAC2 cells (Figure 8c).

The IP-MS analysis in 2 dpf embryos revealed the presence of Hmbox1a/b (Figure 13aⁱ and aⁱⁱ). However, similar to the telomere pull-down performed on BRF41 lysate and the IP-MS on PAC2 lysate, only the Hmbox1b paralogue was successfully detected. Of the 10 Hmbox1a/b peptide sequences detected, 4 were unique to Hmbox1b while 5 were common to both paralogues, and 1 corresponds to the epitope sequence (Figure 13b). No peptides unique to Hmbox1a were detected to suggest its presence. Despite the antibody re-purification effort, the MBP protein was still detectable in the IP with the Hmbox1b-T2 (M) antibody. It was further confirmed that the epitope protein is present in the purified antibody aliquot itself (Figure 14a, right-most column). Therefore, in the subsequent analysis, the Hmbox1a/b sequence corresponding to the epitope sequence will be excluded from any considerations related to Hmbox1a/b detection.

Some of the proteins enriched in the IP-MS analysis of 2dpf embryos were also previously enriched in PAC2 cells by Hmbox1a-T3 and Hmbox1b-T2 (M) antibodies (Figure 8d). This includes Hmbox1b, Myh10, and myosin heavy chain protein (LOC100331639) (Figure 13c). Besides that, Myh9b, and mitochondrial enzymes Dbt and Pcca in both 2 dpf and PAC2 lysates by Hmbox1b-T2 (M) antibody. On the other hand, Frem2a (involved in fin morphogenesis) and Plecb (predicted ankyrin-binding and structural molecule activity) were enriched with both antibodies only in 2 dpf samples but not PAC2. Similar to IP-MS findings in PAC2, most of the enriched proteins in 2 dpf were structural proteins that are related to skeletal and cardiac

muscles, and calcium homeostasis. Apart from it, non-structural proteins such as peroxiredoxins (Prdx1 and Prdx2) and ribosomal protein S26-like (Rps26l) were also enriched.

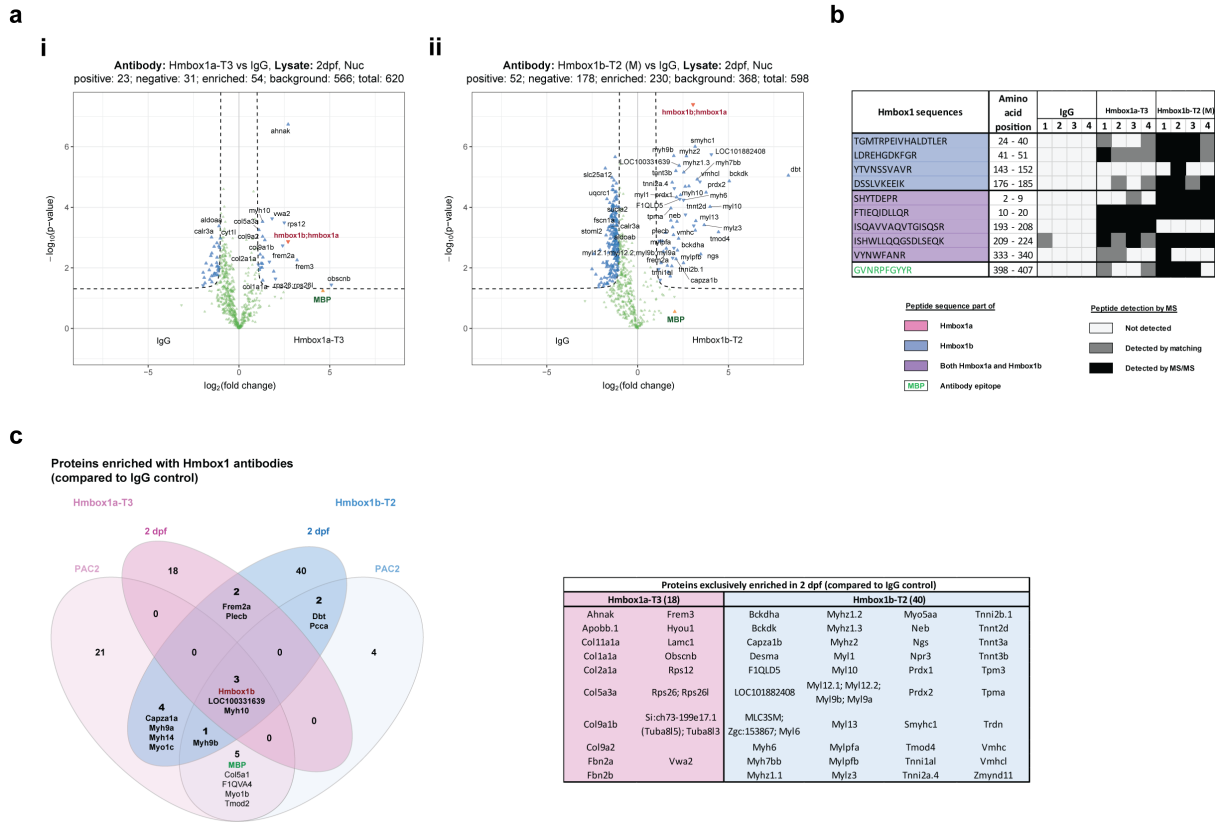


Figure 13. IP-MS on 2 dpf wild-type embryos

a) Volcano plots of IP experiments using **i)** Hmbox1a-T3 and **ii)** Hmbox1b-T2 (M) antibodies compared against an IgG control on nuclear extracts from wild-type 2 dpf embryos. The results are plotted with $\log_2(\text{fold change})$ on the x-axis and $-\log_{10}(p\text{-value})$ on the y-axis. The enrichment threshold was set at a fold change > 2 and a p-value < 0.05 (Welch's t-test) with $c = 0.05$. Proteins are annotated with their gene names, and the MBP protein is an exogenous protein and part of the epitope sequence.

b) The Hmbox1 peptide sequences detected by mass spectrometry from the IP-MS in (i) and (ii).

c) Venn diagram comparing the commonly enriched proteins identified by IP-MS analysis of the nuclear extract from PAC2 cells (Figure 8d) with those from 2 dpf wild-type embryos, using Hmbox1a-T3 and Hmbox1b-T2 antibodies. The table on the right lists the 18 and 40 proteins exclusively enriched in 2 dpf embryos by the respective antibodies.

Given the presence of Hmbox1b in 2 dpf larvae, IP-MS experiments were repeated using nuclear extracts from homozygous mutants ($(hmbox1a^{LD/LD}; hmbox1b^{LD/LD})$, $hmbox1a^{LD/LD}$, and $hmbox1b^{LD/LD}$). However, due to insufficient material, the experiments were conducted with fewer replicates and reduced protein amounts per IP. Consequently, this seems to have compromised the results as only one common sequence between Hmbox1a and Hmbox1b, other than the epitope, was detected (Figure 14a). This common sequence was detected in wild-type lysates by both the Hmbox1a-T3 and Hmbox1b-T2 (M) antibodies but not in $(hmbox1a^{LD/LD}; hmbox1b^{LD/LD})$ mutants or the IgG control, as expected.

To determine whether the antibodies are paralogue-specific, their ability to detect Hmbox1a/b was examined in the individual single mutants. If the Hmbox1a-T3 antibody is specific to the Hmbox1a paralogue, Hmbox1a/b would still be detected in the $hmbox1b^{LD/LD}$ mutant but not in the $hmbox1a^{LD/LD}$ mutant, and vice versa, assuming both paralogues are expressed in 2 dpf larvae. However, the results showed that the Hmbox1a-T3 antibody continued to detect Hmbox1a/b in both single mutants, while the Hmbox1b-T2 (M) antibody did not detect

Hmbox1a/b in either of the single mutants. This conflicting result raises uncertainty about whether the antibodies are specific to individual paralogues and if Hmbox1b is the only paralogue expressed at 2 dpf, as suggested by the IP-MS results in Figure 13b.

When comparing this IP-MS result to the previous one (Figure 13) to identify any reproducible hits from IP using the respective antibodies versus the IgG control in wild-type, very few proteins were commonly enriched across both rounds of IP (Figure 14b). No proteins were consistently enriched across all four experiments. Hmbox1b was detected in all experiments except in the repeated Hmbox1a-T3 antibody IP. However, the IPs using the Hmbox1b-T2 (M) antibody have more common hits between the two rounds of experiments. This includes mitochondrial enzymes like Bckdha, Bckdk, and Dbt, along with a structural protein, Notochord Granular Surface (Ngs). However, these proteins did not appear among those differentially enriched in wild-type when comparing IP results with the (*hmbox1a^{LD/LD}*; *hmbox1b^{LD/LD}*) double mutant for the respective antibodies (Figure 14ci and cii). Instead, proteins such as Rps15a, peroxidase (Prdx1; Prdx2), and Col1a1a were commonly enriched by both antibodies in wild-type sample but not in double mutant sample.

To investigate whether Hmbox1a and Hmbox1b are expressed at other developmental stages and to determine if any of the identified hits were consistent across these stages, IP-MS analysis was extended to 1 dpf and 3 dpf larvae. To minimize material usage for these experiments, only the Hmbox1b-T2 (M) antibody was used on nuclear extracts from wild-type and double mutant larvae.

The IP analysis at 1 dpf (Figure 15a) showed nearly no detection of Hmbox1a/b, with only one shared sequence identified in 1 out of 4 wild-type samples and none in the (*hmbox1a^{LD/LD}*;

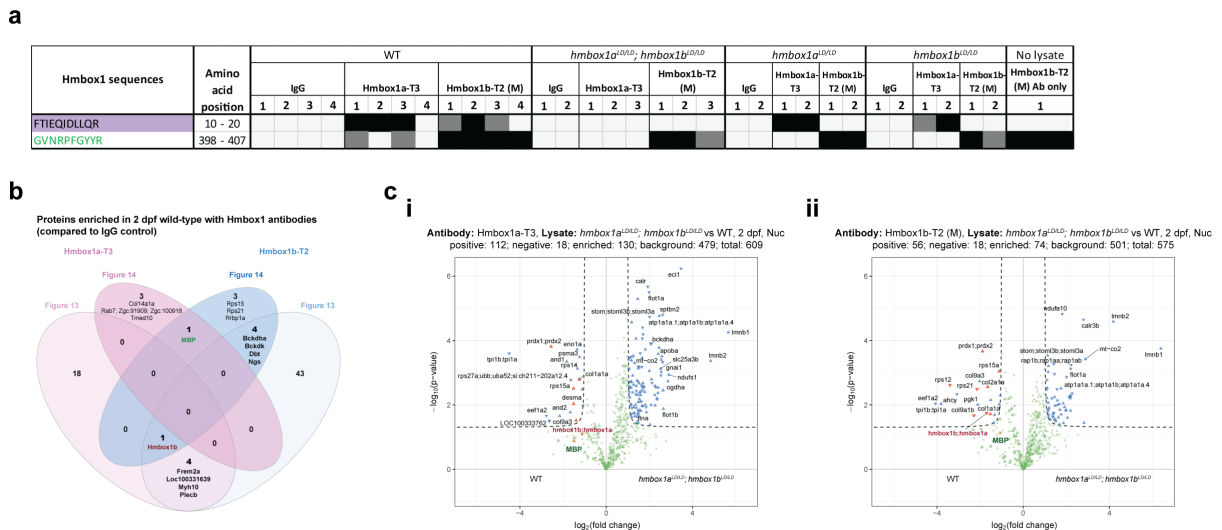


Figure 14. IP-MS on 2 dpf (*hmbox1a^{LD/LD}*; *hmbox1b^{LD/LD}*) double mutant embryos

a The Hmbox1 peptide sequences identified by mass spectrometry from IP performed on nuclear extracts of double mutant (*hmbox1a^{LD/LD}*; *hmbox1b^{LD/LD}*) and wild-type (WT) 2 dpf embryos using Hmbox1a-T3 and Hmbox1b-T2 (M) antibodies.

b Venn diagram comparing the commonly enriched proteins identified by IP-MS analysis using Hmbox1a-T3 and Hmbox1b-T2 antibodies against the IgG control in 2 dpf wild-type embryos from Figure 13 with those from this repeated IP-MS.

c Volcano plots from IP experiments comparing the differential enrichment of proteins between wild-type and double mutant using i) Hmbox1a-T3 and ii) Hmbox1b-T2 (M) antibodies. The plots display $\log_2(\text{fold change})$ on the x-axis and $-\log_{10}(\text{p-value})$ on the y-axis. The enrichment threshold was set at a fold change > 2 and a p-value < 0.05 (Welch's t-test) with $c = 0.05$. Proteins are annotated by their gene names, and those highlighted in orange were also found to be enriched in one or more of the 2 dpf wild-type pull-downs in (b).

hmbox1b^{LD/LD}) double mutant samples or IgG control. In contrast, the 3 dpf IP analysis revealed a notably higher number of Hmbox1a/b peptide sequences, including 4 unique Hmbox1b sequences and 5 shared sequences (Figure 15b). Similarly, none of these sequences were detected in the double mutants or IgG control, which reinforces the loss of Hmbox1b expression in the 3 dpf (*hmbox1a^{LD/LD}; hmbox1b^{LD/LD}*) larvae.

Despite the increased detection of Hmbox1b peptide sequences at 3 dpf, only 2 proteins were differentially enriched in the wild-type. Across all three developmental stages, only phosphoglycerate kinase 1 (Pgk1) was consistently enriched in the wild-type at both 1 dpf and

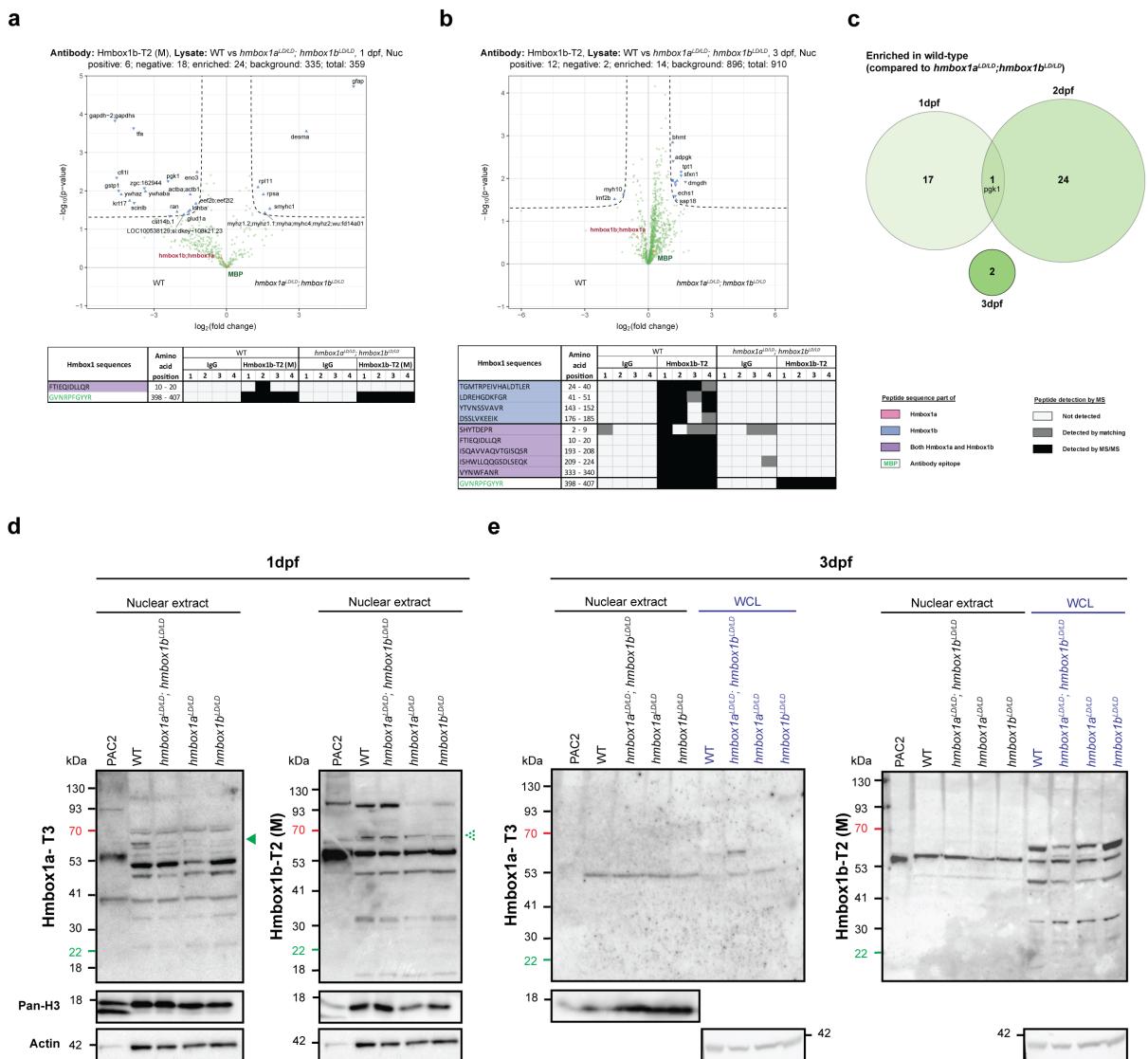


Figure 15 IP-MS and Western blot on 1 dpf and 3 dpf (*hmbox1a^{LD/LD}; hmbox1b^{LD/LD}*) double mutants
a, b Volcano plots illustrating the differential enrichment of proteins between wild-type and *hmbox1a^{LD/LD}; hmbox1b^{LD/LD}* double mutant IP experiments using the Hmbox1b-T2 (M) antibody on nuclear extracts from **a**) 1 dpf and **b**) 3 dpf samples. The plots display $\log_2(\text{fold change})$ on the x-axis and $-\log_{10}(\text{p-value})$ on the y-axis. The enrichment threshold was set at a fold change > 2 and a p-value < 0.05 (Welch's t-test) with $c = 0.05$. Proteins are annotated by their gene names. Below, the identified Hmbox1 peptide sequences are shown.
c Venn diagram comparing the commonly enriched proteins identified by IP-MS analysis in wild-type using the Hmbox1b-T2 (M) antibody compared to the respective *hmbox1a^{LD/LD}; hmbox1b^{LD/LD}* double mutant across 1 dpf, 2 dpf (from Figure 14c), and 3 dpf.
d, e Western blot analysis of **d**) nuclear extract from 1 dpf embryos and **e**) nuclear extract and whole cell lysate from 3 dpf larvae, probed with Hmbox1a-T3 and Hmbox1b-T2 (M) antibodies. Actin and pan-histone 3 (H3) were used as loading controls for total protein and nuclear protein, respectively.

2 dpf (Figure 15c). None of the potential hit proteins previously mentioned in Figure 14c were found to be enriched in the wild-type when compared to the double mutant.

Lastly, Western blot analysis was conducted on nuclear extracts from CRISPR knockout embryos and larvae. A band between the 53 and 70 kDa markers was observed in the Hmbox1a-T3 antibody blot for 1 dpf embryos, which decreased in intensity in the mutants compared to the wild-type (Figure 15d, green arrow). However, this band was not present in the PAC2 nuclear extract and slightly deviates from the expected molecular weight of Hmbox1a and Hmbox1b (49 to 51 kDa). Although a similar band was detected with Hmbox1b-T2 (M) antibody and appeared in the PAC2 nuclear extract, the signal loss was not as pronounced as with the Hmbox1a-T3 antibody.

This band of interest was not observed in any nuclear extract or whole-cell lysate from 3 dpf larvae samples when probed with Hmbox1a-T3 antibody (Figure 15e). Similarly, it was absent in the nuclear extracts probed with Hmbox1b-T2 (M) antibody. Although the band was visible in the whole-cell lysate, there was no notable difference in band intensity between wild-type and mutants. Aside from this mentioned band, no significant differences in banding patterns were observed between wild-type and mutants in the Western blots for both 1 dpf and 3 dpf lysates.

Immunofluorescence (IF) staining on 1 dpf embryos

After the successful detection of Hmbox1b in embryos at 1 to 3 days dpf using IP-MS, I proceeded with immunofluorescence (IF) staining on 1 dpf embryos. This stage was chosen because it is easier to stain and has a simpler anatomy compared to later developmental stages.

In the IF staining of wild-type 1 dpf embryos, nuclear staining was observed in two structures: one resembling the pronephros and the other the notochord (Figure 16a). The structure resembling the pronephros is located above the yolk sac extension and runs along the anterior-posterior axis of the embryo, while the notochord structure is situated along the midline. These signals were absent in the (*hmbox1a^{LD/LD}; hmbox1b^{LD/LD}*) double mutants (Figure 16b). To confirm which paralogue was expressed and responsible for the staining, the IF was also performed on single mutants. The signals remained in the *hmbox1a^{LD/LD}* mutant but were absent in the *hmbox1b^{LD/LD}* mutant, confirming that the observed signals in the pronephros and notochord were due to Hmbox1b expression.

When the experiment was repeated with the small deletion mutants, similar results were observed: signal was present in the *hmbox1a^{SD/SD}* mutant but absent in the (*hmbox1a^{SD/SD}; hmbox1b^{SD/SD}*) and *hmbox1b^{SD/SD}* mutants (Figure 16b). The absence of staining in both the large and small deletion mutants of *hmbox1b* confirms the effectiveness of the CRISPR knockouts in both mutant lines. This further supports that the Hmbox1b-T2 (M) antibody recognizes Hmbox1b, as evidenced by the loss of staining in two independent *hmbox1b* mutant lines.

Since there was no discrepancy in the IF results between the *hmbox1b^{SD/SD}* and *hmbox1b^{LD/LD}* mutant lines, subsequent studies will focus on the *hmbox1b^{LD/LD}* line, which was used in the IP-MS experiments. Additionally, the consistent expression of Hmbox1b in the pronephros and notochord of both *hmbox1a* mutants (*hmbox1a^{LD/LD}* and *hmbox1a^{SD/SD}*) indicates that mutations in *hmbox1a* do not affect the expression or localization of Hmbox1b.

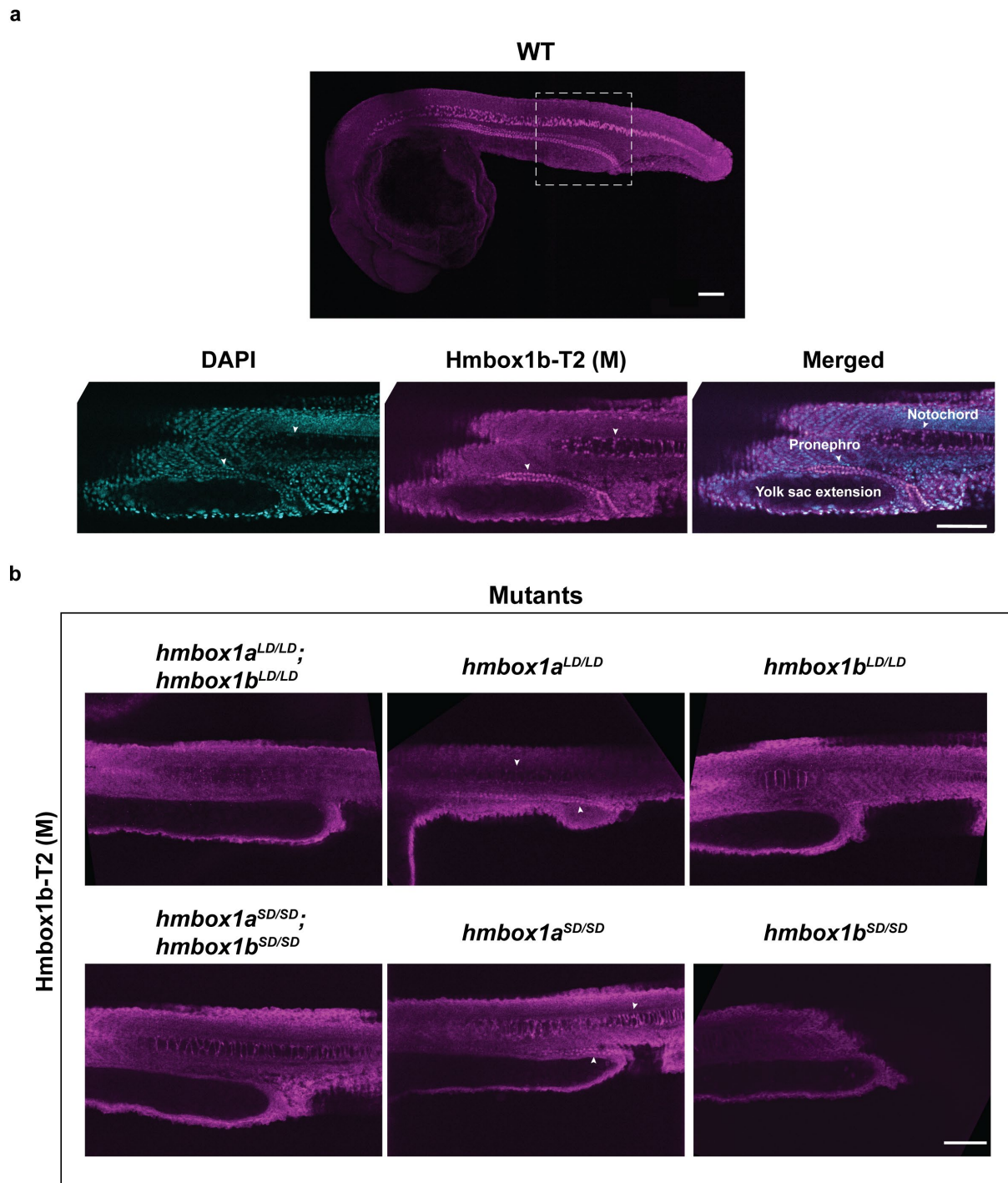


Figure 16. Immunofluorescence (IF) imaging of 1 dpf embryos stained with Hmbox1b-T2 (M) antibody

a) Maximum projection of an IF image of a wild-type 1 dpf embryo stained with the Hmbox1b-T2 (M) antibody, taken with a 10X objective. Below, a single plane, zoomed-in view of the trunk region is shown, co-stained with DAPI, along with a merged image.

b) IF images displaying the staining patterns in various *hmbox1a/b* CRISPR mutants.

All images were captured using a VisiScope confocal microscope at 20X magnification and presented as single-plane images, with scale bar = 100 μ m. White arrowheads indicate staining in the pronephros and notochord.

The Hmbox1a-T3 antibody was previously speculated to be a pan-Hmbox1a/b antibody. To verify if it also recognizes Hmbox1b, similar to what was observed in the IP-MS, staining was performed on 1 dpf embryos using this antibody. The results showed a consistent pattern with the Hmbox1b-T2 (M) staining: both the wild-type and *hmbox1a*^{LD/LD} mutants exhibited staining in the pronephros and notochord, while the staining was absent in the (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) and *hmbox1b*^{LD/LD} mutants (Figure 17, top). This indicates that the Hmbox1a-T3 antibody does indeed recognize Hmbox1b, further confirming Hmbox1b expression in the pronephros and notochord using an independently raised antibody.

Although the Hmbox1a-T3 antibody was raised against an Hmbox1a epitope, its recognition of Hmbox1b is likely due to shared sequence similarities, as previously discussed. However, this cross-reactivity suggests that the antibody should also recognize Hmbox1a. Interestingly, the absence of a unique staining pattern between wild-type and *hmbox1a*^{LD/LD} embryos suggests that Hmbox1a may not be expressed at this stage.

Since our lab had access to the HMBOX1 antibody used in human studies ⁷⁷, this antibody was also tested to determine whether the staining pattern in zebrafish could be recapitulated and to gain any additional insights. Given that zebrafish Hmbox1a and Hmbox1b share approximately 70% sequence identity with their human ortholog, it was anticipated that the human HMBOX1 antibody would exhibit cross-species reactivity.

As expected, the staining pattern was similar, with the pronephros being stained in wild-type and *hmbox1a*^{LD/LD} mutants, while the signal was lost in (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) and *hmbox1b*^{LD/LD} mutants (Figure 17, bottom). However, staining in the notochord was less prominent, and there was a noticeable higher background. Therefore, while the human HMBOX1 antibody demonstrated cross-species reactivity by detecting Hmbox1b in the pronephros, it was not as well stained in the notochord compared to the Hmbox1a-T3 and Hmbox1b-T2 (M) antibodies. Additionally, the staining did not provide any new insights, and due to the suboptimal quality of the results, the human HMBOX1 antibody was not used further in this study.

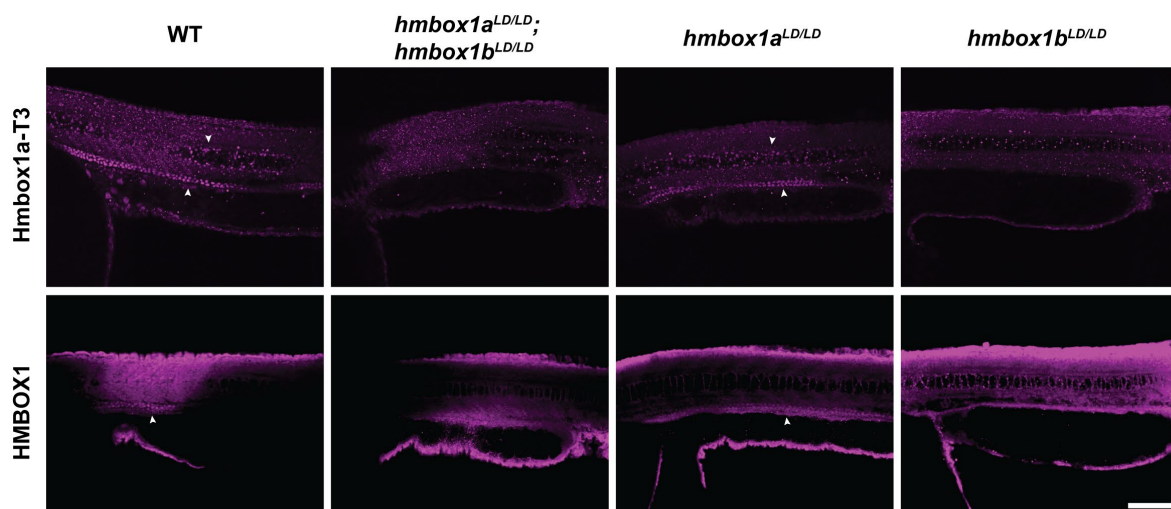


Figure 17. IF imaging of 1 dpf embryos stained with Hmbox1a-T3 and human HMBOX1 antibodies
IF images showing the trunk of wild-type and *hmbox1a/b* mutant 1 dpf embryos stained with Hmbox1a-T3 (top) and human HMBOX1 (bottom) antibodies. All images were captured using a VisiScope confocal microscope at 20X magnification and presented as single-plane images, with scale bar = 100 μ m. White arrowheads indicate staining in the pronephros and notochord.

So far, the localization of Hmbox1b in the pronephros has not been thoroughly validated. To confirm its spatial expression, counter-staining with pronephros markers was employed. This method provides a reliable way to image the pronephros, even in mutants where the Hmbox1b signal is absent, and also enables the assessment of potential malformations in the pronephros of *hmbox1b^{LD/LD}* mutants. To label the entire pronephros, an RNA FISH probe for cadherin 17 (*cdh17*) was designed¹⁴⁰. While *cdh17* covers most of the pronephros, it does not target podocytes¹⁴¹. Therefore, the mutants were crossed with the transgenic line *Tg(wt1b:eGFP)li1*¹⁴², which expresses GFP-tagged Wt1b (Wilms Tumor 1b) in podocytes. This approach ensures comprehensive coverage of the entire pronephros.

The staining was performed using Hmbox1b-T2 (M) antibody and RNA-FISH *cdh17* probe on 1 dpf wild-type and mutants with *Tg(wt1b:eGFP)* background. As expected, the Hmbox1b signal highly co-localized with the *cdh17* marker throughout the entire pronephros, extending to the distal end at the cloaca in wild-type, which such signal was absent in the mutants possessing the *hmbox1b^{LD/LD}* deletion (Figure 18). Yet, the pronephros structure in these mutants was not malformed; it appeared to have the same length and morphology as in the wild-type embryos. A closer examination revealed that no Hmbox1b signal was observed in the podocytes, where *cdh17* is not expressed and Wt1b is exclusively expressed (Figure 19). Additionally, staining in both the pronephros and notochord co-localized with DAPI, further highlighting its nuclear expression.

Next, to investigate whether the Hmbox1b signal co-localizes with the telomere foci, the 1dpf embryos were counterstained with Hmbox1b-T2 (M) antibody and a C-rich telomere probe. When examined under higher magnification, the Hmbox1b staining did not appear as distinct foci but rather seemed homogeneously distributed throughout the nucleus of the notochord and pronephros (Figure 20a and b, respectively). Consequently, the attempt to study co-localization of Hmbox1b with telomere was abolished.

So far, the primary paralogue detected has been Hmbox1b, with no evidence of Hmbox1a expression in the embryos, although it was detected in PAC2 cells via IP-MS (Figure 8c). The IF staining results with Hmbox1a-T3 antibody also indicated the absence of Hmbox1a expression at 1 dpf (Figure 17). To confirm this, an additional IF experiment was conducted using the Hmbox1a-T1 antibody. In agreement with the previous result, there was no staining observed in the wild-type or any of the mutants, nor was there any signal in the pronephros or notochord of any genotype (Figure 18). Notably, IF staining with both Hmbox1a-T1 and Hmbox1a-T3 antibodies on *hmbox1b^{LD/LD}* mutants consistently revealed that the absence of Hmbox1b expression did not lead to the induction of Hmbox1a expression in the pronephros and notochord, suggesting that Hmbox1a does not compensate for the loss of Hmbox1b.

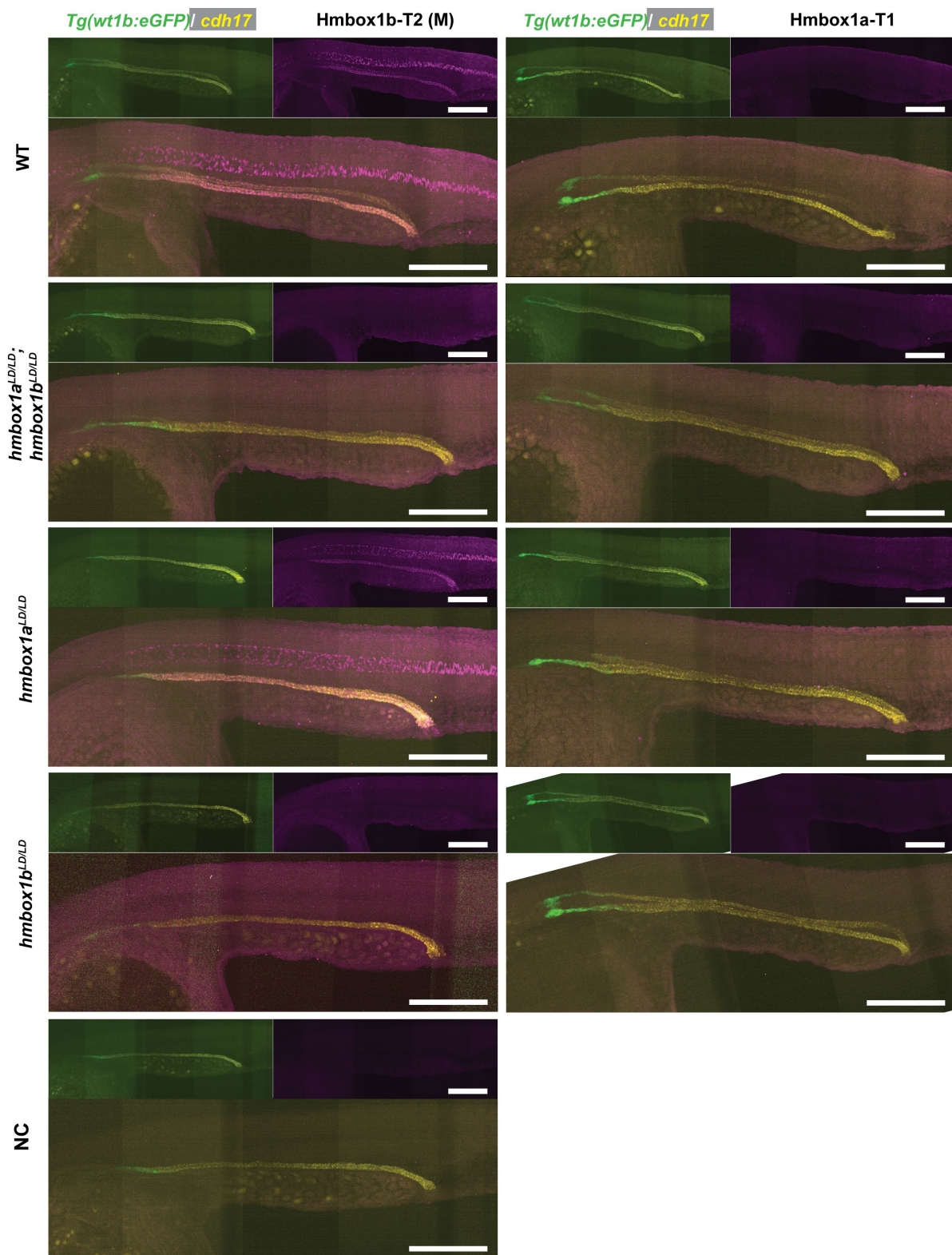


Figure 18. IF imaging of 1 dpf embryos stained with Hmbox1b-T2 (M) and Hmbox1a-T1 antibodies, and RNA probe *cdh17*

Maximum projections of IF images of 1 dpf wild-type and *hmbx1a/b* mutant embryos stained with the Hmbox1b-T2 (M) (left panel) and Hmbox1a-T1 (right panel) antibodies (in magenta). For negative control (NC), the embryo was not subjected to primary antibody staining. All the staining was performed on embryos with a *Tg(wt1b:eGFP)/li1* background (green) and counterstained with RNA probe *cdh17* (yellow) to label the pronephros. The merged channels are presented as a larger image below. Images were taken with a 20X objective and stitched using a BC43 confocal microscope (Andor). Scale bar = 200 μ m.

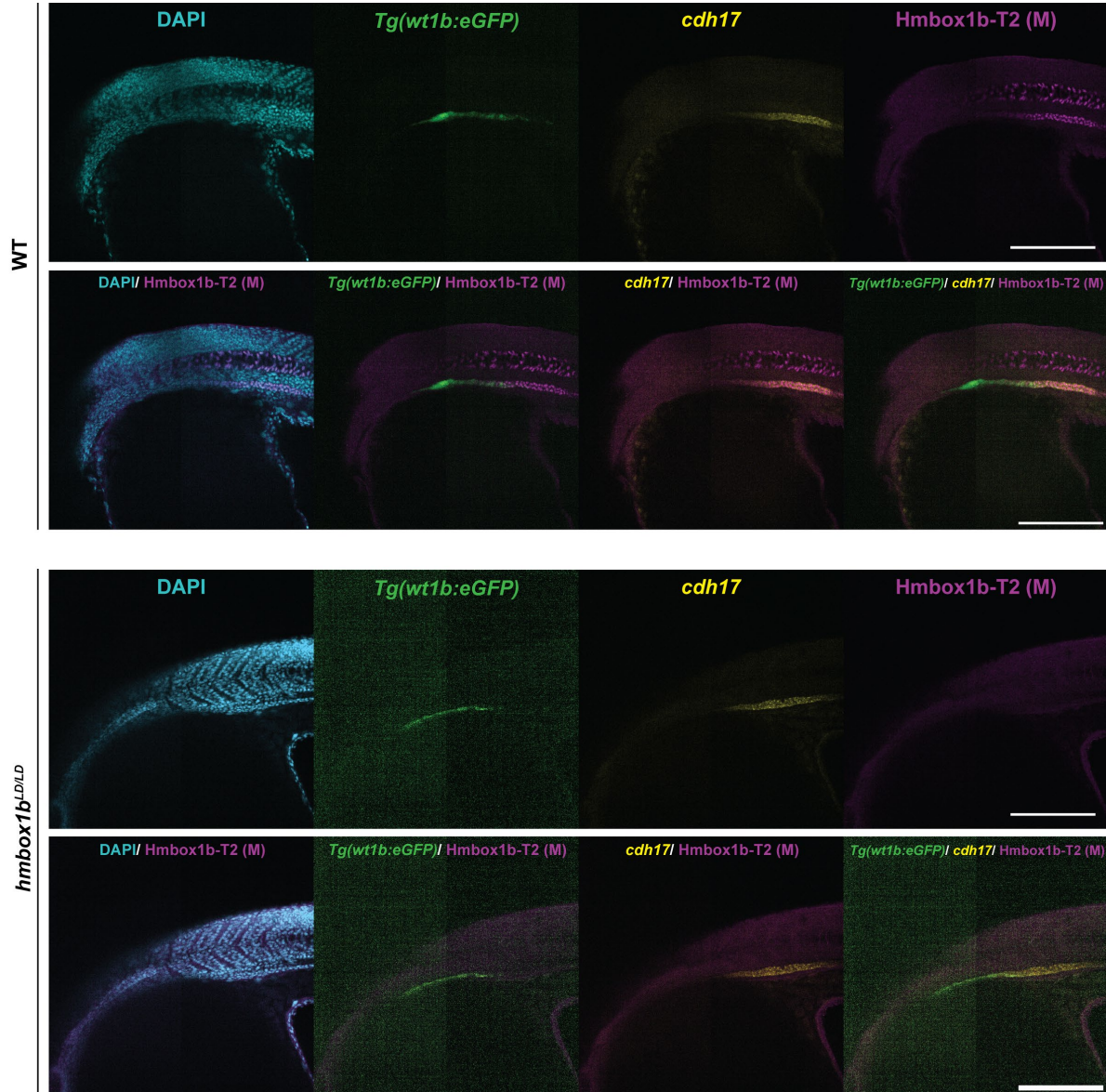


Figure 19. IF images focused on the proximal end of the pronephros in 1 dpf embryos

Maximum projections of IF images focused on the proximal end of the pronephros in 1 dpf wild-type and *hmbox1b^{LD/LD}* mutant embryos from Figure 18. The top row of each panel displays individual channels: DAPI nuclear staining (cyan), Wt1b expression from the *Tg(wt1b:eGFP)/li1* line (green), and RNA probe *cdh17* staining (yellow), both of which label the pronephros, along with Hmbox1b-T2 (M) antibody staining (magenta). The merged channels are presented in the bottom row, with the indicated channels combined. Images were acquired using a 20X objective on a BC43 confocal microscope (Andor). Scale bar = 200 μ m.

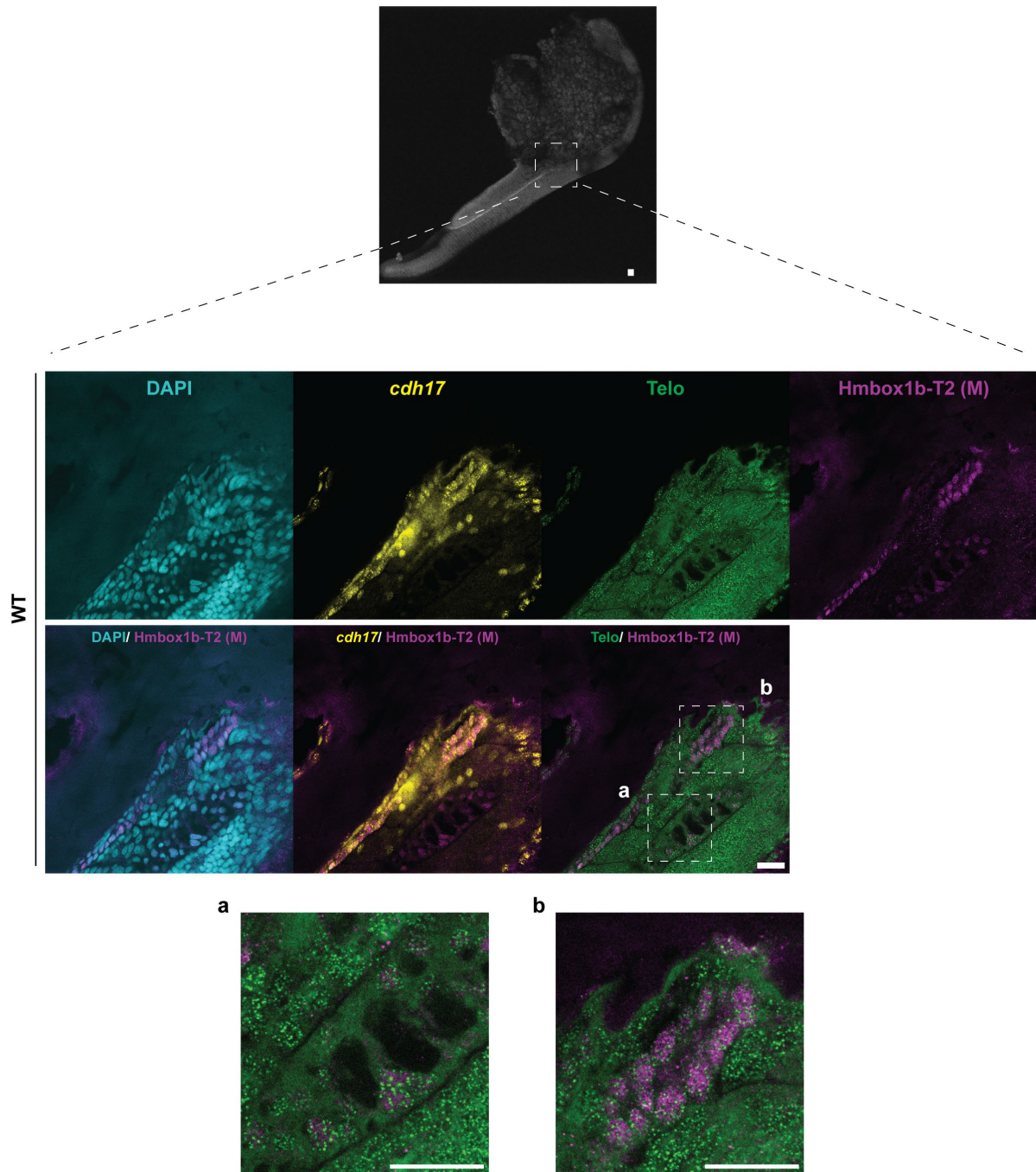


Figure 20. Telomere staining on 1 dpf wild-type embryo

A single-plane image of a 1 dpf wild-type embryo stained with a telomere PNA probe and Hmbox1b-T2 (M) antibody to assess the co-localization of Hmbox1b (magenta) with telomeres (green). The image was focused on two regions: **a**) the notochord and **b**) the pronephros, where Hmbox1b expression is present. Additionally, the embryos were stained with DAPI (cyan) to highlight the nucleus and with the *cdh17* RNA probe (yellow) to mark the pronephros. The top row presents single-channel images, while the bottom row shows merged channels with the indicated combinations. The images were acquired using a 60X objective on a VisiScope confocal microscope. Scale bar = 25 μ m.

IF staining on 4 hours post-fertilization (hpf) embryos

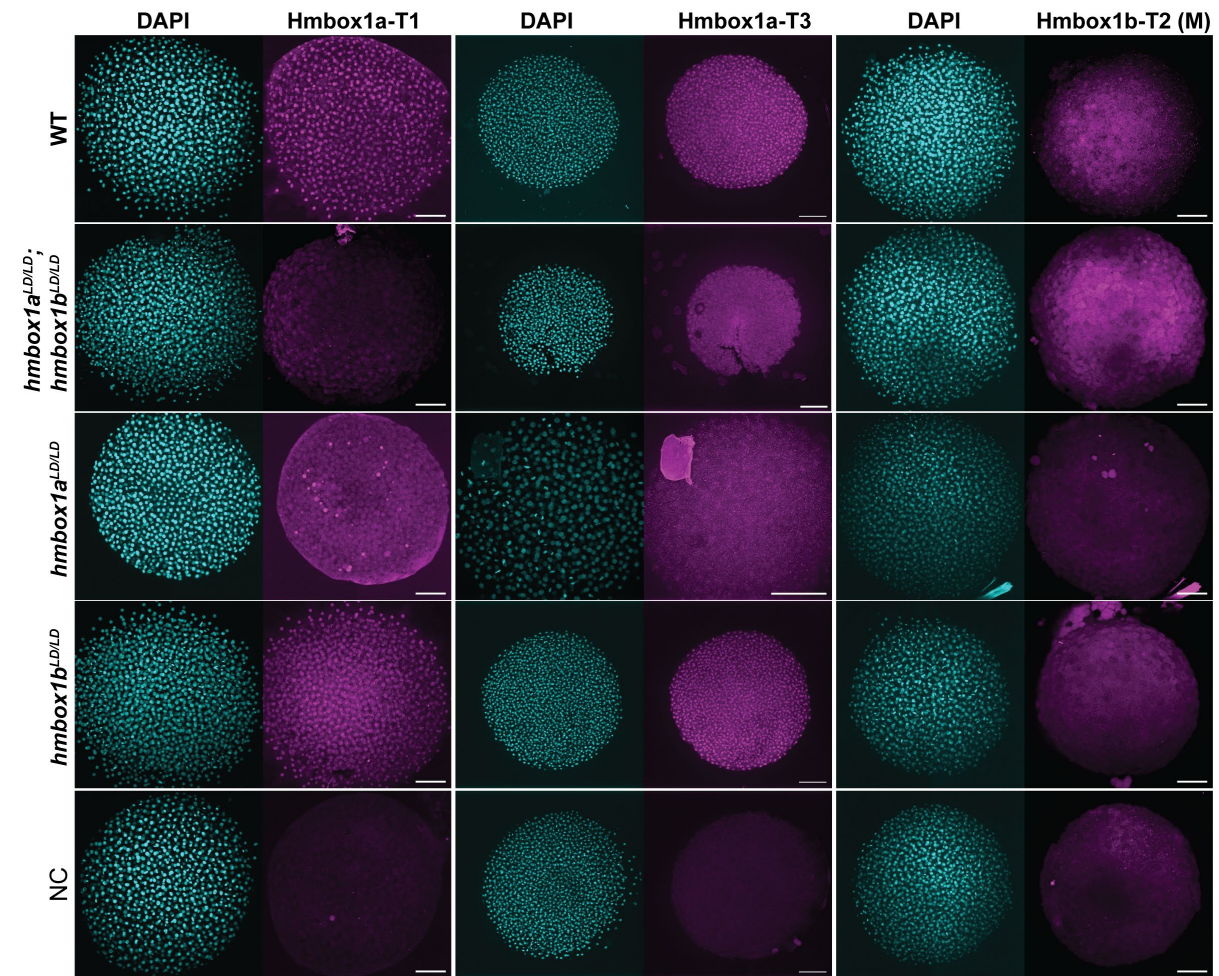
The lack of Hmbox1a detection with the antibodies raised concerns about whether any of the antibodies were successfully raised against Hmbox1a. This uncertainty also hindered the validation of the *hmbox1a* CRISPR knockout lines. Given the complexity of organismal development, it is possible that Hmbox1a is expressed at a specific and narrow time point during development that has not yet been explored in this study. According to the RNA-seq data on embryo development¹³⁵ (Figure 4), *hmbox1a* mRNA expression peaks between 2 to 6 hours post-fertilization (hpf), a stage that has yet to be investigated here.

To validate whether Hmbox1a is indeed expressed during this developmental window, IF experiments were conducted on embryos at sphere stage (4 hpf). The wild-type embryos exhibited nuclear staining when probed with either Hmbox1a-T1 or Hmbox1a-T3 antibodies (Figure 21a). Interestingly, a stronger Hmbox1a signal was observed in the nuclei that were in interphase (Figure 21bi) as compared to those in mitotic stages, such as metaphase (Figure 21bii) and anaphase (Figure 21biii). Such nuclear signal was also observed in *hmbox1b*^{LD/LD} mutants, but not in *hmbox1a*^{LD/LD} and (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutant embryos. Thus, this confirms the Hmbox1a expression at 4 hpf and validates the effectiveness of Hmbox1a-T1 and Hmbox1a-T3 antibodies in recognizing Hmbox1a.

When the embryos were probed with Hmbox1b-T2 (M) antibody, nuclear staining was not observed in the wild-type or any of the mutants (Figure 21a). This implies that the Hmbox1b-T2 antibody does not recognize Hmbox1a, and the loss of Hmbox1a is not compensated by Hmbox1b.

Overall, these IF staining results help to confirmed that Hmbox1a-T1 antibody recognizes Hmbox1a while the Hmbox1b-T2 (M) antibody recognizes Hmbox1b, and the Hmbox1a-T3 antibody is a pan-Hmbox1a/b antibody that recognizes both paralogues. Additionally, these results indicate that the *hmbox1a*^{LD/LD} CRISPR knockout was successful, and the mutant likely does not produce any Hmbox1a protein. Therefore, given the success of the *hmbox1a*^{LD/LD} CRISPR knockout, further testing of the backup *hmbox1a*^{SD/SD} line was not performed.

Next, to determine whether Hmbox1a localizes to the telomere, 4 hpf wild-type and (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutant embryos were co-stained with a telomere probe and the Hmbox1a-T1 antibody. Under 100X magnification, high background signals were observed in both wild-type and mutant embryos (Figure 2.18a). Although the co-localization of telomere foci and Hmbox1a signals was low in wild-type embryos (Figure 2.18bi and ii), it appeared higher compared to double mutant embryos (Figure 2.18biii and iv). Upon analyzing 7 images per genotype using the ImageJ plugin ComDet, the result revealed that an average of 5.4% of telomere foci was co-localizing with Hmbox1a signal (33 Hmbox1a-telomere foci) in wild-type embryos, which was significantly higher than the 1.5% (9 Hmbox1a-telomere foci) observed in the double mutants ($p = 0.017$) (Figure 2.18c). This value is surprisingly higher than the reported 4.5 of HMBOX1-telomeres foci co-localizing in human HeLa cell⁷⁷, and is closer to the 5% of telomeres co-localizing with the shelterin protein Terfa in zebrafish¹³¹.



b

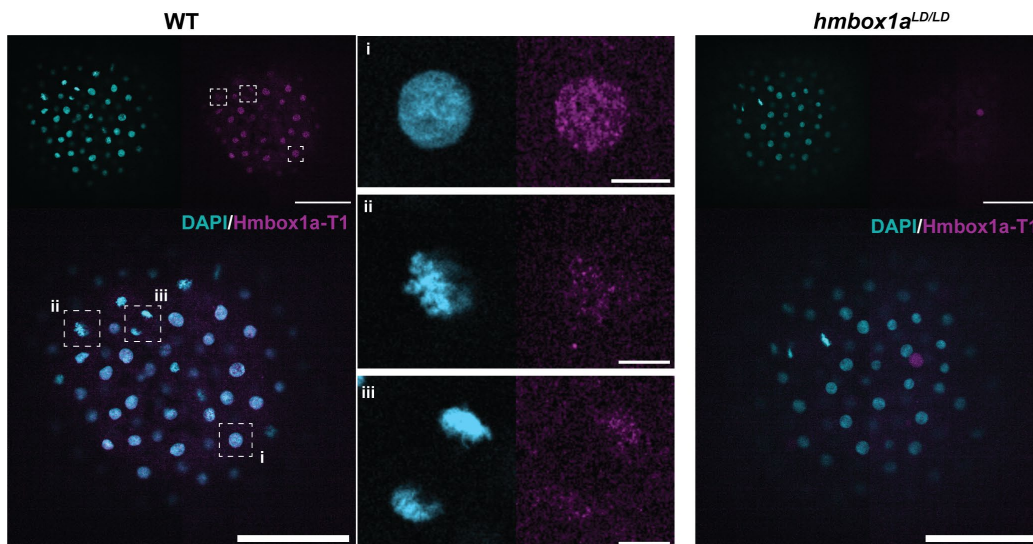


Figure 21. IF imaging of 4 hpf embryos stained with Hmbox1a-T1, Hmbox1a-T3, and Hmbox1b-T2 (M) antibodies

a) Maximum projection images of dorsal views of 4 hpf wild-type and *hmbox1a/b* mutant embryos stained with the Hmbox1a-T1 (left panel), Hmbox1a-T3 (middle panel), and Hmbox1b-T2 (M) (right panel) antibodies. For negative control (NC), the embryo was not subjected to primary antibody staining.

b) Zoomed-in images of embryos showing nuclei at different stages of mitosis: **i)** interphase, **ii)** metaphase, and **iii)** anaphase. Consistent brightness and contrast settings were applied across all images. Images in (a) were acquired using a 20X objective, except for one, and images in (b) were acquired using a 40X objective, both on a BC43 confocal microscope (Andor). All scale bar = 100 μ M, except for (bi-iii) scale bar = 10 μ m.

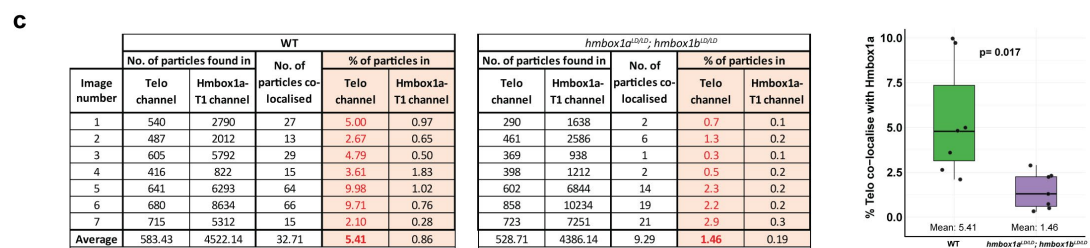
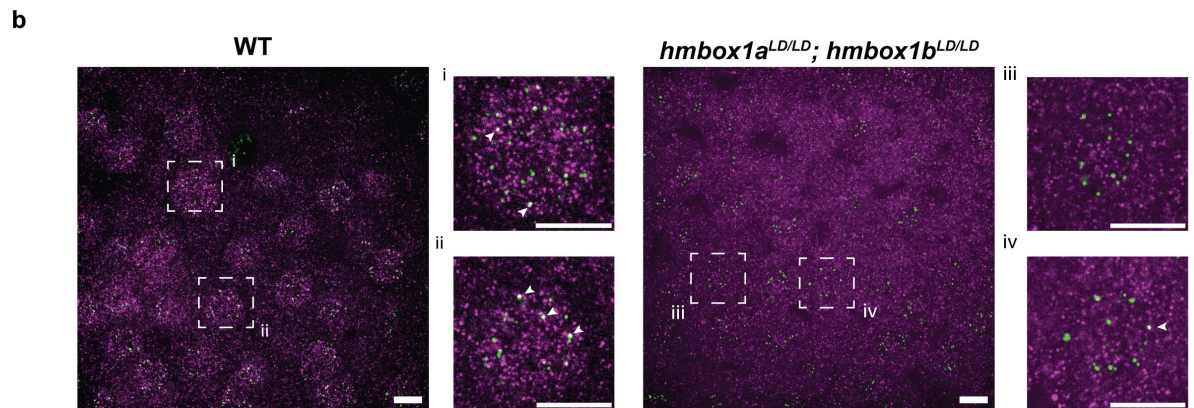
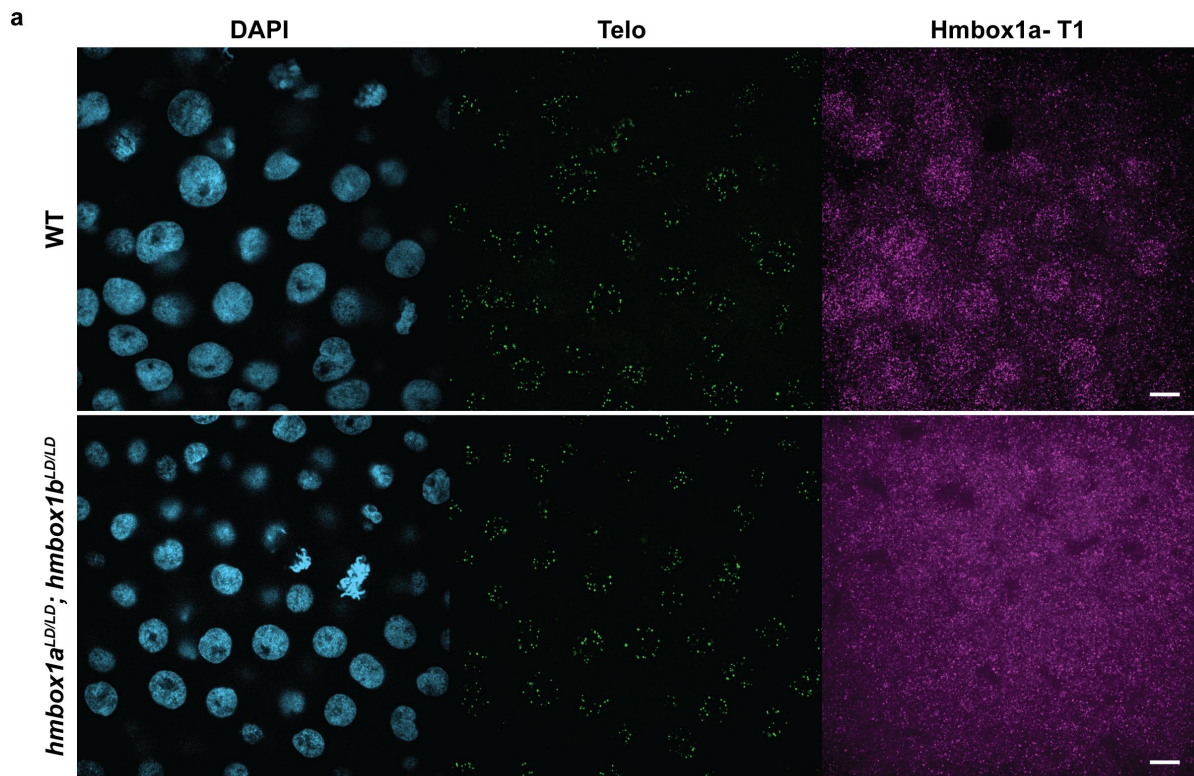


Figure 22. Telomere staining on 4 hpf embryos

a) Single-plane images of 4 hpf wild-type and double mutant embryos stained with a telomere PNA probe and the Hmbox1a-T1 antibody. Scale bar = 10 μ m.

b) Merged channel images showing the co-localization of Hmbox1a (magenta) with telomeres (green). Images were captured using a 100X objective on a BC43 confocal microscope (Andor). Scale bar = 10 μ m.

c) A table presenting the percentage of telomere and Hmbox1a signal co-localization, calculated using the ImageJ plugin ComDet (Katrukha, 2020) on 7 different images for each genotype. The values were plotted into a graph (right), and statistical analysis was performed using Welch's t-test.

Although this yielded promising results for Hmbox1a, telomere co-localization studies with Hmbox1b could not be performed on the 4 hpf embryos because Hmbox1b expression was not observed at this stage. Strikingly, the absence of the Hmbox1b signal at this stage contradicts the RNA-seq data (Figure 4), which indicates higher *hmbox1b* mRNA levels at this stage compared to 1 dpf. This discrepancy may be due to Hmbox1b expression being confined to a small population of cells at 1 dpf, resulting in a lower transcript-to-whole-embryo ratio. In contrast, Hmbox1a is broadly expressed throughout the entire embryo at 4 hpf, leading to higher overall mRNA levels compared to *hmbox1b* at the same stage. Although *hmbox1b* mRNA levels might be the highest around 2 hpf, it is also possible that the transcripts are not translated until much later during the development.

The IF results also suggest that both paralogues have an independent regulation and do not compensate for each other. The lack of Hmbox1a expression in *hmbox1a^{LD/LD}* mutants at 4 hpf did not lead to compensatory expression of Hmbox1b, and conversely, the absence of Hmbox1b expression in *hmbox1b^{LD/LD}* mutants at 1 dpf did not result in Hmbox1a expression. Furthermore, the absence of Hmbox1a expression during early development did not affect the subsequent expression of Hmbox1b in the pronephros and notochord of *hmbox1a^{LD/LD}* and *hmbox1a^{SD/SD}* mutants at 1dpf (Figure 16b).

Spatiotemporal expression of Hmbox1a during embryo development

To study the spatiotemporal expression of Hmbox1a during embryonic development, embryos were stained using the Hmbox1a-T1 antibody at various developmental stages, from the 1k-cell stage (3 hpf) to the bud stage (10 hpf). For more accurate staging in this experiment, the embryos were staged based on their developmental progression rather than solely on the hours post-fertilization, which was used as a reference point. Nuclear staining was consistently observed in wild-type embryos but was absent in the *hmbox1a^{LD/LD}* mutants. The expression of Hmbox1a began as early as the 1k-cell stage (Figure 23a). Earlier stages were not examined due to the challenges of fixing cells at interphase during the rapid cell divisions of embryos, which occur every 15 minutes. The first ten divisions are synchronized, but multiple attempts to fix them at interphase were unsuccessful. Timing the interphase stage and rapidly fixing the embryos was difficult, and most were fixed during the mitotic phase, where the Hmbox1a signal is barely detectable.

However, staining Hmbox1a at later stages was more successful. The signal was observable at the shield stage (6 hpf), 75% epiboly (8 hpf), and continued at least up to the bud stage. Hmbox1a was consistently detected in the majority of cells throughout development. As development progressed, the Hmbox1a signal seems to decline with the size of the nuclei. Eventually, such signal was no longer detectable at 1 dpf (Figure 18).

Along with embryonic cells, the nuclei of the yolk syncytial layer (YSL) were also stained with Hmbox1a. The YSL is an extra-embryonic structure over the yolk. It forms at around 1k-cell stage and plays an important role in embryo development as it is involved in cell signaling, as well as the development and differentiation of embryonic cells¹⁴³. The Hmbox1a signal was observed in the YSL nuclei of embryos at 75% epiboly and bud stages but not at the sphere

stage (4 hpf) (Figure 23b, white arrowhead). This signal was also not observed in the yolk of 1 dpf embryos (Figure 18).

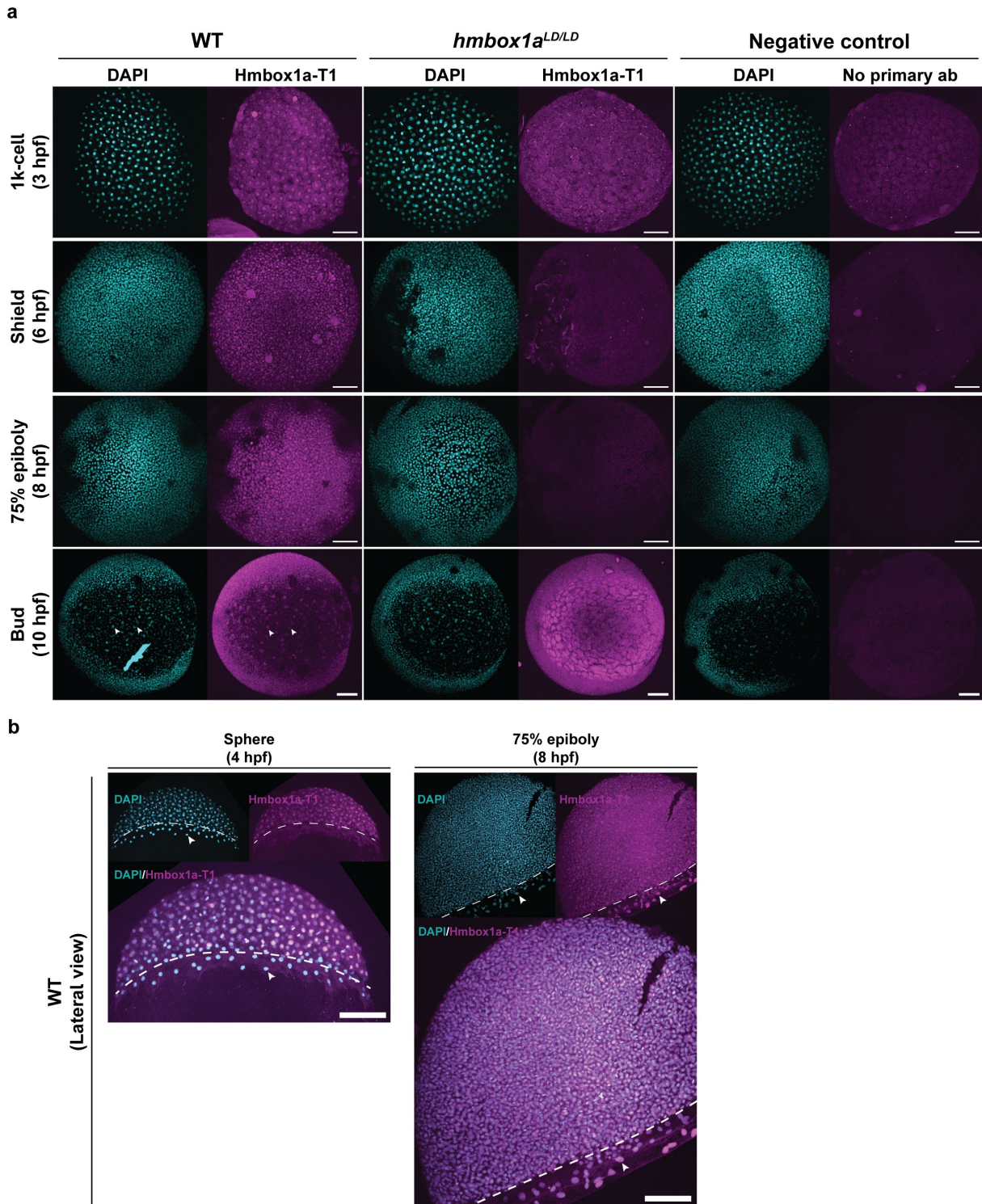


Figure 23. Hmbox1a expression in early embryos

a) Maximum projection images of wild-type and *hmbox1a^{LD/LD}* mutant embryos at various developmental stages, stained with the Hmbox1a-T1 antibody. For negative control, the embryo was not subjected to primary antibody staining. Embryos from the 1k-cell to 75% epiboly stages are shown in dorsal view, while the bud stage is presented in lateral view.

b) Maximum projection images of lateral views of wild-type embryos showing yolk syncytial nuclei staining. The dotted line draws the separation between the embryonic cells and yolk syncytial layer.

Consistent brightness and contrast settings were applied across all images. Images were acquired using a 20X objective, except for the bud stage, which used a 10X objective, on a BC43 confocal microscope (Andor). The white arrowhead indicates the yolk syncytial layer nucleus. Scale bar = 100 μ m.

Noticeably, the expression of Hmbox1a begins at a very early stage in embryonic development, suggesting that it might be maternally deposited. In zebrafish, the maternal-to-zygotic transition (MZT) occurs around 1k-cell stage. During this transition, maternally loaded materials are gradually degraded, while the zygotic genome activation begins and the zygote starts transcribing from its own genome¹⁴⁴. To determine whether Hmbox1a is maternally deposited or expressed by the zygote during early development, crosses between parents with different genotypes (wild-type and *hmbox1a*^{LD/LD}) were conducted.

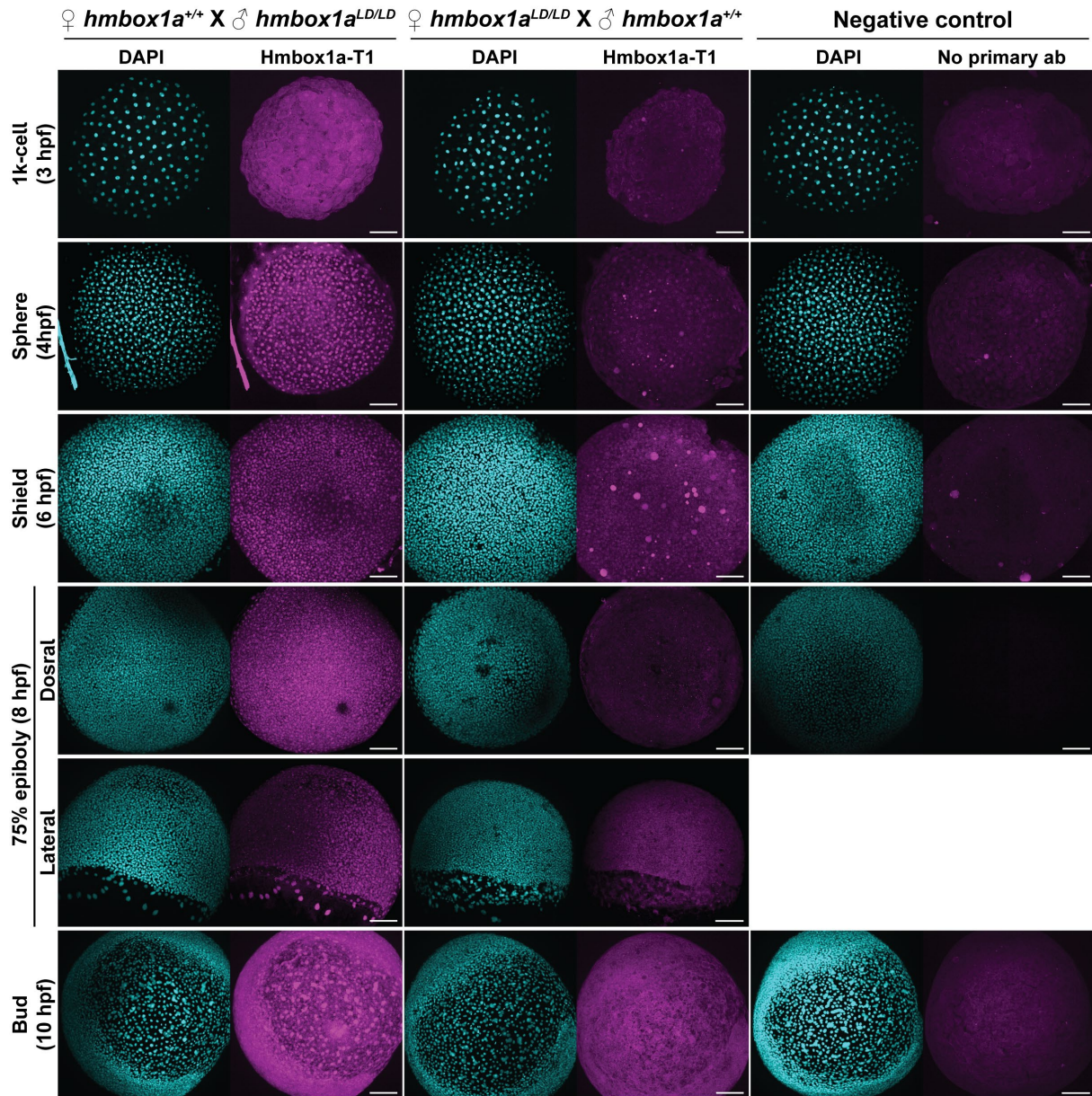


Figure 24. Maternal inheritance of Hmbox1a

Maximum projection images of *hmbox1a*^{LD/LD} heterozygous embryos at various developmental stages, stained with the Hmbox1a-T1 antibody. The embryos were bred from either a wild-type mother and *hmbox1a*^{LD/LD} mutant father (left panel) or an *hmbox1a*^{LD/LD} mutant mother and wild-type father (middle panel). For the negative control, embryos were not subjected to primary antibody staining. The images depict dorsal views of embryos from the 1k-cell to 75% epiboly stages, and lateral views of embryos from the 75% epiboly stage to the bud stage.

Consistent brightness and contrast settings were applied across all images. Images were acquired using a 20X objective on a BC43 confocal microscope (Andor). Scale bar = 100 μ m.

In a cross between a wild-type mother and an *hmbox1a*^{LD/LD} mutant father, a weak nuclear signal was detected at 1k-cell (Figure 24). The signal intensified by sphere stage (4 hpf) and was sustained at least until the bud stage, consistent with observations from in-crosses of wild-type parents. Additionally, staining was observed in the YSL nuclei at 75% epiboly and bud stage. In contrast, the reverse cross (*hmbox1a*^{LD/LD} mutant mother and wild-type father) showed no detectable signal at any stage, including in the YSL nuclei. These results confirm that Hmbox1a is maternally loaded and highlight the importance of considering parental genotypes in studies involving embryos during early development.

It appears that Hmbox1a is solely maternally provided and not zygotically expressed. This is supported by the lack of Hmbox1a signal in the heterozygotes bred from crossing an *hmbox1a*^{LD/LD} mutant mother with a wild-type father, even after the zygotic genome activation event. As with many maternally provided proteins, Hmbox1a is gradually degraded as zygotic development progresses, which aligns with the observed decrease in Hmbox1a expression (Figure 23).

To summarize the observed Hmbox1a expression: Hmbox1a protein is expressed early in development and has been confirmed as a maternally provided protein. No evidence of zygotic expression was found in those stages investigated. Hmbox1a was detected in the nuclei of nearly all embryonic cells, as well as in the YSL nuclei. Its expression is detectable as early as 3 hpf and persists through the bud stage, although it decreases as development progresses, becoming negligible by 1 dpf. Notably, YSL nuclei expression appears to be restricted to the period between the 75% epiboly and bud stages. Despite its early expression, the loss of Hmbox1a expression does not seem to impact embryo development.

Spatiotemporal expression of Hmbox1b during embryo development

Unlike Hmbox1a, which is expressed throughout the embryo, Hmbox1b is highly expressed in a small group of cells, such as the notochord and pronephros, as observed at 1 dpf (Figure 16). But its expression at 4 hpf was not observed (Figure 21). To study the spatiotemporal expression of Hmbox1b, embryos of later stages were probed with the Hmbox1b-T2 (M) antibody.

The first appearance of Hmbox1b was detected in the nucleus of the notochord at the 90% epiboly stage (Figure 25a), during the onset of notochord formation¹⁴⁵. The signal in notochord became more pronounced by the 10-somite stage (Figure 25b), with a higher signal to noise ratio than the 90% epiboly stage in the max projection images. The pronephros also begins to form at this stage, with *Wt1b* and *cdh17* signals present. However, the Hmbox1b signal in the pronephros was minimal (Figure 25b). By the 14-somite stage, Hmbox1b had a high signal in both the notochord and pronephros (Figure 26a). Strong Hmbox1b signal was observed in the densely packed, rod-like nuclei of the notochord (Figure 26bi, top) and the pronephros, co-localizing with *cdh17* and DAPI (Figure 26bi, bottom). These findings suggest that Hmbox1b expression begins in the notochord upon its formation and in the pronephros shortly after its development begins.

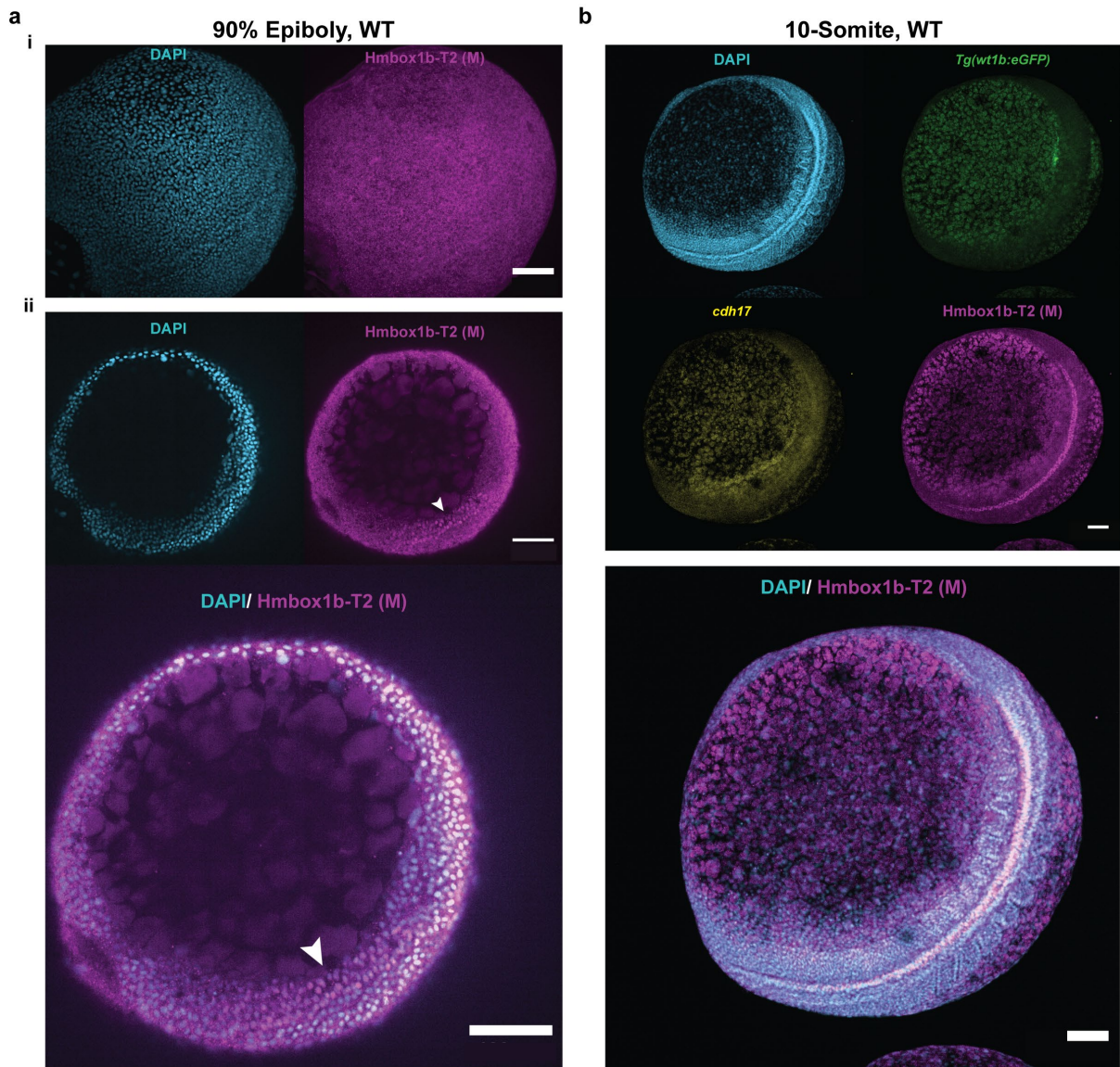


Figure 25. Hmbox1b expression in 90% epiboly and 10-somite stage

a) Lateral view of wild-type embryo at the 90% epiboly stage, stained with the Hmbox1b-T2 (M) antibody (magenta) and DAPI (cyan). **i)** Shows the maximum projection images, while **ii)** displays a single plane of the corresponding image with individual and merged channels. The white arrowhead indicates the presumptive notochord.

b) Maximum projection images of lateral views of wild-type embryo at the 10-somite stage. The embryo, in a *Tg(wt1b:eGFP)*/*li1* background (green), was probed with an RNA probe for *cdh17* (yellow) to label the pronephros, and counterstained with Hmbox1b-T2 (M) antibody (magenta) and DAPI (cyan).

Images were acquired using a 10X objective for (a) and a 20X objective for (b) on a BC43 confocal microscope (Andor). Scale bar = 100 μ m.

The Hmbox1b signal in the pronephros was still detectable at 2 dpf (Figure 27a). However, by 3 dpf, the signal began to fade at both the proximal and distal end of the pronephros (Figure 27bi and ii, respectively), but it has yet to reach the level of absence observed in the *hmbox1b*^{LD/LD} mutants (Figure 27biii and iv, respectively). By 5 dpf, the signal was undetectable (Figure 27bv and vi). This lack of signal in the pronephros was unlikely to be due to antibody penetration issues. This is because measures such as acetone permeabilization and the removal of larvae heads and tails were deployed to enhance permeabilization in the staining of 3 dpf and 5 dpf larvae, but the signal remained low. Therefore, it is likely that Hmbox1b expression in the pronephros progressively declines with development.

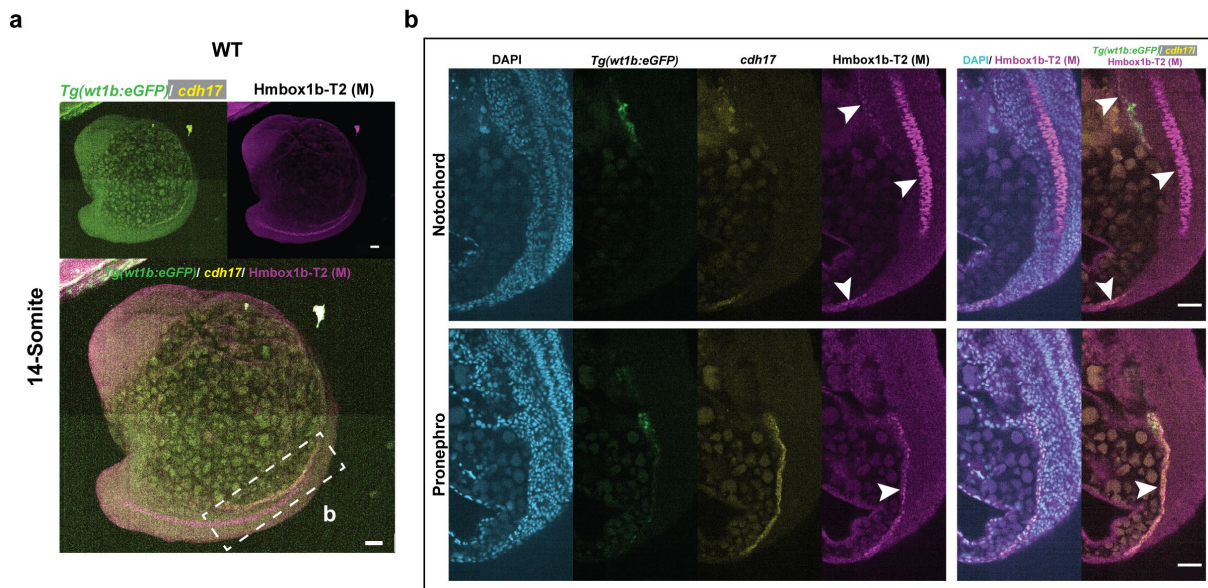


Figure 26. Hmbox1b expression at 14-somite stage

a) Maximum projection image of the lateral view of a wild-type embryo at the 14-somite stage. The embryo, with a *Tg(wt1b:eGFP);ii1* background (green), was probed with an RNA probe for *cdh17* (yellow) to label the pronephros and counterstained with Hmbox1b-T2 (M) antibody (magenta) and DAPI (cyan).

b) Zoomed-in view of the notochord and pronephros, showing individual and merged channels as indicated. The white arrowhead marks the notochord and pronephros.

Images were acquired using a 10X objective for (a) and a 20X objective for (b) on a BC43 confocal microscope (Andor). Scale bar = 50 μ m.

Other than the pronephros, the notochord nuclei also did not appear stained by 3 dpf (Figure 27bvii). Instead, non-specific signals were observed at the tail region, where the notochord is supposedly located, in both wild-type and *hmbox1b^{LD/LD}* mutants at 3 dpf and 5 dpf. This loss of signal coincides with the transformation of the notochord into the vertebral column, which starts after 2 dpf¹⁴⁵.

Although Hmbox1b expression was not detected in the podocytes, a glomerulus filtration test was performed on 4 dpf larvae to confirm that the loss of Hmbox1b does not impact the filtration function of the podocytes. In this study, two different molecular weight dextrans were co-injected into the blood stream of wild-type and (*hmbox1a^{LD/LD}; hmbox1b^{LD/LD}*) double mutants. The lower molecular weight dextran was an Alexa Fluor 647-conjugated 10-kDa dextran, while the higher molecular weight dextran was a FITC-conjugated 500-kDa dextran. The dyes were injected into the heart instead of the caudal vein or the pericardial sac as described by the authors^{146,147}. In a normal functioning glomerulus, the lower molecular weight dextran will be filtered out by 24 hours post-injection (hpi). If this dextran is not filtered out, it would indicate a defect in the filtration system of the pronephros. Conversely, the higher molecular weight dextran should remain in the system because it is too large to be filtered out unless the glomerulus is 'leaky', or too porous.

The injected larvae were imaged at 3 hpi to confirm successful injection and at 24 hpi to assess glomerulus filtration (Figure 28). The images of the different molecular weight dextrans were overlaid and compared between the two time points to evaluate filtration efficiency. As expected, there appeared to be no difference in the results between the wild-type and double mutant larvae at 24 hpi. Most of the lower molecular weight dextran was filtered out of the circulatory system by 24 hpi, while the higher molecular weight dextran was retained in both

wild-type and double mutants. Thus, these results suggest that the glomerulus is functioning normally.

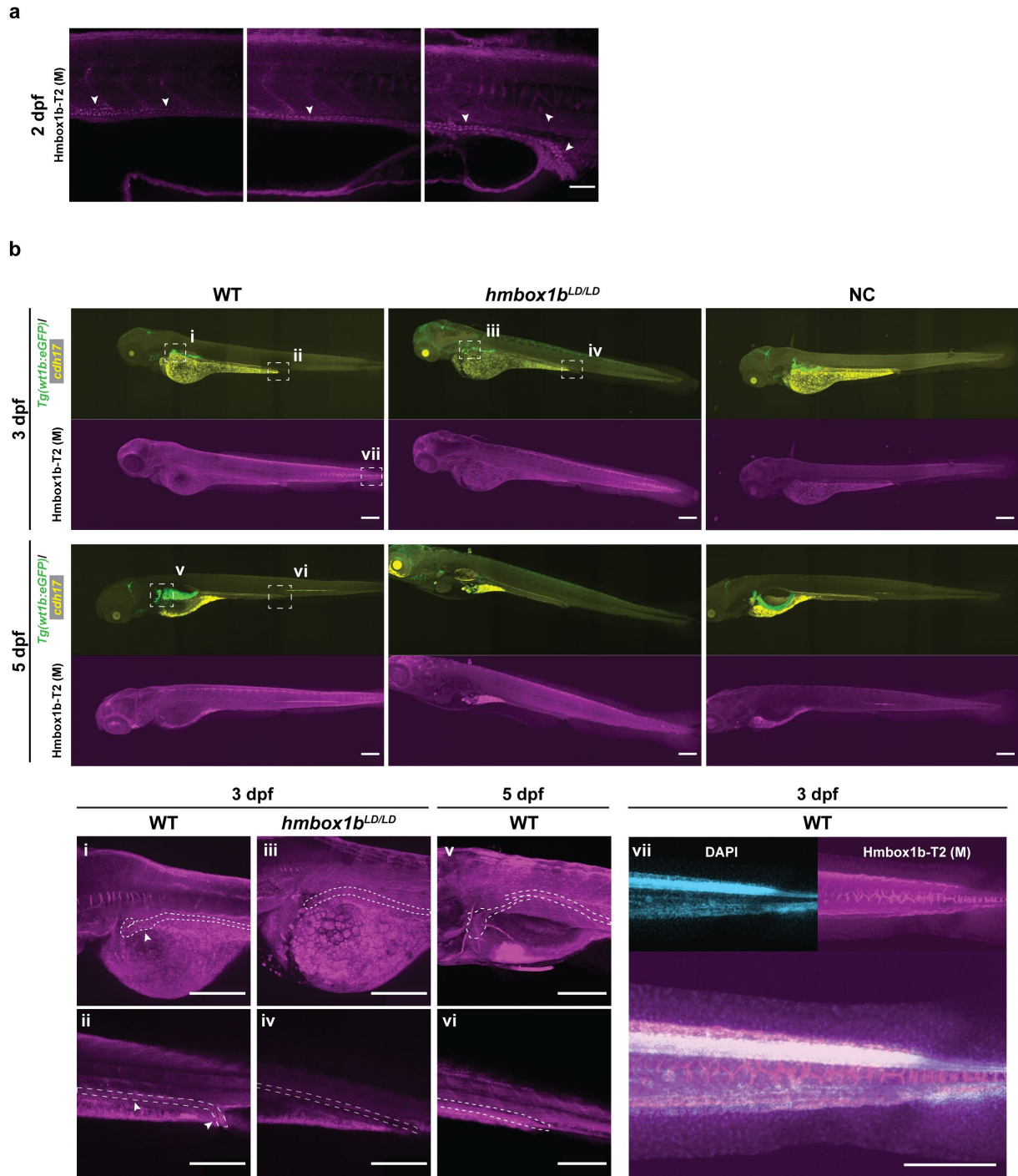


Figure 27. Hmbox1b expression at larval stages

a) Single-plane image of 2 dpf wild-type larvae stained with the Hmbox1b-T2 (M) antibody. The white arrowhead marks the pronephros. Images were captured using a 40X objective on a VisiScope confocal microscope. Scale bar = 50 μ m.

b) Maximum projection images of 3 and 5 dpf wild-type and *hmbox1b^{LD/LD}* mutant larvae, with a *Tg(wt1b:eGFP)/li1* background (green), probed with an RNA probe for *cdh17* (yellow) to label the pronephros and counterstained with Hmbox1b-T2 (M) antibody (magenta). Zoomed-in views of the respective areas are shown in (bi-vii). White dotted lines outline the pronephros based on the Wt1b and *cdh17* signals. Images were acquired using a 10X objective for (b) and a 20X objective for panels (bi-vii) on a BC43 confocal microscope (Andor). Scale bar = 200 μ m.

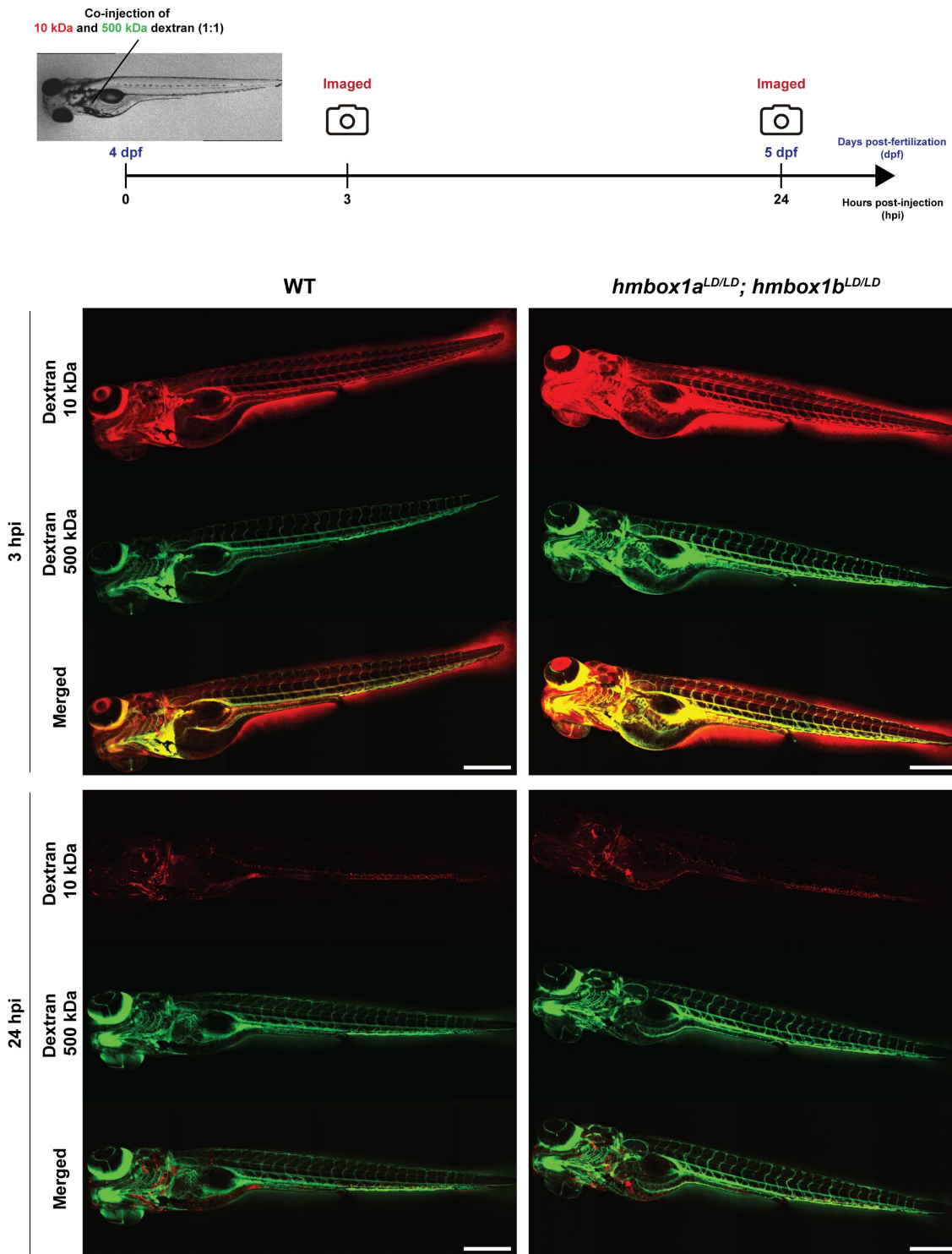


Figure 28. Glomerulus filtration test

The schematic diagram at the top illustrates the timeline and procedure of the test. The images below present maximum projections of wild-type and *hmbox1a^{LD/LD}; hmbox1b^{LD/LD}* double mutants, captured at 3 hours and 24 hours post-injection (hpi). The larvae were injected with Alexa Fluor 647-conjugated 10-kDa dextran (red) and FITC-conjugated 500-kDa dextran (green). All images were acquired using a 10X objective on a VisiScope confocal microscope and were stitched together using plug-in on ImageJ (Preibisch et al., 2009). Consistent brightness and contrast settings were applied across all images. Scale bar = 2 cm.

Distribution of *hmbox1a/b* genotype and survival

Based on the results obtained so far, knock out of either *Hmbox1a*, *Hmbox1b*, or both did not lead to embryo lethality or developmental defects, despite their early expression during embryogenesis. These homozygous mutants not only survive to adulthood, but have also reproduce, which provided some samples for IP-MS and IF analyses.

To evaluate whether the different genotypes have same survival fitness during early adulthood, the genotype distribution of offspring bred from double heterozygous (*hmbox1a*^{+LD}; *hmbox1b*^{+LD}) parents was analyzed through fin clipping between 1.5 and 3 months of age. In a normal distribution, the probabilities of having wild-type, double mutants (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}), and either of the single mutants (*hmbox1a*^{LD/LD} or *hmbox1b*^{LD/LD}) are each 6.25% (1 in 16), the heterozygotes (*hmbox1a*^{+LD}; *hmbox1b*^{+LD}) are 25%, and the remaining genotypes are 12.5% each. In the three crosses evaluated, they all followed Mendelian inheritance patterns. The chi-square tests revealed no significant differences in the genotype distribution for individual crosses or the combined analysis of all three (Figure 29a, $\chi^2 < 15.51$).

In the first cross (Fish ID: 9907), only one double mutant was detected instead of the expected four, which could raise speculation about a potential reduction in fitness for the double mutants. However, this discrepancy is likely due to the small sample size (n=79). In subsequent crosses with larger sample sizes, such as cross 3 (Fish ID: 736), the genotype follows normal distribution, with the observed count of double mutants matching the expected count of 28. This larger sample refutes the speculation that double mutants have reduced fitness.

a

Cross	Age genotyped	Fish ID	GT	Observed										Total	Individual cross χ^2
				<i>hmbox1a</i> <i>hmbox1b</i>	+/+	-/-	-/-	+/+	-/-	+/-	+/-	+/-	+/-		
1	6 wpf	9907		10	1	4	2	14	8	7	10	23	79	13.71	
2	7 wpf	ded1		17	12	6	11	20	14	18	22	38	158	9.39	
3	12 wpf	736		26	28	33	24	52	58	52	56	128	457	4.27	
Total Observed (O)				53	41	43	37	86	80	77	88	189	694		
Expected (E)				43.34 (6.25%)				86.75 (12.50%)				173.5 (25%)			
Overall $\chi^2 = \sum (O-E)^2/E$				6.24											

For a test of significance at $\alpha = 0.05$ and $df = 8$, the χ^2 critical value is 15.51

b

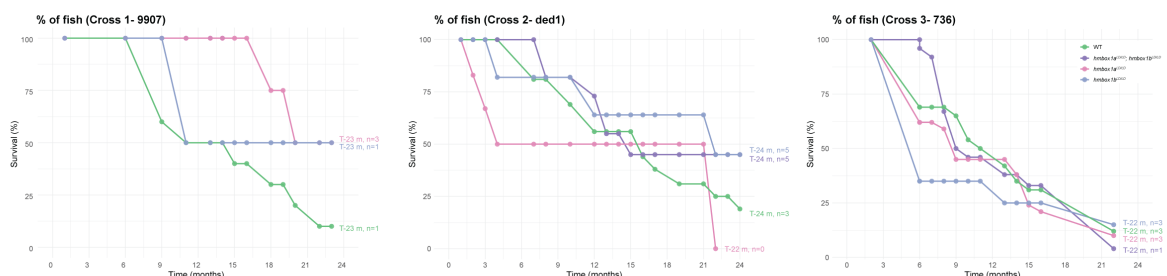


Figure 29. First generation of CRISPR-Cas9 *hmbox1a/b* mutant lines

a) A table displaying the number of fish of each genotype from three different crosses between *hmbox1a*^{+LD}; *hmbox1b*^{+LD} heterozygous parents. A Chi-square (χ^2) test was conducted for each individual cross as well as an overall test for all three crosses combined. If the observed values significantly differ from the expected values, the χ^2 statistic must exceed the critical value of 15.51 for an alpha level of 0.05 and 8 degrees of freedom (df). Age of the fish was represented in week-post fertilization (wpf).

b) The graph plots the percentage of fish remaining in the tank over time for the first mutant generations of different crosses. Fish were periodically counted, with the age (T) and number (n) of the fish at the end of each count noted on the side of the graph. For the starting age and number of fish, refer to the data in (a).

Since the genotype distribution follows Mendelian patterns at 1.5 to 3 months old, it is unlikely that there are survival disadvantages during embryonic and juvenile development. This normal distribution also indicates that double mutants do not experience a survival disadvantage compared to single homozygous mutants ($hmbox1a^{LD/LD}$ and $hmbox1b^{LD/LD}$) or three-allele mutants ($hmbox1a^{LD/LD}; hmbox1b^{+/LD}$ and $hmbox1a^{+/LD}; hmbox1b^{LD/LD}$).

Next, to assess whether there was a reduction in survival fitness during later life stages, the number of fish per genotype in each tank was regularly monitored. Graphs were created to display the percentage of fish remaining in each genotype over time. The decrease in the number of fish was not discriminated between deaths and illness, or escapes. As such, Kaplan-Meier survival curves were not used due to this missing information. There appears to be minimal difference in survival rates among the different genotypes from the three first-generation mutant crosses (Figure 29b). All homozygous mutants survived to at least two years of age in at least one of the crosses. Interestingly, although wild-type fish were presumed to be the fittest, they did not consistently exhibit the highest survival rates at all time points compared to the homozygous mutants across all three crosses.

Aside from having only minor impacts on lifespan, these homozygous mutants were both fertile and capable of reproducing for at least three additional inbred generations. Figure 30a shows the family tree of the CRISPR-Cas9 *hmbox1a/b* mutants, with survival curves for each successive generation plotted in Figure 30b. The mutants were bred up to the fourth generation (G4), with their offspring remaining viable and appearing healthy. However, inbreeding depression¹⁴⁸ became evident by the second generation (G2) of mutants. Of the 50 embryos initially placed in the aquarium at 5 dpf, fewer than 40% survived to 3-month-old. This contrasts with the first generations, where at least 50% survival at 3-month-old in most genotypes. Although the wild-type fish had the highest survival rate at 3 months compared to the other G2 mutants, its population slowly declines after 1 year old to eventually dip below all the mutants at 16-month-old.

Similarly, in the third generation (G3), a significant decline in all populations was observed within the first 6 months. Both $hmbox1a^{LD/LD}$ and $hmbox1b^{LD/LD}$ mutants, as well as the wild-type, survived to at least 17 months. In contrast, a double mutant clutch (Fish ID: 879) exhibited abnormal swimming behavior around four months of age. Some were struggling to swim and maintain balance, displaying signs of swim bladder issues, and they often gathered at the top of the tank (Figure 31). Following animal welfare guidelines, this clutch was euthanized before reaching five months. When a new cross (Fish ID: 959) was made from the same parental fish four months later, this phenotype was not observed, and the fish survived for at least a year before the study concluded. Thus, the observed phenotype might not be due to the double mutant effect.

In summary, the study found no strong evidence that knocking out of *Hmbox1a* and *Hmbox1b* impacts fitness or lifespan in zebrafish. The mutants showed no lethality during embryonic or juvenile stages and were able to reproduce successfully across multiple generations, with no lifespan differences compared to wild-type fish.

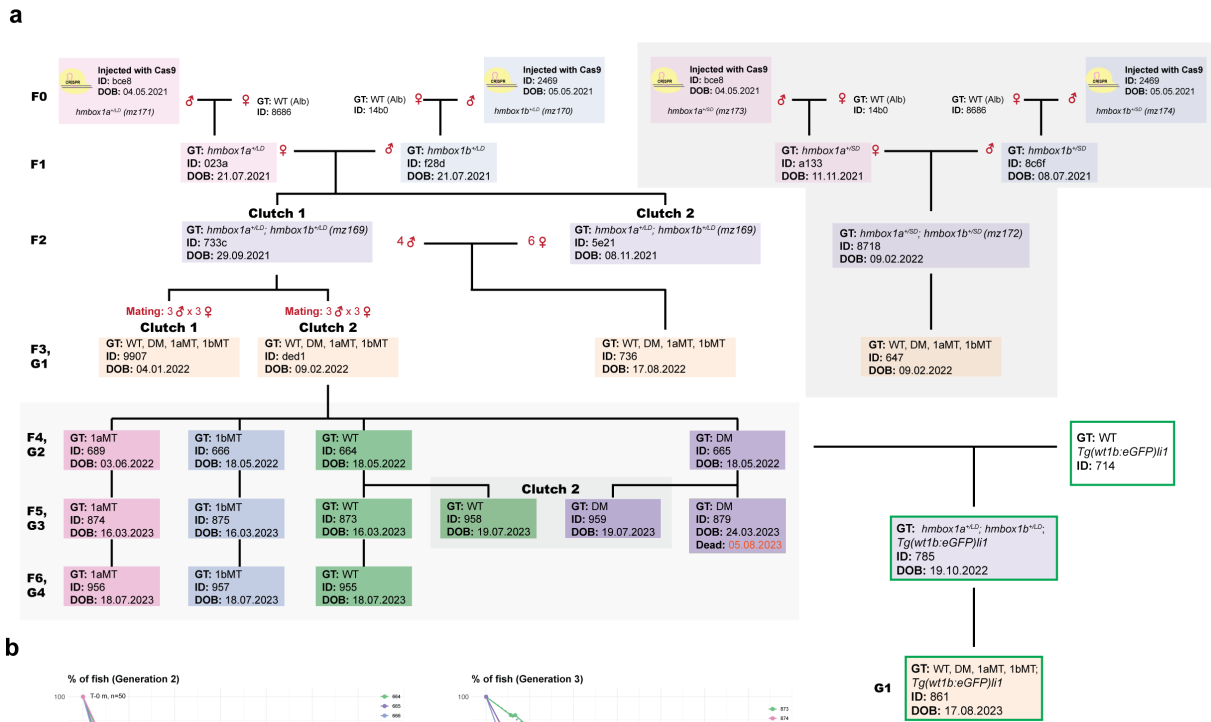


Figure 30. Establishing CRISPR-Cas9 *hmbox1a/b* mutant lines

a) The family tree of CRISPR-Cas9 *hmbox1a/b* mutant lines generated during this study. Vertical lines represent a new generation, while horizontal lines indicate matings between parents with different fish IDs (noted with the number of females and males beside them). If only a vertical line is shown without a horizontal line, it represents incrosses. When the number of males and females is not specified, it refers to a single mating. Each box represents a clutch and is color-coded based on genotype. It details the fish genotype, ID, and the date of birth (DOB). Only the tank that dies prematurely was noted with the date of death. Orange-filled boxes represent offspring of multiple genotypes due to *hmbox1a^{h/LD}; hmbox1b^{L/D}* heterozygous parent crosses. A green border indicates a transgenic line, *Tg(wt1b:eGFP)l1*. Generations are labeled as "F" (filia), starting from the CRISPR-injected embryos, while "G" labels denote homozygous mutant generations resulting from incrossing. The genotype was coded as follows: WT: wild-type, DM: *hmbox1a^{L/D/LD}; hmbox1b^{L/D/LD}*, 1aMT: *hmbox1a^{L/D/LD}*, 1bMT: *hmbox1b^{L/D/LD}*.

b) The graph plots the percentage of fish remaining in the tank over time for the second and third mutant generations that arose from incrossing. Fish were periodically counted, with the age (T) and number (n) of the fish at the start and end of each count noted on the side of the graph.

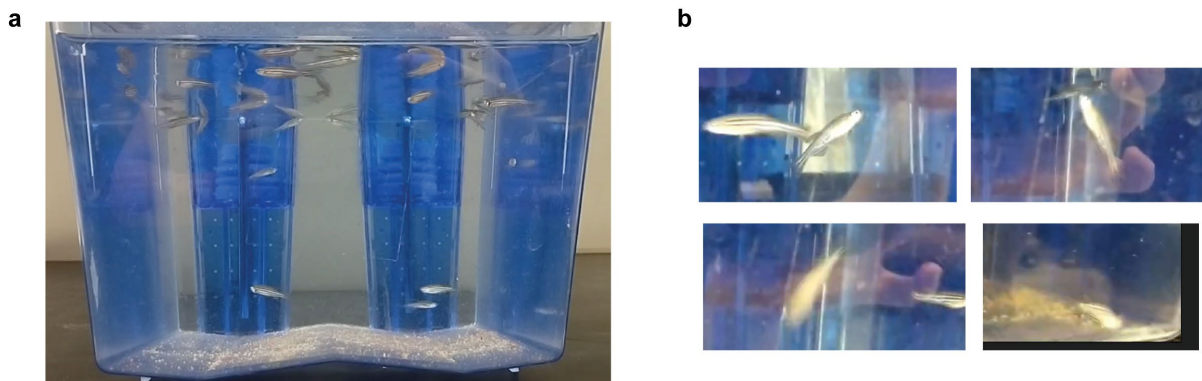


Figure 31. Abnormal swimming behavior of third generation double mutant (Fish ID: 879)

a) The tank containing the third-generation (*hmbox1a^{L/D/LD}; hmbox1b^{L/D/LD}*) double mutant (Fish ID: 879) at approximately 4 months of age, gathering near the top.

b) Snap shot of an individual fish exhibiting abnormal swimming behavior.

Telomere length of (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutants

To investigate the influence of *hmbox1a/b* knockout on telomere length, tail (caudal) fins were clipped from three wild-type and (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutant males at 8 months of age. These mutants were from the first generation, specifically from cross 2, and were bred through group mating. Telomere restriction fragment (TRF) analysis using Southern blotting was performed on the clipped samples by our collaborator, Naz Şerifoğlu from Miguel Godinho Ferreira's lab. However, the genomic material from one of the double mutants was not successfully detected and was excluded from the analysis. Based on the remaining samples, the results showed no significant change in telomere length ($p = 0.49$) (Figure 32).

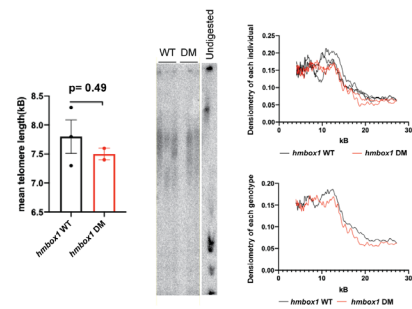


Figure 32. Telomere length of double mutants at 8 month old

Telomere restriction fragment (TRF) analysis was conducted on wild-type males ($n=3$) and (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutant males ($n=2$), from Cross 2, at 8 months of age. The graph on the left illustrates the average telomere length for each genotype. Error bars represent the standard error of the mean (SEM), and p -value was determined by an unpaired two-tailed t -test (Mann-Whitney). The central panel displays the Southern blot results of the TRF analysis. The two graphs on the right show the densitometry analysis: the top graph represents the band intensity for individual samples, while the bottom graph groups the data by genotype.

Proteome analysis at 4 hpf

The knockout of *hmbox1a/b* does not appear to impact embryonic development, lifespan, or telomere length in those mutants examined. To further explore the role of Hmbox1a/b, proteome analysis was conducted at 4 hpf, when Hmbox1a expression was observed. The samples included wild-type, (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutants, *hmbox1a*^{LD/LD}, and *hmbox1b*^{LD/LD} embryos, with four technical replicates for each group. Embryos were collected from incrosses of parents with the corresponding genotypes. The embryos were pooled and deysolked.

Despite immunofluorescence confirming Hmbox1a expression at 4 hpf, it was not detected in the proteome and neither was Hmbox1b. Pairwise comparisons were made between embryos of *hmbox1a*^{+/+} (wild-type and *hmbox1b*^{LD/LD}) and *hmbox1a*^{LD/LD} genotypes (*hmbox1a*^{LD/LD} and (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD})), based on the assumption that Hmbox1b is not expressed at this developmental stage and the *hmbox1b* transcript does not influence the regulation of Hmbox1a or other genes. The goal was to identify proteins commonly dysregulated between *hmbox1a*^{+/+} and *hmbox1a*^{LD/LD} genotypes. This analysis revealed a consistent upregulation of 10 proteins and downregulation of 14 across all comparisons (Figure 33a).

To better visualize the gene expression of proteins dysregulated by the *hmbox1a*^{LD/LD} mutation across the different genotypes, a heatmap was generated (Figure 33b). Proteins like Cyp2x7, Fetub, and the uncharacterized protein Si:dkey-19b23.11 were exclusively detected in embryos carrying the *hmbox1a*^{LD/LD} mutation. Conversely, proteins such as Glrx, Gpn3, Si:dkey-111e8.1, and Strada, were absent in both *hmbox1a*^{LD/LD} and (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutants. Interestingly, some proteins undetected in the *hmbox1a*^{LD/LD} mutant were found in some of the technical replicate(s) of the double mutants, including Cnot4b, Ddx10, Si:ch1073-412h12.3, Si:ch211-91p5.3, Vcpkmt, and Zgc:65997. The functions of these proteins are listed in Table 1, most of them are uncharacterized.

Two of these proteins are particularly noteworthy as they may be connected to pronephros development and the paralogue *Hmbox1b* expression. One is the upregulation of retinol-binding protein 4 (*Rbp4*) in *hmbox1a^{LD/LD}* mutants, which is responsive to retinoic acid. The other is the downregulation of beta-ureidopropionase (*Upb1*), known to be expressed in the pronephros at 1 dpf ¹⁴⁹.

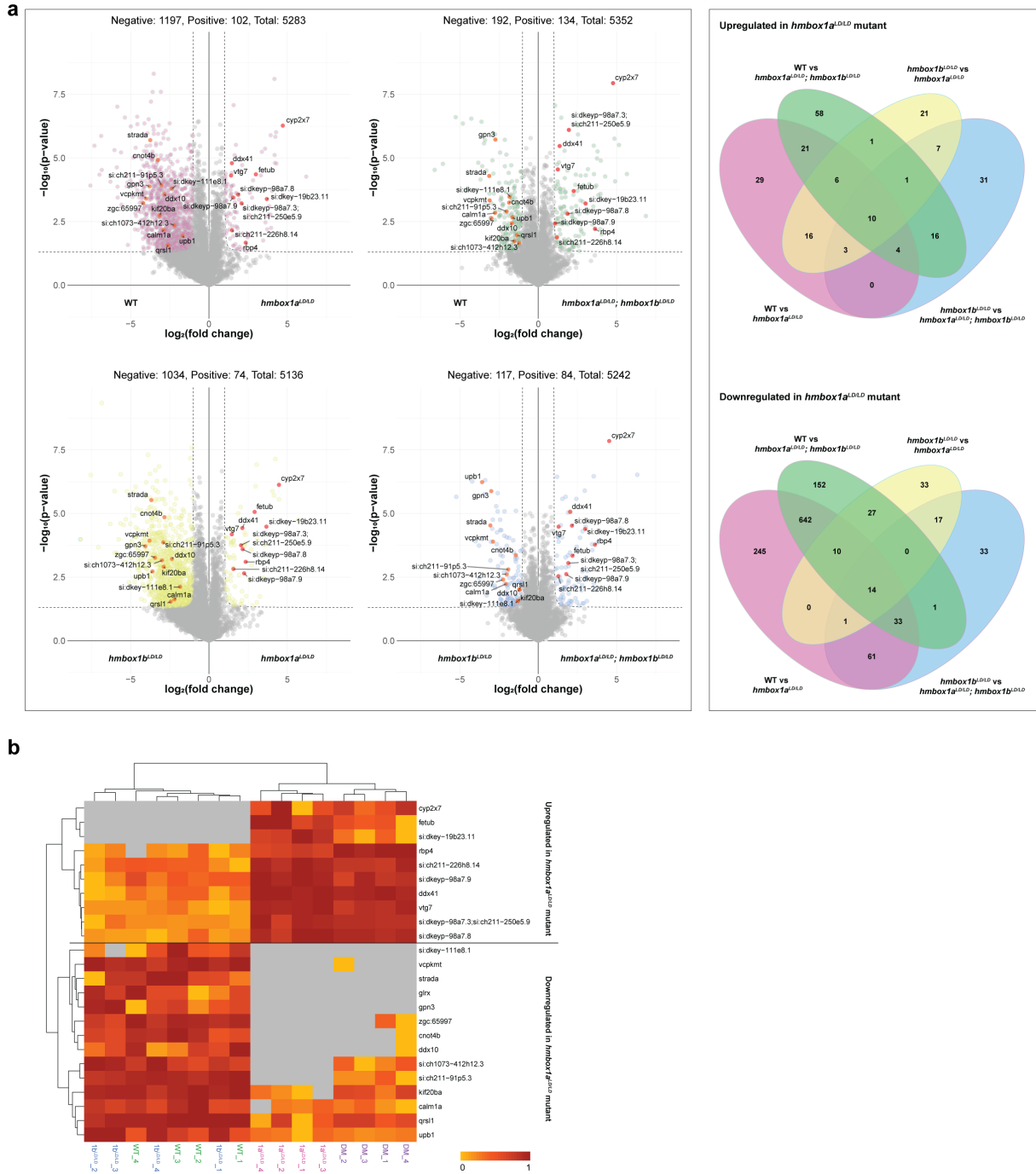


Figure 33. Proteome analysis of wild-type and *hmbox1a/b* mutants at 4 hpf

Proteome analysis was conducted on whole-cell lysates from 4 hpf wild-type, *hmbox1a^{LD/LD}*, *hmbox1b^{LD/LD}* double mutants, *hmbox1a^{LD/LD}* single mutants, and *hmbox1b^{LD/LD}* single mutants.

a) Volcano plots show the proteome analysis results comparing embryos possessing the *hmbox1a^{+/+}* genotype (wild-type and *hmbox1b^{LD/LD}*) to those with the *hmbox1a^{LD/LD}* genotype (*hmbox1a^{LD/LD}*, *hmbox1a^{LD/LD}; hmbox1b^{LD/LD}*). A Venn diagram to the right presents the number of commonly dysregulated proteins across all pairwise comparisons. Proteins consistently upregulated or downregulated in all comparisons are marked in red and orange, respectively, in the volcano plots. The enrichment threshold was set at a fold change > 2 and a p-value < 0.05 (Welch's t-test) with $c = 0.05$. Proteins are annotated with their gene names.

b) A heatmap displays the expression levels of proteins that are commonly dysregulated in all comparisons in (a), across different technical replicates of the various genotypes.

Table 1. Summary of dysregulated proteins identified in the proteomic comparison between *hmbox1a^{+/+}* and *hmbox1a^{LDLD}* mutants at 4 hpf

	Gene name	Full gene name	Zfin description
Proteins dysregulated in embryos possessing <i>hmbox1a^{LDLD}</i>	<i>si:ch211-226h8.14</i> <i>si:dkeyp-98a7.9</i>		
	<i>rbp4</i>	retinol binding protein 4, plasma	Predicted to have retinol binding activity. Involved in liver development and response to retinoic acid. Predicted to localize to extracellular space. Human ortholog(s) of this gene implicated in type 2 diabetes mellitus. Is expressed in several structures, including apical ectodermal ridge; digestive system; integument; reproductive system; and yolk. Orthologous to human RBP4 (retinol binding protein 4).
	<i>ddx41</i>	DEAD (Asp-Glu-Ala-Asp) box polypeptide 41	Predicted to have ATP binding activity; carbohydrate binding activity; and nucleic acid binding activity. Involved in NIK/NF-kappaB signaling; innate immune response; and response to virus. Localizes to cytoplasm and nucleus. Is expressed in brain; digestive system; gill; heart; and immune system. Orthologous to human DDX41 (DEAD-box helicase 41).
	<i>vtg7</i>	vitellogenin 7	Predicted to have lipid transporter activity. Predicted to be involved in cellular response to estrogen stimulus and response to estradiol. Is expressed in liver; pleuroperitoneal region; skeletal muscle cell; and unfertilized egg.
	<i>si:dkeyp-98a7.3</i> <i>si:ch211-250e5.9</i> <i>si:dkeyp-98a7.8</i>		
	<i>cyp2x7</i>	cytochrome P450, family 2, subfamily X, polypeptide 7	Predicted to have heme binding activity; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen; and steroid hydroxylase activity. Predicted to be involved in several processes, including exogenous drug catabolic process; organic acid metabolic process; and xenobiotic metabolic process. Predicted to localize to cytoplasm and intracellular membrane-bounded organelle.
	<i>fetub</i> <i>si:dkey-19b23.11</i>	fetuin B	Predicted to have endopeptidase inhibitor activity. Predicted to be involved in negative regulation of endopeptidase activity. Predicted to localize to extracellular region. Is expressed in liver and yolk syncytial layer. Orthologous to human FETUB (fetuin B).
	<i>kif20ba</i>	kinesin family member 20Ba	Predicted to have ATPase activity; microtubule binding activity; and microtubule motor activity. Predicted to be involved in microtubule-based movement. Predicted to localize to kinesin complex; microtubule; and nucleus. Orthologous to human KIF20B (kinesin family member 20B).
	<i>grs1</i>	glutaminyI-tRNA amidotransferase subunit QRSL1	Predicted to have ATP binding activity; amidase activity; and glutaminyI-tRNA synthase (glutamine-hydrolyzing) activity. Predicted to be involved in glutaminyI-tRNA Gln biosynthesis via transamidation and mitochondrial translation. Predicted to localize to glutaminyI-tRNA(Gln) amidotransferase complex and mitochondrion. Orthologous to human QRSL1 (glutaminyI-tRNA amidotransferase subunit QRSL1).
	<i>ddx10</i> <i>si:dkey-111e8.1</i>	DEAD (Asp-Glu-Ala-Asp) box polypeptide 10	Predicted to have RNA binding activity and RNA helicase activity. Predicted to be involved in rRNA processing. Predicted to localize to nucleus. Orthologous to human DDX10 (DEAD-box helicase 10).
	<i>strada</i>	STE20 related adaptor alpha	Predicted to have protein serine/threonine kinase activity. Predicted to be involved in several processes, including activation of protein kinase activity; signal transduction by protein phosphorylation; and stress-activated protein kinase signaling cascade. Predicted to localize to cytoplasm. Orthologous to human STRADA (STE20 related adaptor alpha).
	<i>glrx</i>	glutaredoxin (thioltransferase)	Predicted to have electron transfer activity; protein disulfide oxidoreductase activity; and transferase activity. Predicted to be involved in cell redox homeostasis. Orthologous to human GLRX (glutaredoxin).
	<i>gpn3</i> <i>si:ch1073-412h12.3</i>	GPN-loop GTPase 3	Predicted to have GTPase activity. Orthologous to human GPN3 (GPN-loop GTPase 3).
	<i>upb1</i>	ureidopropionase, beta	Predicted to have beta-ureidopropionase activity. Predicted to be involved in beta-alanine biosynthetic process via 3-ureidopropionate. Is expressed in liver; pronephric duct; and pronephros. Orthologous to human UPB1 (beta-ureidopropionase 1).
<i>vcpkmt</i> <i>si:ch211-91p5.3</i> <i>zgc:65997</i>	valosin containing protein lysine (K) methyltransferase	Orthologous to human VCPKMT (valosin containing protein lysine methyltransferase). Orthologous to human CXorf38 (chromosome X open reading frame 38).	
<i>calm1a</i>	calmodulin 1a	Involved in midbrain-hindbrain boundary morphogenesis. Human ortholog(s) of this gene implicated in catecholaminergic polymorphic ventricular tachycardia 4 and long QT syndrome 14. Is expressed in several structures, including mesoderm; midbrain neural keel; nervous system; neural tube; and trigeminal placode. Orthologous to human CALM1 (calmodulin 1).	
<i>cnot4b</i>	CCR4-NOT transcription complex, subunit 4b	Predicted to have ubiquitin-protein transferase activity. Predicted to be involved in protein ubiquitination. Predicted to localize to CCR4-NOT complex. Orthologous to human CNOT4 (CCR4-NOT transcription complex subunit 4).	

Zbtb48

Zebrafish Zbtb48

Another direct telomere binder examined in this project is ZBTB48 (also known as TZAP or HKR3), which has been identified as a negative regulator of telomere length and a transcriptional activator in human cancer cell lines^{73,109}. ZBTB48 has demonstrated phylogenetic binding to the TTAGGG telomeric sequence in 13 out of 16 vertebrate species, including zebrafish⁷⁹. In my earlier telomere pull-down experiments in zebrafish (Figure 2b), Zbtb48 was repeatedly detected in the zebrafish cell line BRF41, as well as in embryos during the first 5 days of development. Importantly, Zbtb48 was detected in embryos more frequently than any of the shelterin proteins, though it occasionally fell below the enrichment threshold. Given that ZBTB48 is a phylogenetically conserved telomere binder across vertebrate species and is consistently present during zebrafish embryonic development, this suggests its importance in telomere biology and prompting further investigation into its role at the organismal level.

Unlike *hmbx1a/b*, no paralogue was found for *zbtb48* in zebrafish. The zebrafish Zbtb48 also features an N-terminal BTB/POZ domain and 11 C2H2-type zinc fingers (ZNFs) at the C-terminus (Figure 34). A comparison of the amino acid sequences reveals 80% identity in the ZNF domain between the orthologs (V322 – N637, boxed in blue), which is notably higher than the 42% identity in the BTB domain (C28 – D93, boxed in beige). Notably, the telomere sequence recognition motif RxxHxxR⁷⁵ along with the 7 other residues in ZNF11 and the adjacent C-terminal region that interact with the telomeric TTAGGG sequence⁹⁶ are all conserved in zebrafish Zbtb48. The zebrafish Zbtb48 has 72 more amino acids than the human ZBTB48, which these additional residues are primarily distributed outside of the BTB and ZNF domains.

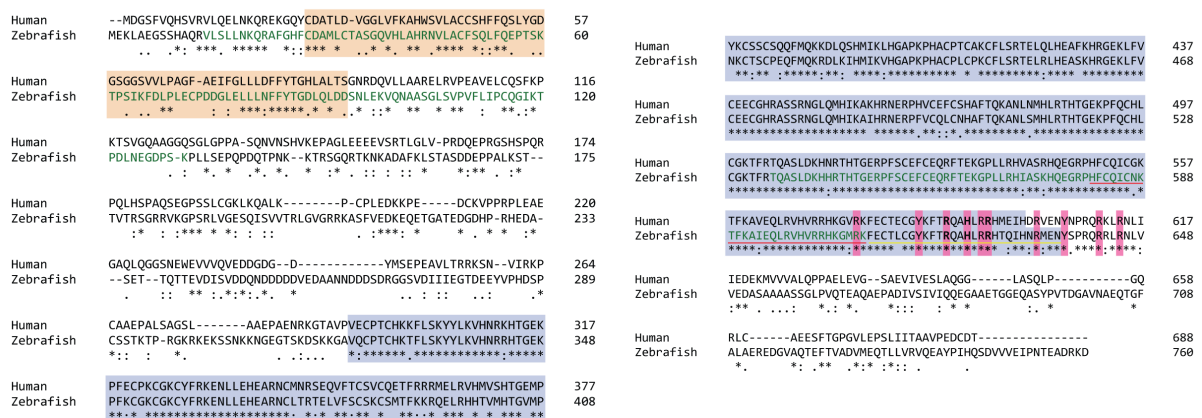


Figure 34. Amino acid sequence alignment of ZBTB48 homologs

Amino acid alignment of the human ZBTB48 protein with the zebrafish Zbtb48 protein. The BTB domain is highlighted in beige, and the ZNF domain is highlighted in blue. Amino acids involved in telomere-binding activity in humans are highlighted in pink, while residues in bold are part of the telomeric sequence recognition motif (RxxHxxR). Sequences underlined in red correspond to ZNF domain 10, and those underlined in yellow correspond to ZNF domain 11. Amino acids written in green represent the epitope sequences used for antibody production.

Creating and validating CRISPR-Cas knockout zebrafish lines

In zebrafish, the *zbtb48* gene is located on chromosome 23 and spans across 11 exons. The BTB/POZ domain is encoded by exon 2, while the eleven adjacent C2H2-type zinc fingers (ZNFs) are located in exons 3 to 11 (Figure 35a). To generate a *zbtb48* CRISPR-Cas knockout zebrafish line, a pair of guide RNAs was designed to target the second exon of *zbtb48*.

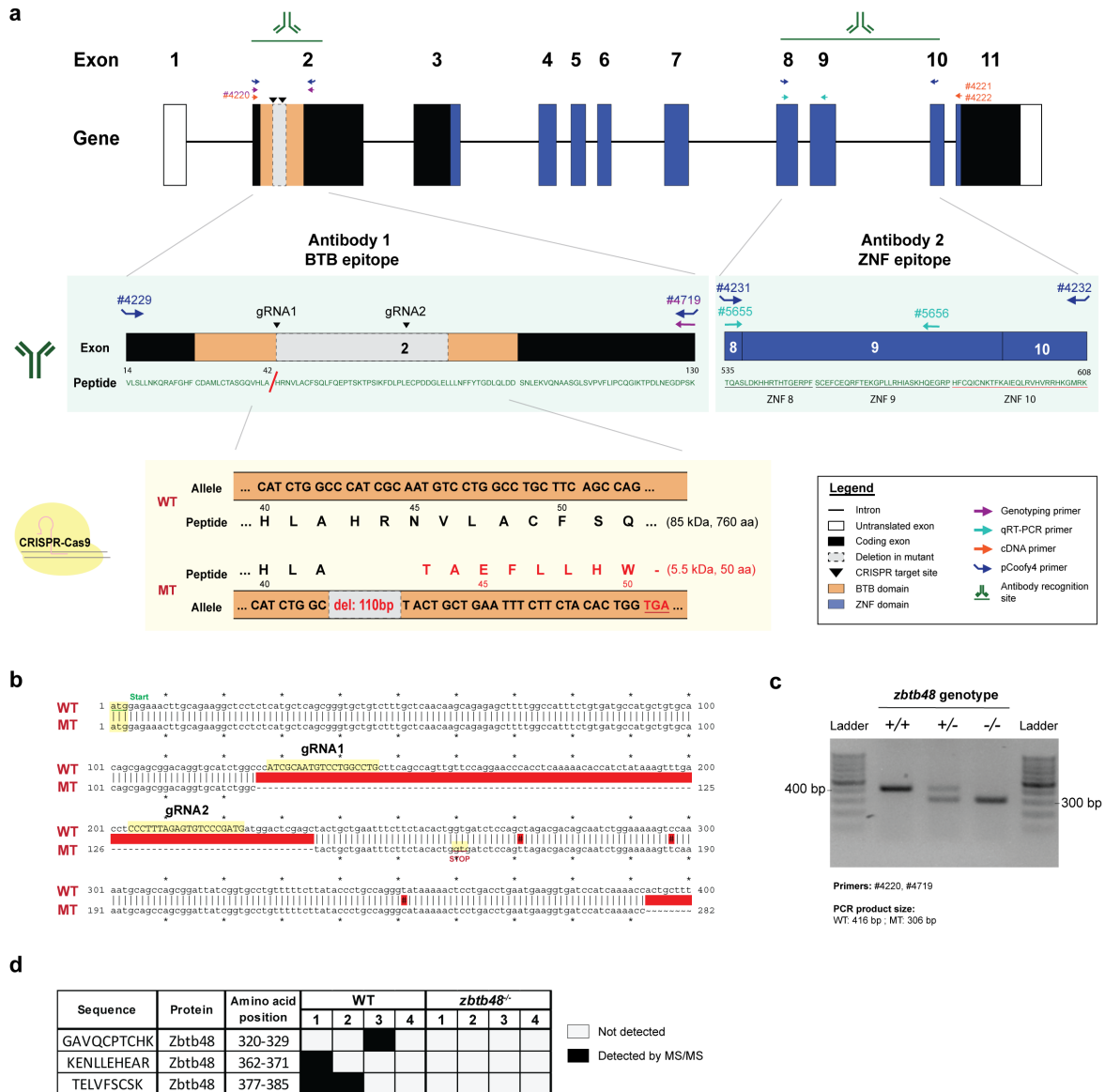


Figure 35. *zbtb48* CRISPR knockout fish line

a) Top: Diagram of the zebrafish *zbtb48* gene constructs, highlighting key features such as CRISPR gRNA target sites, mutant deletion site, antibody epitope site, genotyping and qRT-PCR primer annealing sites, as well as the BTB and ZNF domains. **Middle:** Epitope sequence and its relationship to other key features. **Bottom:** *zbtb48*^{-/-} mutant cDNA sequence and its corresponding amino acid sequence.

b) Sequencing results of the *zbtb48*^{-/-} mutant cDNA sequence compared to the wild-type, with the start, stop, and gRNA sequences highlighted in yellow.

c) Genotyping results using the primers described in (a).

d) Table of Zbtb48 peptides detected in telomere pull-down assays from 5 dpf wild-type and mutant larvae.

Out of the 71 injected embryos, 28 survived to adulthood. However, only 1 founder was identified with successful germline transmission. This mutant line has 110 base pairs deletion in the second exon as confirmed via cDNA sequencing (Figure 35b) and genotyping with PCR (Figure 35c). Deciphering from the cDNA sequence, this deletion would result in a frameshift mutation that produces 50 amino acids long Zbtb48 mutant protein (5.5 kDa), with a truncated BTB domain that lacks the ZNF domain essential for telomere binding (Figure 35a, bottom).

To confirm the loss of telomere binding, a telomere pull-down assay coupled to mass spectrometric analysis was performed on protein lysates from 5 dpf wild-type and *zbtb48*^{-/-} larvae. The result showed Zbtb48 peptides detection in the pull-down with lysates from the wild-type larvae but not in the *zbtb48*^{-/-} larvae (Figure 35d) despite the heightened *zbtb48* mRNA expression level. Taken together, these findings confirmed the success of generating a *zbtb48* knockout zebrafish line.

Generating and validating Zbtb48 antibodies on PAC2

In addition to creating a CRISPR knockout line, zebrafish-specific Zbtb48 antibodies were generated. The antibodies were raised against two unique epitope sequences of zebrafish Zbtb48. One epitope consists of the entire BTB domain along with its flanking sequences and the other contains part of ZNF8 to ZNF10 (residues written in green in Figure 34 and 35a).

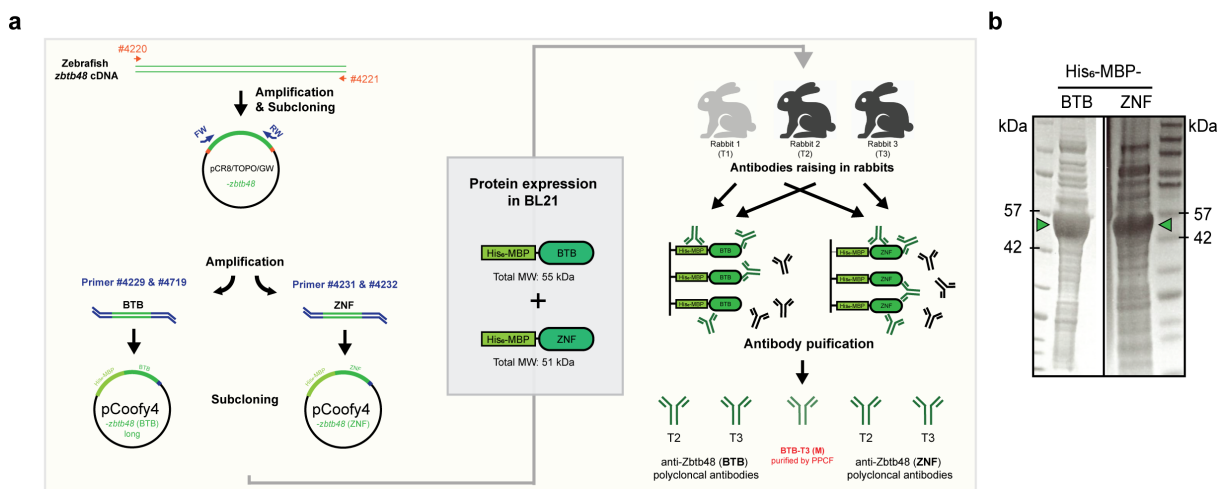


Figure 36. Raising zebrafish Zbtb48 antibodies in rabbits

a) Schematic diagram illustrating the procedures for raising antibodies against Zbtb48. The epitope sequences were amplified from the pCR8/TOPO/GW-*zbtb48* vector and subcloned into a pCoofy4 expression vector, which tags the synthesized epitope sequence with His₆-MBP at the N-terminus. The tagged epitopes were then overexpressed in BL21 *E. coli*. Both epitopes were co-injected into three rabbits, and only sera from animal (T) 2 and 3 were purified after immunization.

b) Gel images showing the overexpression of epitope proteins in BL21 after overnight IPTG induction. The recombinantly expressed His₆-MBP-Zbtb48-BTB epitope is 55 kDa, while His₆-MBP-Zbtb48-ZNF is 51 kDa (marked with green arrowhead).

An overview of the antibody generation process is shown in Figure 36a. The full-length *zbtb48* wild-type cDNA sequence was cloned into pCR8/TOPO/GW before amplifying the epitope sequences from it. The corresponding pairs of primers were used, and the amplicons were subcloned into the expression vector pCoofy4, then transformed into BL21 *E. coli*. Subsequently, the peptides were overexpressed through IPTG induction. This resulted in the production of 55 kDa His₆-MBP-BTB and 51 kDa His₆-MBP-ZNF peptides (Figure 36b). An

equal amount of these peptides were mixed and injected into three rabbits. After two months of immunization, the sera were collected. The different epitopes and sera were purified separately, and 10 elution fractions were collected for each antibody purification. In this study, only rabbit 2 (T2) and rabbit 3 (T3) sera were purified. Thus, yielding four types of polyclonal antibodies that were accordingly named: T2-BTB, T2-ZNF, T3-BTB, and T3-ZNF.

These antibodies were tested in Western blot application on whole-cell protein lysate from the zebrafish cell line PAC2 (Figure 37a). The expected bands corresponding to the molecular weight of zebrafish Zbtb48 protein (85 kDa) were notably absent across all blots. However, two strong bands were observed between the 53 kDa and 70 kDa markers for the T2-BTB antibody (Figure 37ai) while only a faint band was observed above 53 kDa in elution 3 of the T2-ZNF antibody (Figure 37aii), despite a higher chemiluminescent substrate concentration had been used to develop the blot. Meanwhile, the blots with T3-BTB (Figure 37aiii) and T3-ZNF antibodies (Figure 37aiv) have two other similar bands, one at 41 kDa marker and another slightly above 53 kDa marker. However, it is unsure whether the band observed above 53 kDa is identical in both T2 and T3 sera. The coincidence in similar banding patterns observed between the BTB and ZNF antibodies of each serum may suggest that the ZNF antibodies are diluted versions of the respective BTB antibodies. Unfortunately, with the results gathered, it is hard to shortlist any of the antibodies for Western blot use.

Following the Western blot analyses, IP-MS was deployed to assess the antibodies binding to Zbtb48. As with the Hmbox1a/b IP-MS experiments, these were conducted on the nuclear extracts from PAC2 cells, with an enrichment threshold set to a fold change >2 and a p-value < 0.05 relative to the IgG control. To prevent the epitope sequence from influencing Zbtb48 detection, both epitope and MBP sequences were included in the search library.

The IP using T2-BTB and T3-BTB antibodies showed Zbtb48 levels below the detection threshold (Figures 37bi & ii), while no Zbtb48 was detected with the T3-ZNF antibody (Figure 37biii). As anticipated, the exogenous MBP protein was highly enriched in all IP samples. However, both the Zbtb48 peptide sequences identified by MS correspond exclusively to the BTB epitope sequences in the IP that used T2-BTB and T3-BTB antibodies (Figure 37c). This suggests that the endogenous Zbtb48 was not successfully detected by the antibodies in these IPs. Additionally, the quality of the IP experiments appeared to be suboptimal, as only a minimal number of proteins were enriched.

To improve the quality of the antibody, the Protein Production Core Facility (PPCF) was engaged for the purification process. The T3 serum was purified against the BTB epitope, and the second batch of the purified antibody was validated and hereafter labeled as the T3-BTB (M) antibody. The IP-MS experiment was repeated but using twice the protein concentration of the PAC2 nuclear extract and at a higher antibody concentration than before. An aliquot of the T3-BTB (M) antibody was also measured by MS to check for the presence of the epitope. As a result, approximately ten times more proteins were quantified than in Figure 37b, and a drastically higher number of proteins was enriched using the T3-BTB (M) antibody compared to the IgG control (Figure 37di). Despite the continued detection of MBP in the IP with the T3-BTB (M) antibody, the enrichment of Zbtb48 was notably more significant. Detection of endogenous Zbtb48 was evident and fourteen unique Zbtb48 peptide sequences, distinct from the epitope sequences, were identified in the IP-MS (Figure 37dii). Although the MBP protein

remains detectable in the T3-BTB (M) antibody aliquot itself, none of the BTB epitope sequence was detected (Figure 37dii). Therefore, the results from this repeated IP-MS showed a significant improvement compared to the previous result in Figure 37a. Building on this success, the antibody validation was also carried out with *zbtb48* CRISPR knockout zebrafish.

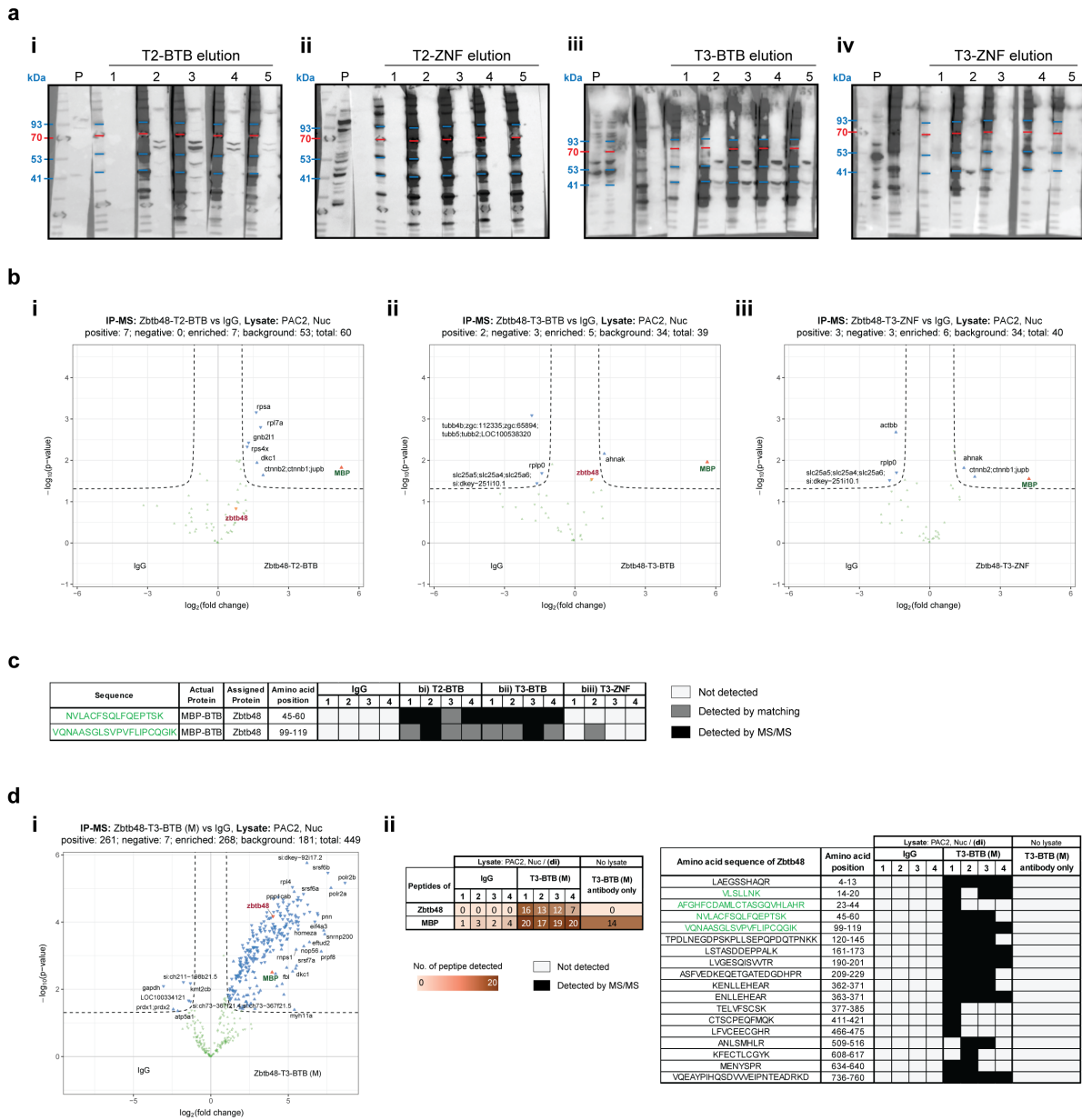


Figure 37. Validating antibodies on PAC2 zebrafish cells

a) Western blot analysis showing the first five elutions for each of the following purified antibodies: **i)** T2-BTB, **ii)** T2-ZNF, **iii)** T3-BTB, and **iv)** T3-ZNF. These antibodies were used as primary antibodies against whole cell lysates from the PAC2 cell line. 'P' represents pre-immunized serum. The expected molecular weight of Zbtb48 is 85 kDa.

b) Volcano plots visualizing the results from the label-free quantification of proteins bound in the IP experiments using **i)** T2-BTB, **ii)** T3-BTB, and **iii)** T3-ZNF antibodies against IgG control, with nuclear extract of the PAC2 zebrafish cell line. The results are plotted with $\log_2(\text{fold change})$ on the x-axis and $-\log_{10}(\text{p-value})$ on the y-axis. The enrichment threshold was set at a fold change > 2 and a p-value < 0.05 (Welch's t-test) with $c = 0.05$. Proteins are annotated with their gene names. Zbtb48 is labelled in red while MBP in green.

c) Zbtb48 peptide sequences detected by mass spectrometry in the IP from (b).

d) Repeat of IP-MS using the Zbtb48-T3-BTB (M) antibody on nuclear extracts from the PAC2 zebrafish cell line. **i)** A volcano plot displaying the IP-MS results, and **ii)** a table on the left showing the number of Zbtb48 and MBP peptides detected in the IP-MS, along with the detected Zbtb48 peptide sequences listed in the table on the right. The epitope sequences were written in green.

Validating Zbtb48 antibodies on *zbtb48*^{-/-} 5 dpf larvae

The generated CRISPR *zbtb48* knockout fish has a deletion that disrupts the BTB domain. However, the mutant Zbtb48 protein, if synthesised, may still be recognizable by T3-BTB (M) antibody because the first 29 amino acids (~25%) of the epitope sequence are still present in the mutant (Figure 35a).

To assess whether the antibody could differentiate between the wild-type Zbtb48 and the mutant version, it was first tested using Western blot analysis. The antibody was probed against whole-cell protein lysates from wild-type and *zbtb48*^{-/-} larvae at 5 dpf, along with the nuclear-enriched protein lysate from PAC2. Unfortunately, no difference in the band pattern was observed between the wild-type and *zbtb48*^{-/-} samples (Figure 38a). This included the faint bands observed between 70 and 93 kDa markers, corresponding to the molecular weight of Zbtb48 (85 kDa), as well as the band above 53 kDa marker, which was previously seen in Figure 37a. Additionally, the two bands seen in the PAC2 nuclear-enriched lysate at 30 kDa and below 53 kDa markers were also present in the wild-type and *zbtb48*^{-/-} mutant lysates. Given the low molecular weight of the mutant Zbtb48 protein (5.5 kDa), it was challenging to detect such a band in the *zbtb48*^{-/-} mutant lysate on the blot. Moreover, due to the low expression levels of Zbtb48 at 5 dpf (as observed in Figure 2) and the presence of numerous nonspecific bands on the blot made it difficult to identify a band that corresponds to Zbtb48 using the Western blot method.

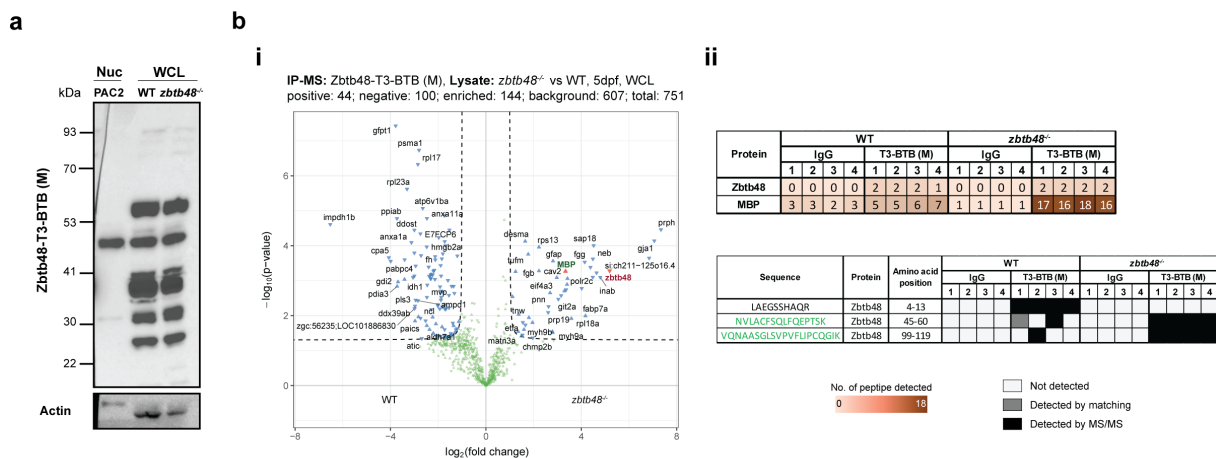


Figure 38. Validating Zbtb48-T3-BTB (M) antibody on the *zbtb48* CRISPR-Cas9 zebrafish line

a) Western blot analysis using Zbtb48-T3-BTB (M) antibodies on whole-cell lysates from 5 dpf wild-type and *zbtb48*^{-/-} mutant larvae. Nuclear extracts from the PAC2 zebrafish cell line were also loaded for comparison. The expected molecular weight of Zbtb48 is 85 kDa, and actin was used as the loading control.

b) IP-MS analysis performed with the Zbtb48-T3-BTB (M) antibody on whole-cell lysates from 5 dpf wild-type and *zbtb48*^{-/-} mutant larvae. **i**) A volcano plot illustrating the IP-MS results, and **ii**) a table at the top showing the number of Zbtb48 and MBP peptides detected, with the Zbtb48 peptide sequences listed in the table at the bottom. The epitope sequences were written in green.

Next, IP-MS was conducted on 5 dpf wild-type and *zbtb48*^{-/-} whole-cell protein lysates using the T3-BTB (M) antibody. Unexpectedly, Zbtb48 was detected in the *zbtb48*^{-/-} sample rather than in the wild-type (Figure 38bi). MBP was also found to be enriched in the *zbtb48*^{-/-} sample, despite the same amount of antibody being used in both IP experiments. Although the number of Zbtb48 peptides detected was similar in both samples, the *zbtb48*^{-/-} sample had three times more MBP peptides than the wild-type (Figure 38bii). Closer analysis revealed that Zbtb48

detection in the *zbtb48*^{-/-} sample was entirely due to the BTB epitope sequences, identical to those identified in Figure 37c. Conversely, these epitope peptide sequences were occasionally detected in the wild-type sample. Notably, one of the sequences corresponds to the CRISPR deletion site of the *zbtb48*^{-/-} mutant, which is more likely from the epitope peptide than being expressed by the mutant. The only other Zbtb48 peptide sequence detected was from amino acids 4 to 13, which was exclusively detected in the wild-type sample and not in the *zbtb48*^{-/-} sample.

Overall, the T3-BTB (M) antibody showed some success in detecting endogenous Zbtb48 with IP-MS while its effectiveness in Western blot cannot be confirmed. However, caution should be taken when interpreting the IP-MS results as it was demonstrated that the epitope peptide used to raise the antibody and for purification was present in the antibody aliquot. Consequently, this epitope peptide was being detected as Zbtb48 and could lead to misinterpretation of results.

Distribution of *zbtb48* genotype and survival

To examine the genotype distribution of the offspring bred from *zbtb48*^{+/-} parents, 500 offspring from eight crosses were genotyped at various developmental stages. Generally, the distribution of all crosses follows the Mendelian inheritance pattern, except for cross 6. Additionally, the data suggested no significant reduction in the fitness of *zbtb48*^{-/-} mutants during development (Figure 39a, $\chi^2 < 5.99$).

Only the larvae from cross 7 (Fish ID: 3032) and cross 8 (Fish ID: 836) were placed in the aquarium to develop into adulthood. These fish were counted routinely, and the survival data were expressed as a percentage of the initial number for each genotype. The *zbtb48*^{-/-} mutants from cross 7 showed a higher survival rate during the first 18 months, but their numbers declined sharply afterward (Figure 39b). A similar trend was observed in cross 8, where *zbtb48*^{-/-} mutants exhibited better survival rates during the first year. However, tracking of this cross was discontinued at one year due to the low number of fish remaining after sacrificing four males from each genotype for proteomic analysis. Overall, these results suggest that *zbtb48*^{-/-} mutants exhibit slightly better survival during the first year of life compared to their wild-type counterparts. However, a larger sample size and additional crosses are needed for a statistically robust analysis to draw a definitive conclusion.

Telomere length in *zbtb48*^{-/-} mutants was assessed in both embryos and adults. Measurements were taken from tail biopsies of 1.5 dpf embryos (cross 1) and 5 dpf larvae (cross 4) using quantitative fluorescence in situ hybridization (q-FISH). To minimize individual variability in mean telomere length, siblings from a single mating of *zbtb48*^{+/-} parents were used. Although *zbtb48*^{-/-} mutants showed a slight increase in telomere length compared to their wild-type siblings at both 1.5 dpf and 5 dpf, the differences were not statistically significant in either case ($p = 0.33$ and 0.55 , respectively) (Figure 39ci). Similarly, no significant differences in telomere length were observed at 8 months of age ($p = 0.46$), as determined by telomere restriction fragment (TRF) analysis (Southern blotting) on caudal fin samples from male fish (cross 7) (Figure 39cii). Therefore, there is no evidence of telomere lengthening in the first generation of *zbtb48*^{-/-} zebrafish mutants, in contrast to reports of telomere lengthening in mammalian ZBTB48 knockout cell lines.

The *zbtb48*^{-/-} mutants were also fertile and capable of propagating for at least three inbred generations, with each generation surviving to adulthood. No physical or behavioral abnormalities were observed in any of the inbred generations of *zbtb48*^{-/-} mutants. For instance, the appearance, weight ($p = 0.4$), and length ($p = 0.44$) of 10.5-month-old second-generation *zbtb48*^{-/-} male mutants (Fish ID: 872) were comparable to their wild-type counterparts (Fish ID: 871) (Figure 39d). Figure 40 summarizes all the crosses and the propagation of the *zbtb48*^{-/-} (*mz114*) mutant line.

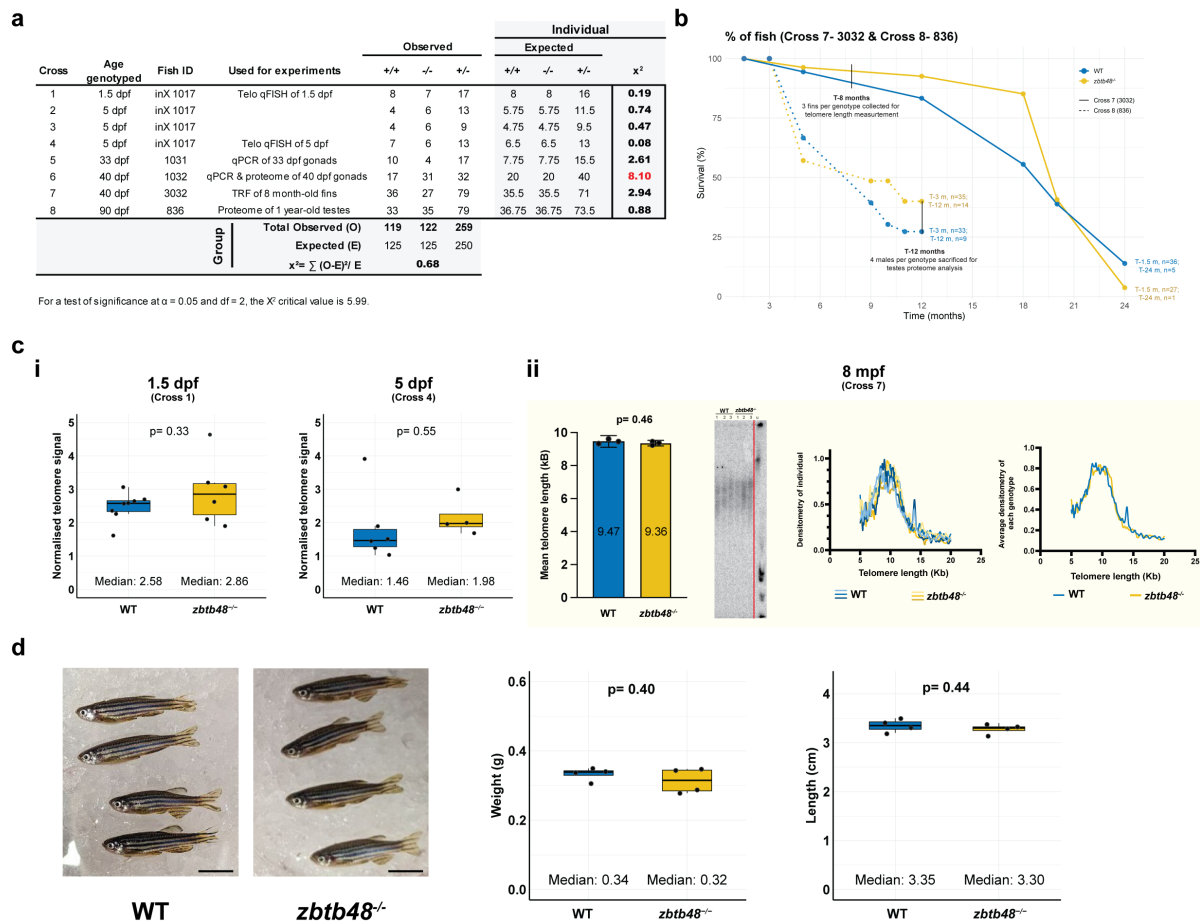


Figure 39. Genotype distribution, survival, phenotypes, and telomere length of CRISPR-Cas9 *zbtb48* mutant lines

a) A table displaying the number of fish of each genotype from eight different crosses between *zbtb48*^{+/-} heterozygous parents. A Chi-square (χ^2) test was conducted for each individual cross as well as an overall test for all three crosses combined. If the observed values significantly differ from the expected values, the χ^2 statistic must exceed the critical value of 5.99 for an alpha level of 0.05 and 2 degrees of freedom (df). Age of the fish was represented in days-post fertilization (dpf).

b) A graph showing the percentage of fish remaining in the tank over time for the first mutant generation from Cross 7 (Fish ID: 3032) and Cross 8 (Fish ID: 836). Fish counts were taken periodically, with the age (T) and number (n) of fish at the start and end of each count displayed alongside the graph.

c) Telomere length was measured in the first generation of *zbtb48*^{-/-} mutants and their wild-type siblings. **i)** Telomere length was assessed in 1.5 dpf embryos (Cross 1) and 5 dpf larvae (Cross 4) using the qFISH method, with telomere signals from tail fin biopsies normalized against DAPI staining. Each dot represents the normalized mean telomere signal of an individual, with p-values calculated by Welch's t-test. **ii)** Telomere length was quantified by telomere restriction fragment (TRF) analysis from caudal fin samples of 8-month-old first-generation males ($n = 3$ per genotype; Cross 7, Fish ID: 3032). The mean telomere length for wild-type and *zbtb48*^{-/-} mutants was represented in a bar plot on the left, with error bars showing the standard error of the mean (SEM) and the p-value determined by an unpaired two-tailed Mann-Whitney test. The central panel shows Southern blot results from the TRF analysis, and two graphs on the right present densitometry results: one for individual samples and the other grouped by genotype.

d) A photograph of second-generation *zbtb48*^{-/-} male zebrafish (Fish ID: 872) and their wild-type counterparts (Fish ID: 871) at 10.5 months of age. The scale bar represents 1 cm. The weight and size of these fish were measured and displayed in a box plot on the right, with p-values calculated using Welch's t-test.

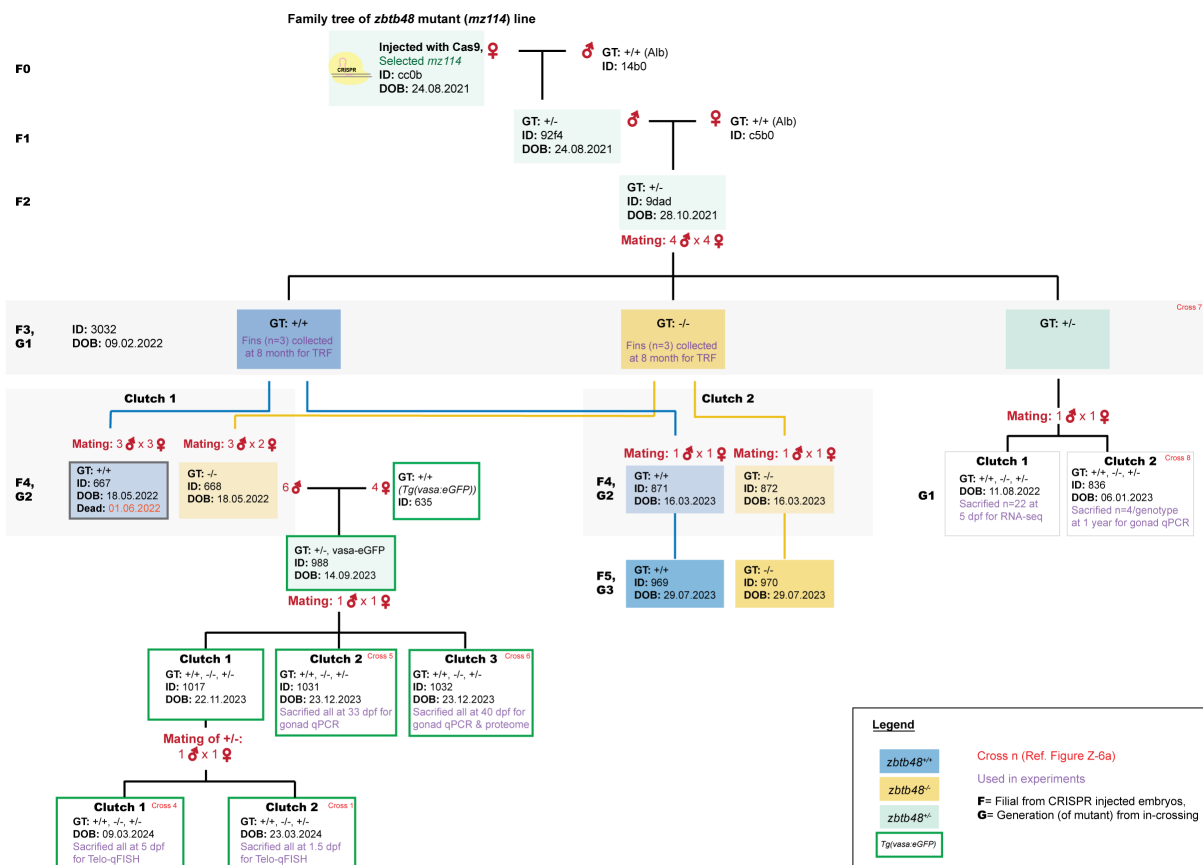


Figure 40. Family tree of CRISPR-Cas9 *zbtb48* mutant lines

The family tree of CRISPR-Cas9 *zbtb48* mutant lines generated during this study. Vertical lines represent a new generation, while horizontal lines indicate matings between parents with different fish IDs (noted with the number of females and males beside them). If only a vertical line is shown without a horizontal line, it represents incrosses. When the number of males and females is not specified, it refers to a single mating. Each box represents a clutch and is color-coded based on genotype. It details the fish genotype, ID, cross number, and the date of birth (DOB). Only the tank that dies prematurely was noted with the date of death. A green border indicates a transgenic line, *Tg(vasa:eGFP)*. Generations are labeled as "F" (filia), starting from the CRISPR-injected embryos, while "G" labels denote homozygous mutant generations resulting from incrossing.

Transcriptomic and proteomic profiling of 5 dpf

Apart from regulating telomere length, ZBTB48 has been reported to be a transcriptional regulator in human cancer cells⁷³. To investigate whether this function is conserved in zebrafish, transcriptome and proteome analyses were conducted on 5 dpf *zbtb48*^{-/-} mutant and wild-type larvae.

For RNA sequencing (RNA-seq) analysis (Figure 41a), four wild-type and *zbtb48*^{-/-} siblings bred from *zbtb48*^{+/-} parents through single mating were collected and used as biological replicates. Among the 32,058 genes quantified by RNA-seq, only 25 genes exhibited downregulation, and 7 genes showed upregulation in the mutant (fold change > |2|, p-value < 0.01). Notably, *zgc:153284*, which is predicted to be involved in cell redox homeostasis, was the most strongly deregulated gene, with a 163-fold ($p = 4.14e^{-44}$) increase in expression in the *zbtb48*^{-/-} larvae.

Interestingly, one of the transcripts that was upregulated in *zbtb48*^{-/-} mutant is *zbtb48*, with 3.2-fold increase ($p = 6.59e^{-08}$). This upregulation was also reflected in the qPCR, which showed a significant 4.6-fold increase ($p = 0.004$) in *zbtb48*^{-/-} mutant (Figure 41b). An investigation of the RNA-seq tracks revealed a lack of sequencing reads in the CRISPR deletion region for the *zbtb48*^{-/-} mutants (Figure 41c).

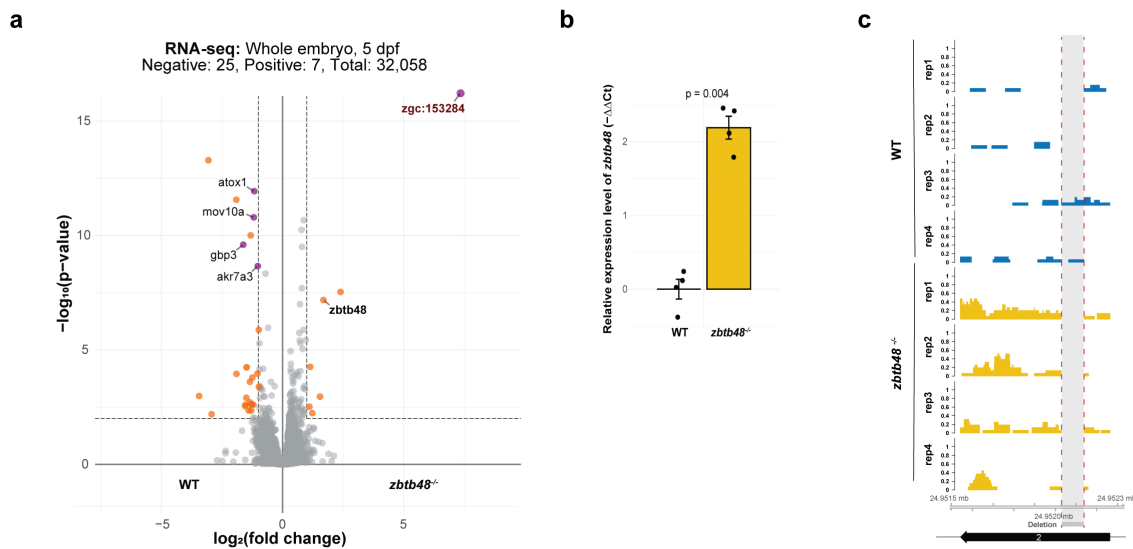


Figure 41. Transcriptomic profiling of 5dpf *zbtb48*^{-/-} mutant larvae

a) Volcano plot with quantitative comparison of RNA sequencing (RNA-Seq) data from 5 dpf wild-type and *zbtb48*^{-/-} mutant larvae ($n=4$). The results were log-transformed and plotted on the x-axis as $\log_2(\text{fold change})$ and on the y-axis as $-\log_{10}(\text{p-value})$ (Welch's t-test). The transcript enrichment thresholds were set at a fold change $> |2|$ and a p-value < 0.01 . Genes annotated in this plot were also differentially expressed in the proteome analysis (b). Note: the p-value ($4.14e^{-44}$) of *zgc:153284* is beyond the upper limit of the y-axis.

b) qRT-PCR analysis comparing *zbtb48* mRNA levels between wild-type and mutant larvae at 5 dpf. The experiment was conducted in biological quadruplicate (25 pooled larvae) with technical triplicates each. Error bars represent the standard error of the mean (SEM), and p-values were calculated using two-tailed Welch's t test.

c) RNA-seq data showing *zbtb48* mRNA levels in 5 dpf wild-type and mutant larvae. The CRISPR deletion site is marked with red dotted lines on the track.

Next, two biological replicates of proteome analysis were performed on independently collected 5 dpf *zbtb48*^{-/-} mutant larvae. Wild-type and *zbtb48*^{-/-} larvae were bred from parents of the respective genotype, pooled, and used as technical replicates in each collection. In the first replicate (Figure 42ai), a total of 8,075 proteins were quantified, with 80 downregulated and 109 upregulated (fold change $> |2|$, p-value < 0.05). In the second experiment from a new crossing (Figure 42aii), a total of 7,601 proteins were quantified, 99 of which were downregulated and 152 of which were upregulated. Between the two replicates, some variation was observed. However, 11 proteins (*Akr7a3*, *Cpa4*, *Ctrb.2*, *Cyp2aa7*, *Cyp2p8*, *Golgb1*, *Hdac10*, *Lmf2b*, *Mtfp1*, *Olfml3a*, and *Zgc:111983*) were consistently downregulated, while 10 proteins (*Coch*, *Col28a2a*, *Emc1*, *Myl7*, *Slc27a1a*, *Vtg2*, *Vtg3*, *Vtg5*, *Vtg7*, and *Zgc:153284*) were consistently upregulated in *zbtb48*^{-/-} mutants in both proteome analyses (Figure 42b).

A small set of genes exhibited significant changes in transcript and protein expression levels (Figure 42b), including *akr7a3*, *atox1*, *gbp3* and *mov10a*, which were downregulated, while *zgc:153284* was upregulated. The downregulation of *akr7a3* and *atox1*, as well as the upregulation of *zgc:153284* transcripts were validated on 5 dpf *zbtb48*^{-/-} larvae by qRT-PCR (Figure 42c). Additionally, consistent with the proteome data, the downregulation of the *hdac10*

and *mtfp1* transcripts and upregulation of the *col28a2a* transcript were detected in *zbtb48*^{-/-} larvae by qRT-PCR (Figure 42c), which were not significantly different at the transcript level according to the chosen RNA-seq cut-offs.

These genes cover diverse biological aspects (Table 2), but one of them, *Mtfp1*, is particularly interesting because its downregulation has also been reported in human ZBTB48 knockout cell lines⁷³. Thus, the investigation was extended to other genes reported to be regulated by ZBTB48 in humans, namely, *PXMP2*, *PP3CB*, and *VWA5A*, by examining the transcription levels of their zebrafish orthologs via qRT-PCR. No differences in the expression of *pp3cb* or *vwa5a* between *zbtb48*^{-/-} and wild-type zebrafish larvae at 5 dpf was observed, and only a slight 1.2-fold ($p = 0.03$) decrease in *pxmp2* expression was detected (Figure 42c). Therefore, strong *mtfp1* downregulation upon *zbtb48* knockout at the transcript and proteome levels was the most pronounced similarity between humans and zebrafish.

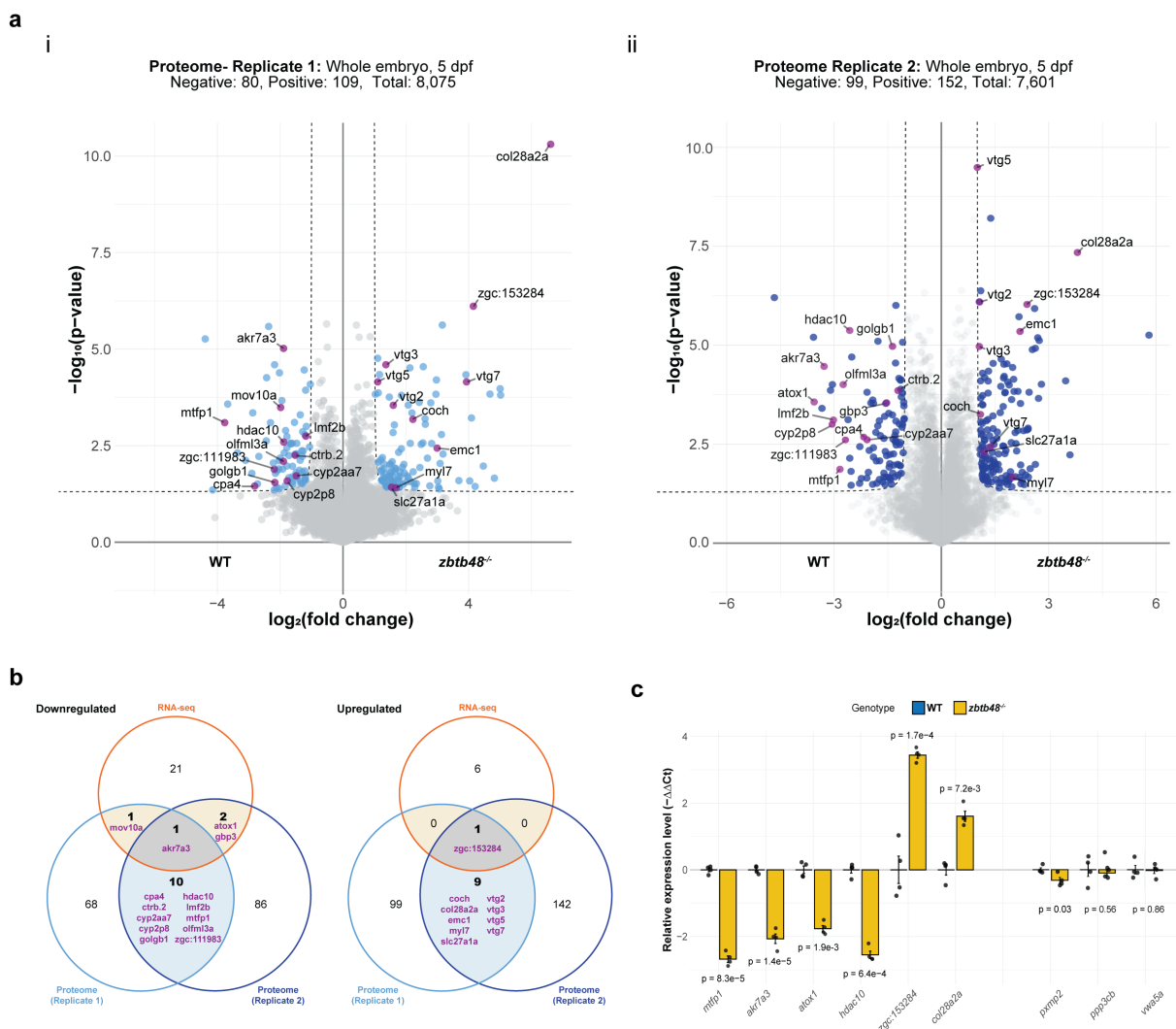


Figure 42. Proteomic profiling of 5dpf *zbtb48*^{-/-} mutant larvae

a Volcano plots for two biologically independent proteome analyses (**i** and **ii**) conducted on pooled 5 dpf wild-type and *zbtb48*^{-/-} mutant larvae. The results were log-transformed and plotted on the x-axis as $\log_2(\text{fold change})$ and on the y-axis as $-\log_{10}(p\text{-value})$ (Welch's t-test). The protein enrichment threshold was set at a fold change $> |2|$ and a $p\text{-value} < 0.05$.

b Venn diagrams showing downregulated (left) and upregulated (right) transcripts or proteins in the *zbtb48*^{-/-} mutants compared to the wild-type.

c qRT-PCR analysis of selected genes in 5 dpf larvae of wild-type and *zbtb48*^{-/-} mutant zebrafish. The set includes differentially expressed genes from this study (left) and previously reported human target genes (Jahn et al., 2017) (right). The experiment was conducted in biological quadruplicate (25 pooled larvae) with technical triplicates each. Error bars represent the standard error of the mean (SEM), and $p\text{-values}$ were calculated using two-tailed Welch's t test.

Table 2. Summary of dysregulated proteins identified in 5 dpf *zbtb48*^{-/-} mutant larvae

	Gene name	Full gene name	Zfin description	
Upregulated	<i>coc</i>	coagulation factor C homolog, cochlinal (Limulus polyphemus)	Predicted to be involved in growth plate cartilage chondrocyte morphogenesis. Predicted to localize to collagen-containing extracellular matrix and extracellular space. Human ortholog(s) of this gene implicated in autosomal dominant nonsyndromic deafness 9 and autosomal recessive nonsyndromic deafness 110. Orthologous to human COCH (cochlinal).	
	<i>col28a2a</i>	collagen, type XXVIII, alpha 2a	Predicted to have extracellular matrix structural constituent. Predicted to be involved in extracellular matrix organization. Predicted to localize to extracellular matrix and extracellular space. Is expressed in floor plate; integument; muscle; rhombomere 5; and white matter. Orthologous to human COL28A1 (collagen type XXVIII alpha 1 chain).	
	<i>emc1</i>	ER membrane protein complex subunit 1	Predicted to localize to ER membrane protein complex. Orthologous to human EMC1 (ER membrane protein complex subunit 1).	
	<i>myl7</i>	myosin, light chain 7, regulatory	Predicted to have calcium ion binding activity. Involved in several processes, including cardiac muscle cell proliferation; heart contraction; and myofibril assembly. Is expressed in several structures, including corpuscles of Stannius; heart primordium; mesoderm; muscle; and pericardial region. Orthologous to human MYL7 (myosin light chain 7).	
	<i>slc27a1a</i>	solute carrier family 27 member 1a	Predicted to have fatty acid transmembrane transporter activity; long-chain fatty acid transporter activity; and long-chain fatty acid-CoA ligase activity. Predicted to be involved in several processes, including fatty acid transport; glycerophospholipid biosynthetic process; and regulation of phosphate metabolic process. Predicted to colocalize with mitochondrial inner membrane. Orthologous to human SLC27A1 (solute carrier family 27 member 1).	
	<i>vtg2</i>	vitellogenin 2	Exhibits lipopolysaccharide binding activity and lipoteichoic acid binding activity. Involved in positive regulation of phagocytosis, engulfment and response to estradiol. Is expressed in atrium; intestine; liver; and ovary.	
	<i>vtg3</i>	vitellogenin 3, phosphatidyl	Predicted to have lipid transporter activity. Involved in cellular response to estrogen stimulus. Is expressed in several structures, including axis; digestive system; hatching gland; ovary; and sensory system.	
	<i>vtg5</i>	vitellogenin 5	Predicted to have lipid transporter activity. Predicted to be involved in cellular response to estrogen stimulus and response to estradiol. Is expressed in intestine; liver; and ovary.	
	<i>vtg7</i>	vitellogenin 7	Predicted to have lipid transporter activity. Predicted to be involved in cellular response to estrogen stimulus and response to estradiol. Is expressed in liver; pleuropertoneal region; skeletal muscle cell; and unfertilized egg.	
	<i>zgc:153284</i>		Predicted to have electron transfer activity and protein disulfide oxidoreductase activity. Predicted to be involved in cell redox homeostasis.	
	Downregulated	<i>akr7a3</i>	aldo-keto reductase family 7, member A3 (afatoxin aldehyde reductase)	Is expressed in gut; liver; myotome; and optic vesicle. Orthologous to several human genes including AKR7A3 (aldo-keto reductase family 7 member A3).
		<i>atox1</i>	antioxidant 1 copper chaperone	Predicted to have copper chaperone activity. Predicted to be involved in cellular copper ion homeostasis and response to oxidative stress. Predicted to localize to cytosol. Orthologous to human ATOX1 (antioxidant 1 copper chaperone).
<i>cpa4</i>		carboxypeptidase A4	Predicted to have metallocarboxypeptidase activity. Predicted to be involved in proteolysis. Predicted to localize to extracellular space. Is expressed in pancreas and somite. Orthologous to human CPA4 (carboxypeptidase A4).	
<i>ctrb.2</i>		chymotrypsinogen B, tandem duplicate 2	Predicted to have serine-type endopeptidase activity. Predicted to be involved in proteolysis.	
<i>cyp2aa7</i>		cytochrome P450, family 2, subfamily AA, polypeptide 7	Predicted to have heme binding activity; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen; and steroid hydroxylase activity. Predicted to be involved in several processes, including exogenous drug catabolic process; organic acid metabolic process; and xenobiotic metabolic process. Predicted to localize to cytoplasm and intracellular membrane-bounded organelle.	
<i>cyp2p8</i>		cytochrome P450, family 2, subfamily P, polypeptide 8	Predicted to have heme binding activity; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen; and steroid hydroxylase activity. Predicted to be involved in several processes, including exogenous drug catabolic process; organic acid metabolic process; and xenobiotic metabolic process. Predicted to localize to cytoplasm and intracellular membrane-bounded organelle. Human ortholog(s) of this gene implicated in hypertension; myocardial infarction; pulmonary hypertension; renal fibrosis; and renal hypertension. Orthologous to human CYP2J2 (cytochrome P450 family 2 subfamily J member 2).	
<i>gpb3</i>		guanylate binding protein 3	Predicted to have GTP binding activity and GTPase activity. Predicted to be involved in inflammatory response; innate immune response; and regulation of apoptotic process.	
<i>golgb1</i>		golgin B1	Predicted to localize to Golgi apparatus and integral component of membrane. Used to study hyperphosphatemic familial tumoral calcinosis. Is expressed in axis; floor plate; hypochord; otic placode; and polster. Orthologous to human GOLGB1 (golgin B1).	
<i>hdac10</i>		histone deacetylase 10	Exhibits acetylsermidine deacetylase activity and zinc ion binding activity. Involved in spermidine deacetylation. Predicted to localize to cytoplasm. Orthologous to human HDAC10 (histone deacetylase 10).	
<i>lmf2b</i>		lipase maturation factor 2b	Predicted to be involved in protein maturation. Predicted to localize to endoplasmic reticulum membrane. Is expressed in several structures, including central nervous system; hatching gland; hindbrain neural plate; notochord; and polster. Orthologous to human LMF2 (lipase maturation factor 2).	
<i>mov10a</i>		Mov10 RNA helicase a	Predicted to have RNA binding activity. Predicted to be involved in posttranscriptional gene silencing by RNA. Predicted to localize to P granule and cytosol. Human ortholog(s) of this gene implicated in hypertension. Orthologous to human MOV10 (Mov10 RISC complex RNA helicase).	
<i>mtfp1</i>		mitochondrial fission process 1	Predicted to be involved in mitochondrial fission. Predicted to localize to mitochondrial inner membrane. Is expressed in visceral fat. Orthologous to human MTFP1 (mitochondrial fission process 1).	
<i>olfml3a</i>		olfactomedin-like 3a	Predicted to be involved in multicellular organism development. Predicted to localize to extracellular region. Orthologous to human OLFML3 (olfactomedin like 3).	
<i>zgc:111983</i>			Predicted to localize to integral component of membrane. Is expressed in epidermis; periderm; pharynx; post-vent region; and sensory system. Orthologous to human APOBR (apolipoprotein B receptor).	

Validating the downregulation of *mtfp1* and its phenotype in *zbtb48*^{-/-} embryos

MTFP1 (mitochondrial fission process 1), a nuclear-encoded protein located in the inner mitochondrial membrane, acts as a negative regulator of inner mitochondrial membrane fusion. ZBTB48 has been reported to bind to the promoter regions of MTFP1 in U2OS and HeLa cells⁷³. Knocking out ZBTB48 in these cells resulted in a loss of MTFP1 expression, accompanied by mitochondrial fusion and hyperbranching of the mitochondrial network, similar to the phenotype observed in MTFP1 knockout cells^{73,150}. To investigate whether these findings could be recapitulated in zebrafish, experiments were conducted on *zbtb48*^{-/-} mutant embryos.

Unfortunately, the binding of Zbtb48 to the *mtfp1* promoter in zebrafish could not be verified due to limitations with the zebrafish Zbtb48 antibody for chromatin immunoprecipitation sequencing (ChIP-seq) applications. Notably, the promoter sequence of MTFP1 differs between humans and zebrafish, and no match for the MTFP1 promoter region bound by ZBTB48 in humans was found in zebrafish when using BLAST.

Despite the inability to confirm Zbtb48 binding to the *mtfp1* promoter in zebrafish, significant downregulation of *mtfp1* was observed in *zbtb48*^{-/-} mutant zebrafish, as assessed by qRT-PCR (Figure 42c). However, this downregulation was not statistically significant in RNA-seq (Figure 41a), although both experiments were performed on larvae of the same developmental stage. Nevertheless, the proteome analyses (Figure 42a) consistently showed a significant

downregulation of Mtfp1 protein in the *zbtb48*^{-/-} mutant larvae, aligning with proteomic results reported in ZBTB48 knockout U2OS and HeLa cells ⁷³.

In addition to identifying the downregulation of MTFP1 in human cell lines, Jahn et al. ⁷³ used Western blot analysis to demonstrate the loss of MTFP1 expression in *zbtb48* knockout cells. To recapitulate this result in zebrafish, Western blot analysis was attempted on whole-cell lysates from 5 dpf larvae. However, due to the lack of a zebrafish-specific Mtfp1 antibody, a human MTFP1 antibody was used, assuming it would cross-react due to the 75% identity between zebrafish Mtfp1 and its human ortholog (Figure 43a). While a strong MTFP1 band was observed at 18 kDa in U2OS cell lysates, it was only weakly detected in the zebrafish wild-type lysates, which this band was also weakly detected in *zbtb48*^{-/-} (Figure 43b). Therefore, this attempt of using human MTFP1 antibody failed to confirm the loss of Mtfp1 in *zbtb48*^{-/-} mutant larvae. Nevertheless, the proteome data from 5 dpf larvae (Figure 42a) revealed a reduced number of Mtfp1 peptides in *zbtb48*^{-/-} mutants, compared to the 3 to 8 peptides found in wild-type samples (Figure 43c).

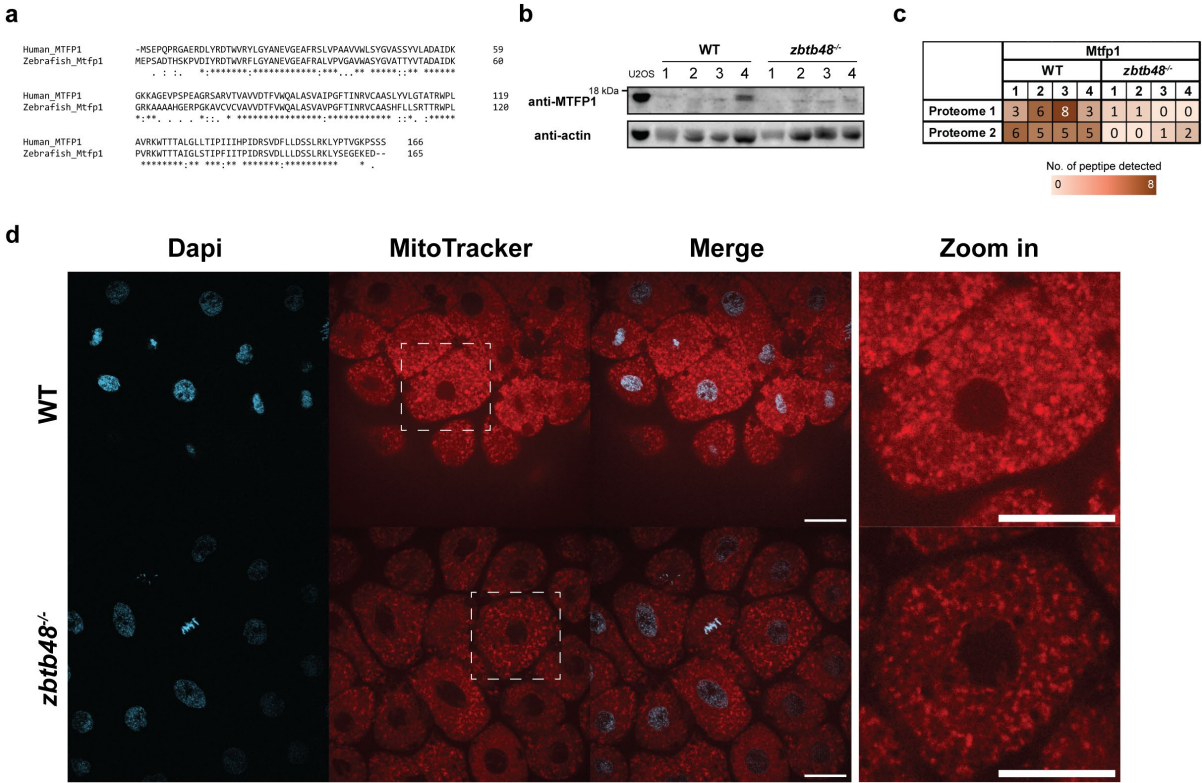


Figure 43. Validation of Mtfp1 loss in *zbtb48*^{-/-} mutants
a) Amino acid alignment of the human MTFP1 protein with the zebrafish Mtfp1 protein.
b) Western blot analysis of whole-cell protein lysates from 5 dpf wild-type and *zbtb48*^{-/-} mutant larvae, probed with an human MTFP1 antibody. U2OS, a human cell line, was used as a positive control. Actin served as a loading control. The expected molecular weight of Mtfp1 is 18 kDa.
c) Table showing the number of Mtfp1 peptides detected in the 5 dpf proteome analyses by MS.
d) Wild-type and *zbtb48*^{-/-} embryos stained with MitoTracker Deep Red FM and DAPI at 3 hpf. The boxed cells are enlarged in the images on the right. Images were acquired using a 100X objective on a BC43 confocal microscope (Andor). Scale bar = 20 μm.

Lastly, to investigate whether the downregulation of Mtfp1 caused by the knockout of *zbtb48* in zebrafish would lead to mitochondrial elongation, a MitoTracker assay was performed on embryos at the blastula stage (approximately 3 hpf) to visualize mitochondrial morphology. Unlike the tubular mitochondrial network typically observed in human cell lines, the mitochondria in wild-type zebrafish embryos exhibited a clustered organization. No distinguishable morphological differences were observed between *zbtb48*^{-/-} mutants and wild-type embryos, and mitochondrial network elongation was not detected in the *zbtb48*^{-/-} mutants (Figure 43d).

Spatiotemporal expression of *zbtb48*

Thus far, most experiments have been conducted on 5 dpf larvae. However, the expression of *Zbtb48* appears to be low at 5 dpf, as it was not detected in both proteome analyses (Figure 42a). Similarly, RNA-seq data from a developmental zebrafish study (Figure 4) shows that *zbtb48* mRNA levels are among the lowest at 5 dpf. Higher *zbtb48* transcript levels were observed during earlier development, aligning with the detection of *Zbtb48* protein in the telomere pull-down assay (Figure 2b). The highest expression of *zbtb48* was seen at the 128-cell to 1k-cell stage (approximately 2 to 3 hpf), prior to zygotic genome activation.

To further explore the spatiotemporal expression of *zbtb48*, publicly available single-cell RNA-seq (scRNA-seq) data was analyzed. Generally, a gradual decline of *zbtb48* expression was observed during zebrafish embryonic development^{151,152} (Figure 44a). At 5-6 hpf, the highest *zbtb48* expression was observed in the endoderm, while by 10-12 hpf, the spinal cord (derived from ectoderm) had the highest transcription levels, followed by the liver (derived from endoderm). *zbtb48* was also expressed in primordial germ cells (PGCs), though at about half the level seen in the spinal cord and liver. By 5 dpf (or 120 hpf), *zbtb48* transcript levels had decreased in both the spinal cord and liver, with expression predominantly observed in PGCs.

To investigate whether this strong expression of *zbtb48* in larval PGCs is sustained in adult gonads, analysis was extended to a published scRNA-seq dataset of adult tissues¹⁵³. *zbtb48* was prominently expressed in the testis, although it was detected in less than 0.25% of cells (Figure 44b). Strikingly, the ovary exhibited 30-fold lower expression compared to the testis. In addition to the gonads, low levels of *zbtb48* transcripts were detected in the skin, brain, fins, and spleen. These findings collectively suggest that *zbtb48* expression is predominantly localized to the gonads, specifically within a niche population of cells.

To identify the specific cell populations expressing *zbtb48* in the gonads, scRNA-seq studies of zebrafish testes¹⁵⁴ and ovaries¹⁵⁵ were analyzed. In the testes, the highest expression of *zbtb48* was found in spermatocytes (undergoing meiosis), followed by spermatogonia (stem cells) and early round spermatids, regardless of age (Figure 44c). Notably, *zbtb48* expression was high but confined to a small number of cells compared to *cdh17*, a gene expressed in nephrons, liver, and intestine¹⁴⁰ used as a negative control, and *ddx4* (also known as *vasa*), strongly expressed in the germline^{156,157}, used as a positive control. Similarly, analysis of juvenile ovaries at 40 dpf revealed the highest *zbtb48* expression in the germline stem cell population, followed by meiotic cells and early oocytes, with comparisons again made to *cdh17* and *ddx4* (Figure 44d).

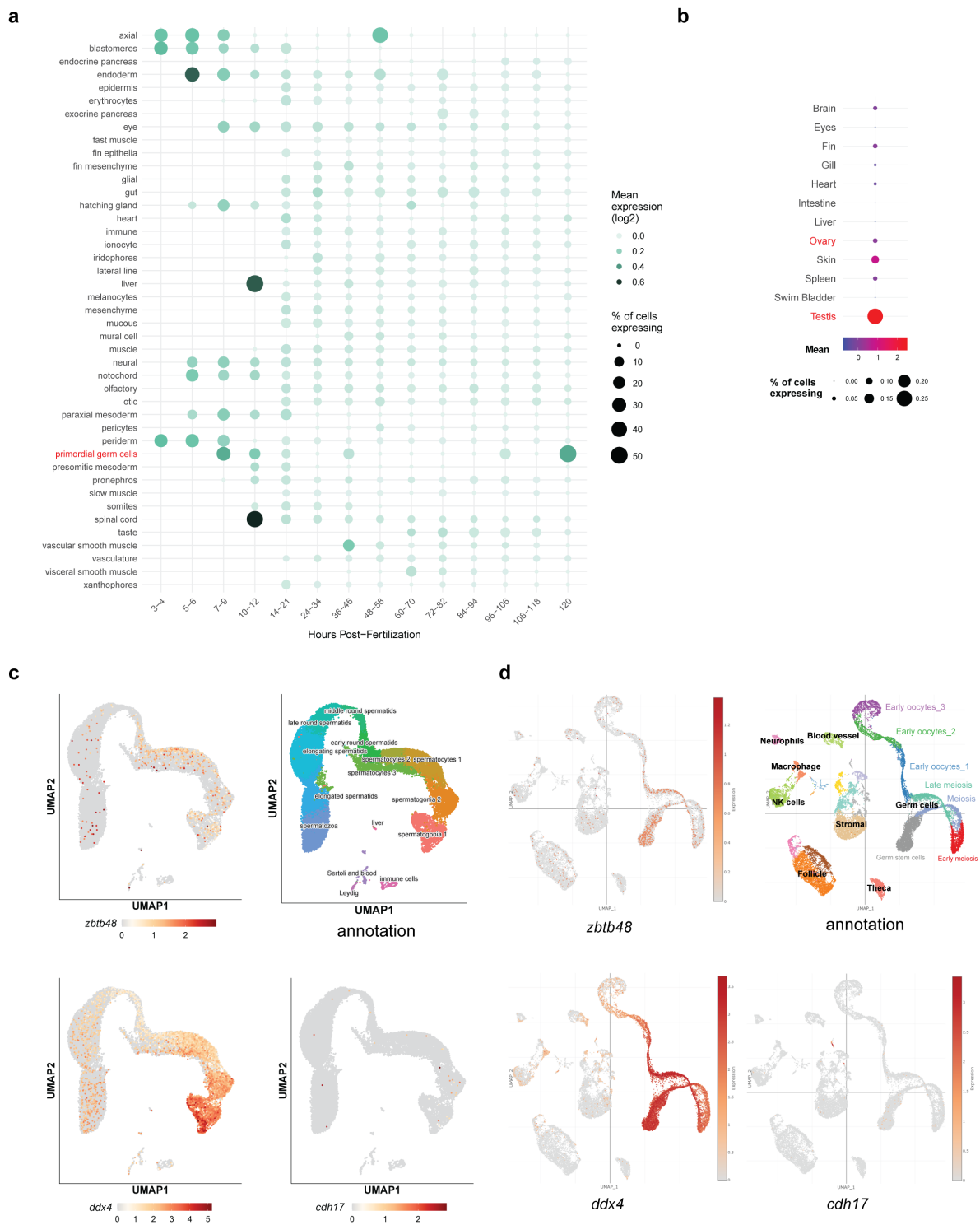


Figure 44. Spatiotemporal expression of *zbtb48* transcripts in larvae and adults

a) Heatmap showing *zbtb48* expression across various tissues during the first five days of development, based on scRNA-seq data from Farrell et al. (2018) and Sur et al. (2023).

b) Heatmap depicting *zbtb48* expression in adult tissues, based on scRNA-seq data from Jiang et al. (2021).

c) UMAP plots from scRNA-seq of testes from fish aged 5 to 22 months. Figures were obtained from the public website (Sposato et al., 2023).

d) UMAP plots from scRNA-seq data of ovaries from 40 dpf fish. Figures were obtained from a public website (Liu et al., 2022). For panels c) and d), the top left shows *zbtb48* mRNA expression, the top right shows cell population annotations, the bottom left shows *ddx4* (also known as *vasa*, highly expressed in germ cells), and the bottom right shows *cdh17* (absent in germ cells).

Overall, these data underscore the expression of *zbtb48* transcripts in the gonads, albeit in a small subset of cells and at lower levels compared to the germ cell marker *ddx4*. The higher expression of *zbtb48* in these highly proliferative tissues may be attributed to its role in telomere regulation.

Validating *Zbtb48* expression in embryo through IF

Following the identification of high *zbtb48* expression during early embryonic development and in germ cells, IF experiments were carried out to assess the specificity of the generated *Zbtb48* antibodies. As previously noted in IP-MS, BTB antibodies may detect the truncated BTB domain in the mutant *Zbtb48* protein, potentially producing a signal in *zbtb48*^{-/-} mutants. To address this concern, ZNF antibodies, which should not recognize the mutant protein due to the absence of the ZNF domain, were included in the test. The primary goal of the experiment is to assess the suitability of these antibodies for IF. At the same time, this will help to address whether the BTB antibodies detect the mutant *Zbtb48* protein and if the ZNF antibodies were successfully produced or simply diluted versions of the BTB antibodies.

Prior to the specialization of cells into PGCs, germline materials are localized within germplasms at the cleavage furrows during the cleavage stages¹⁵⁸, which are identifiable at the 4-cell stage. In this IF experiment, both T2-BTB and T2-ZNF antibodies stained the cleavage furrows of wild-type 4-cell stage embryos. However, this signal was not diminished in the *zbtb48*^{-/-} mutants when using both antibodies (Figure 45a, arrowhead). In contrast, no cleavage furrow staining was detected in wild-type embryos probed with T3-BTB and T3-ZNF antibodies. Since all embryos were stained and imaged simultaneously with the same brightness and contrast settings applied to the images, the weaker but consistent staining pattern observed with ZNF antibodies suggests that they may be a diluted version of the respective BTB antibodies. Therefore, the ZNF antibodies are unsuitable for validating *Zbtb48* expression in *zbtb48*^{-/-} mutant as initially intended. However, given the differing staining patterns between antibodies the two sera, and the lack of any difference between T2-BTB and T2-ZNF staining in both wild-type and *zbtb48*^{-/-} mutant embryos, it remains unclear whether the cleavage furrow staining is truly a result of *Zbtb48* recognition or due to an off-target effect.

To verify that the T2 antibody signal in the cleavage furrow was indeed located within the germplasm, 4-cell stage embryos were counterstained with the Ziwi antibody, a germ cell marker¹⁵⁹. As anticipated, the *Zbtb48*-T2-BTB signal co-localized with Ziwi in both wild-type and *zbtb48*^{-/-} embryos (Figure 45b). This finding provided several important pieces of information. First, it confirms that the T2 antibody's signal, presumably representing *Zbtb48*, is localized in the germplasm. Second, the knockout of *zbtb48* does not affect the morphology or localization of the germplasm at the cleavage furrow. Lastly, the mutation does not displace the supposed *Zbtb48* signal from the germplasm.

This co-localization was also observed in the PGCs of both wild-type and *zbtb48*^{-/-} mutants at 1 dpf (Figure 45c). Notably, an additional *Zbtb48* signal outside the Ziwi foci was occasionally detected in both wild-type and mutant embryos (arrow), with this signal being more pronounced in the mutants. However, the exact cellular compartment where this additional staining occurs could not be determined.

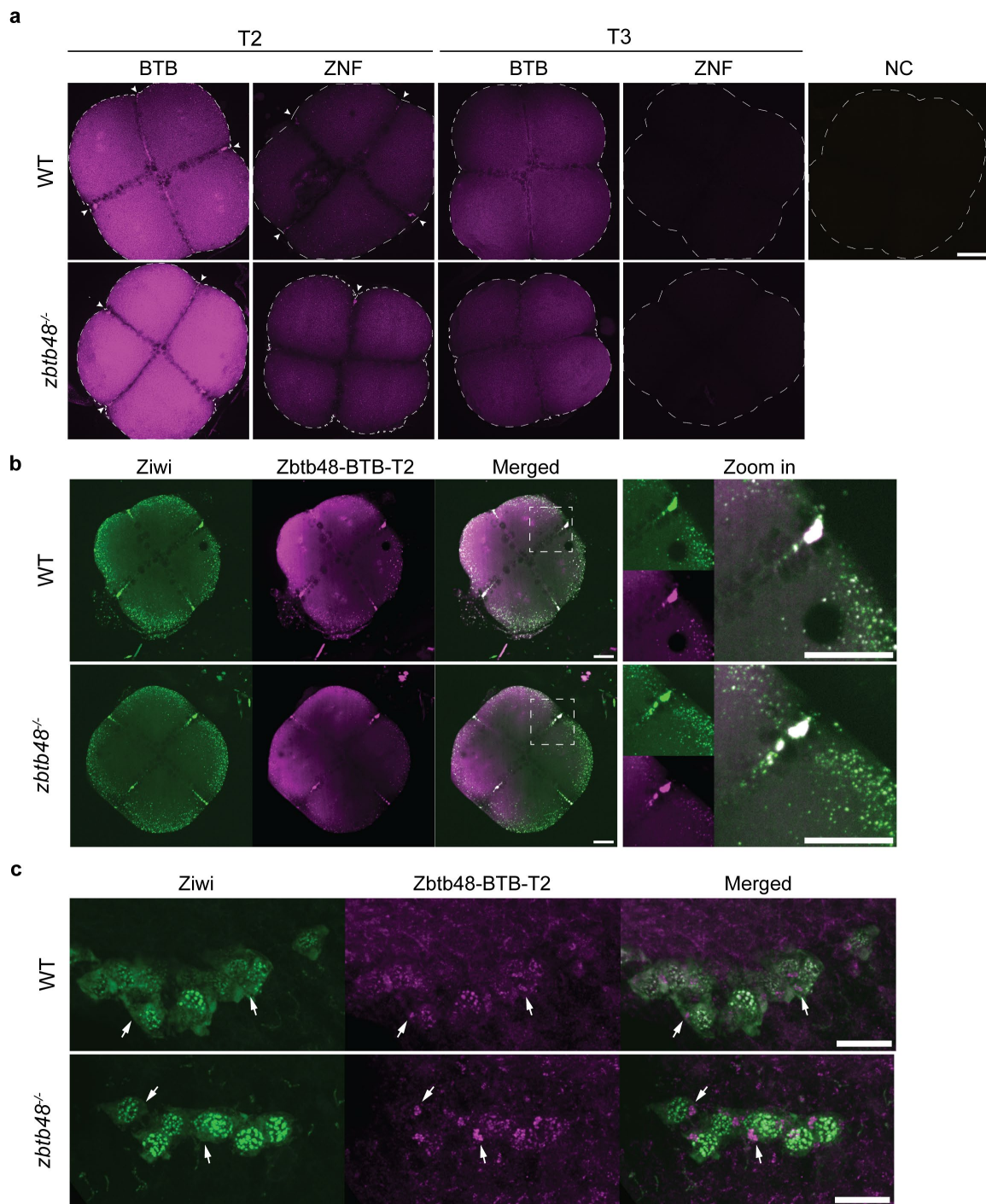


Figure 45. Validating Zbtb48 expression in germplasm and germ cells of embryos using Zbtb48 antibodies

a) Embryos at the 4-cell stage stained with various Zbtb48 antibodies in wild-type and *zbtb48*^{-/-} mutants. The negative control (NC) embryo was not treated with the primary antibody. The white dotted line outlines the embryo, and the white arrowhead indicates staining at the cleavage furrow. Same brightness and contrast settings were applied across all images. Scale bar = 100 μ m.

b) 4-cell stage wild-type and *zbtb48*^{-/-} embryos co-stained with the Zbtb48-BTB-T2 antibody (magenta) and Ziwi antibody (green), where the latter marks the germplasm. Individual and merged channel images are shown. The germplasm region, boxed by a white dotted outline, is magnified and displayed in the rightmost images. Scale bar = 100 μ m.

c) Close-up images of primordial germ cells (PGCs) in 1 dpf embryos co-stained with Zbtb48-BTB-T2 antibody (magenta) and Ziwi antibody (green) in wild-type and *zbtb48*^{-/-} mutants. The white arrow highlights the signal detected by the Zbtb48-BTB-T2 antibody but not with Ziwi antibody. Scale bar = 20 μ m.

Images were acquired using a 20X (a), 10X (b), and 60X (c) objective lens on a BC43 confocal microscope (Andor).

Given Zbtb48's roles as both a telomere binder and a transcription factor, its nuclear localization was examined. Staining was performed on 4 hpf embryos, when *zbtb48* mRNA levels were still high, as shown in Figure 4. The T2-BTB antibody revealed signals in both the nucleus (Figure 46ai) and the cytoplasm (Figure 46aaii). During mitosis, the signal did not co-localize with DAPI. Instead, most of the signal appeared localized in the cytoplasm, possibly at the spindle poles (Figure 46aaiii). Similar to the 4-cell stage, no signal was observed in the 4 hpf embryos probed with T3-BTB antibody.

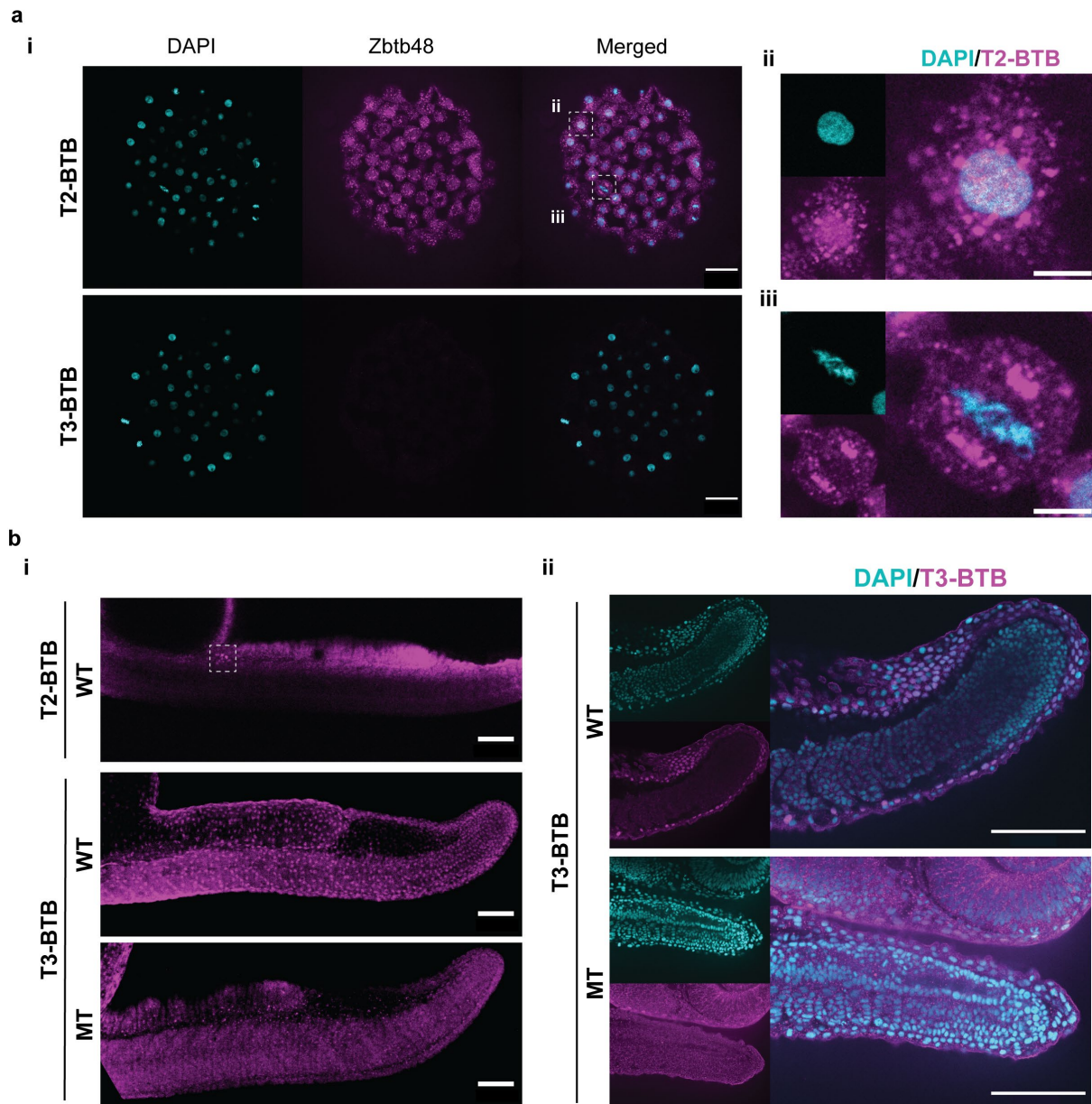


Figure 46. IF imaging of 4 hpf and 1 dpf embryos stained with Zbtb48 antibodies

a) 4-cell stage wild-type embryos stained with Zbtb48 antibodies. **i)** Single plane images displaying embryos stained with the respective Zbtb48 antibodies and DAPI. Scale bar = 50 μm. Cells outlined by a white dotted line are magnified and shown in the rightmost images in **ii)** at interphase and **iii)** at metaphase, with individual and merged channel images shown. Scale bar = 10 μm. **b)** 1 dpf embryos stained with Zbtb48 antibodies. Scale bar = 100 μm. **i)** Maximum projection images of the trunk and tail of 1 dpf embryos, with the PGCs boxed in white dotted line. **ii)** Single plane images of the tail region of wild-type and *zbtb48*^{-/-} mutant embryos stained with Zbtb48-T3-BTB antibody and DAPI.

Images were acquired using a 40X objective lens (a and bii) and a 10X objective lens (bii) on a BC43 confocal microscope (Andor).

When staining was repeated on 1 dpf embryos, a different pattern emerged. The T2-BTB antibody showed no nuclear or cytoplasmic signal but revealed prominent staining in the PGCs at the gonadal ridge (Figure 46bi, boxed), consistent with Figure 45c. In contrast, the wild-type embryos probed with T3-BTB antibody showed a distinct signal in what appears to be the epidermal layer, but such signal was absent in *zbtb48*^{-/-} mutants (Figure 46bi). Closer examination of the tail region (Figure 46bii) showed that the signal localized to the nucleus of epidermal cells. This epidermal staining is intriguing because the scRNA-seq data indicates low *zbtb48* mRNA levels in the epidermis like the other tissues at 1 dpf (Figure 44a). However, adult scRNA-seq data also indicate that the skin has the second highest *zbtb48* expression level after the testis (Figure 44b), which might explain the observed epidermal expression.

In summary, the T2-BTB antibody showed germplasm staining at the 4-cell stage and in PGCs at 1 dpf. This result aligned with the high *zbtb48* mRNA levels observed in germ cells, as indicated in Figure 44. However, the signal was not diminished in mutant samples, suggesting that the T2-BTB antibody might either recognize both wild-type and mutant Zbtb48 proteins, or that the observed staining in the germplasm and PGCs may not be specific to Zbtb48. The latter could also explain why such staining was not seen with the T3-BTB (M) antibody, which was the antibody previously used in IP-MS on PAC2 cells and detected Zbtb48 (Figure 37d).

Conversely, the T3-BTB antibody exhibited nuclear staining in the epidermal layer of 1 dpf embryos, which the loss of signal in *zbtb48*^{-/-} mutants supports the antibody's specificity. However, the absence of staining in the germplasm and PGCs with the T3-BTB antibody remains puzzling, given the strong evidence supporting *zbtb48* expression in germ cells.

The conflicting results between the IF staining obtained with the two different antibodies and the scRNA-seq data further underscore the need for validation to resolve issues with antibody specificity and effectiveness for IF applications. Although attempts were made to perform staining on 5 dpf larvae, the complexity of the organism and high background signal in other tissues made imaging and interpreting of results at this stage challenging.

Western blot analysis on gonads

To evaluate the suitability of the antibody for Western blot analysis, a blot was performed using protein lysates from adult gonads, as these tissues show the highest *zbtb48* expression among adult organs. Testes from both wild-type and *zbtb48*^{-/-} mutants, collected at one year of age, were probed with BTB-T2 and BTB-T3 antibodies. However, the overall quality of the blot was poor, making it difficult to assess the loss of bands in the mutants. Inconsistent amounts of proteins were observed across the wild-type and mutant lysates, despite the calculated concentration loaded were equal.

A strong band was detected between the 53 and 70 kDa markers with the BTB-T2 antibody (Figure 47ai), which corresponded to a band previously observed in Figure 37ai. However, this band was also present in the *zbtb48*^{-/-} mutant lysates. When the blot was probed with BTB-T3 antibodies, very few bands were observed in the wild-type samples (Figure 47aai). Re-probing with BTB-T3 (M) antibody revealed multiple bands, but no clear differences in banding patterns

between wild-type and *zbtb48*^{-/-} mutant testes were observed (Figure 47aⁱⁱⁱ). When the analysis was extended to the ovary of a nine-month-old wild-type, a large smear appeared on the blot (Figure 47aⁱⁱⁱ), likely due to the large yolk content. Unfortunately, these Western blot results were inconclusive and did not provide any clarity on the effectiveness of the antibodies against Zbtb48. This lack of specificity could be attributed to the antibodies being polyclonal, which may have resulted in the detection of non-Zbtb48-specific bands. Additionally, Zbtb48 may be a low-abundance protein, making its detection through Western blotting inherently difficult.

qRT-PCR analysis on gonads

Next, to determine whether the dysregulated genes identified in the 5 dpf omics analysis were also extended to the gonads, qRT-PCR was performed on gonad samples. To facilitate identification at early developmental stages, the *zbtb48*^{-/-} mutant line was crossed with a transgenic *vasa-egfp* line (*Tg(vasa-egfp)*). (*Tg(vasa-egfp); zbtb48*^{+/-}) mutant parents were then in-crossed to produce *zbtb48*^{-/-} mutant offspring, with wild-type siblings serving as controls.

Since scRNA-seq data indicate that *zbtb48* is highly expressed in germ stem cells and meiotic cells (Figure 44c and d), the gonads were harvested at earlier developmental stages to maximize the ratio of germ stem cells and meiotic cells to total cells. Zebrafish gonads begin differentiating after 25 dpf, with ovaries completing differentiation by approximately 35 dpf. In contrast, testes differentiation is typically completed by 45 dpf, as it requires an additional process of oocyte removal before differentiating^{160,161}.

To assess whether dysregulation could already be detected in differentiating gonads, samples were collected at 33 dpf (WT n = 13, *zbtb48*^{-/-} n = 3) and from differentiating testes or non-ovary gonads at 40 dpf (n = 4 per genotype). A significant 2.4-fold downregulation of *mtfp1* was observed in *zbtb48*^{-/-} mutants at both 33 dpf and 40 dpf (p = 0.007 and 0.05, respectively) (Figure 47bⁱ and Figure 47bⁱⁱ). Conversely, *akr7a3* showed a significant 5.4-fold downregulation in the 33 dpf differentiating gonads (p = 0.0013), but the stronger 11.8-fold downregulation at 40 dpf was not statistically significant (p = 0.08). Other genes, such as *atox1*, *hdac10*, *zgc:153284*, and *col28a2a*, showed no statistically significant differences between wild-type and mutant gonads at either 33 dpf or 40 dpf.

To investigate whether similar results would be observed in mature testes, where highest *zbtb48* expression was observed in Figure 44b, qRT-PCR was repeated on testes harvested from 1-year-old males (n = 4 per genotype). Interestingly, more genes exhibited significant differences in the testes of 1-year-old males compared to the differentiating gonads at 33 dpf and 40 dpf. In *zbtb48*^{-/-} mutants, a more than 2-fold downregulation of *mtfp1*, *akr7a3*, and *atox1* (p = 0.04, 0.001, and 0.04, respectively), along with an 8.5-fold upregulation of *zgc:153284* (p = 0.003), was observed (Figure 47bⁱⁱⁱ). However, levels of *hdac10* and *col28a2a* remained unchanged.

Unfortunately, due to an insufficient number of female fish, qRT-PCR was not performed on ovaries from 40 dpf or older samples.

Overall, the qRT-PCR results consistently showed significant downregulation of *mtfp1* in all tested gonads. While downregulation of *akr7a3* and *atox1*, as well as upregulation of *zgc:153284*, was observed, these changes were not consistent across all samples. So far, *hdac10* and *col28a2a* have not shown significant changes in the tested gonads.

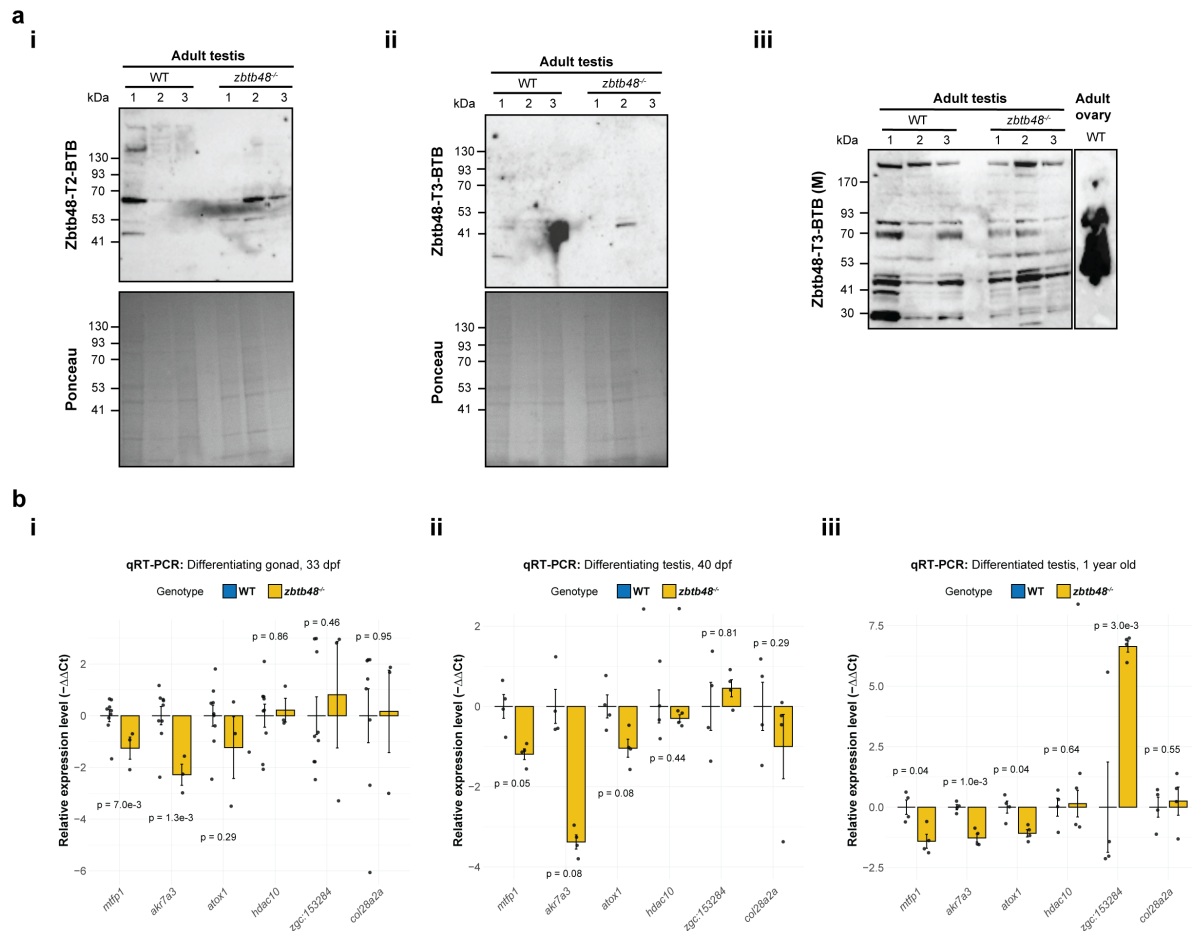


Figure 47. Western blot and qRT-PCR analysis of gonads from *zbtb48*^{-/-} mutants

a) Western blot analysis was performed on 1-year-old wild-type and *zbtb48*^{-/-} mutant testes using three antibodies: **i)** T2-BTB, **ii)** T3-BTB, and **iii)** T3-BTB (M), where the blot from **(ii)** was re-probed with the T3-BTB (M) antibody. Additionally, 9-month-old wild-type ovaries were included in the analysis, with Ponceau staining serving as a loading control.

b) qRT-PCR analysis was conducted on gonads to evaluate the expression of genes previously found to be dysregulated in 5 dpf samples. This was performed on RNA extracted from **i)** 33 dpf undifferentiated gonads (WT n=12, MT n=3), **ii)** 40 dpf testes (n=4 per genotype), and **iii)** 1-year-old testes (n=4 per genotype). Error bars represent the standard error of the mean (SEM), and p-values were calculated using two-tailed Welch's t test.

Proteome analysis on gonads

Proceeding from the promising qRT-PCR results, proteome analysis was also conducted on gonad samples. Differentiated ovaries (n=3 per genotype) and differentiating testes (n=4 per genotype) were collected from those 40 dpf siblings used in qRT-PCR. The proteome analysis of ovaries identified 850 downregulated proteins and 38 upregulated proteins out of 6,015 quantified (Figure 48ai). Conversely, in the testes proteome analysis with 5,891 quantified proteins using the same threshold (fold change >|2|, p-value < 0.05), it has much lower number

of protein that were downregulated, with 30 proteins downregulated and 22 upregulated in *zbtb48*^{-/-} mutants (Figure 48aii). Notably, there was a stark difference in the number of downregulated proteins between the differentiated ovaries and differentiating testes in 40 dpf zebrafish.

To investigate whether the fully differentiated, mature testes exhibit greater downregulation compared to 40 dpf differentiating testes, proteome analysis was also conducted on testes harvested from 10.5-month-old second-generation *zbtb48*^{-/-} mutants and its wild-type counterpart. Interestingly, the analysis revealed an increase in the number of downregulated proteins, with the downregulation of 227 proteins while 9 proteins were upregulated out of the 5,419 quantified (Figure 48aiii). This suggests that gonad maturity increases the number of downregulated proteins, potentially correlating with the expression level of Zbtb48.

Among the three types of gonads examined, Zbtb48 was only detected in the proteome of wild-type 40 dpf ovaries, further highlighting its low expression levels and explaining the difficulty in detecting it via Western blot (Figure 47a). It is intriguing that, despite the ovary and differentiating testes being from siblings of the same age, Zbtb48 was only detectable in the differentiated ovary, suggesting higher Zbtb48 expression in fully differentiated gonads. However, this hypothesis would require additional analysis of recently differentiated testes, such as those from 45 dpf samples.

A comparative analysis of the three gonad proteomes revealed a distinct set of commonly dysregulated proteins than those identified in the 5 dpf proteome analysis (Figure 48b). The functions of these proteins are detailed in Table 3. Notably, Mtfp1 showed consistent and significant downregulation across all gonad proteomes, aligning with previous findings from the 5 dpf proteome analysis (Figure 42a) and further supported by qRT-PCR results (Figure 47b). A deeper analysis of Mtfp1 peptides identified by mass spectrometry revealed a drastic reduction of Mtfp1 in the *zbtb48*^{-/-} mutants, particularly in the testes of 10.5-month-old samples, where Mtfp1 was not detected in mutants but highly expressed in wild-type samples (Figure 48c). In addition to Mtfp1, a chitinase, Chia.6, was significantly downregulated across all gonad proteomes.

Interestingly, the proteome of 10.5-month-old testes shared more commonly downregulated proteins with 40 dpf differentiated ovaries than with 40 dpf differentiating testes. This may be attributed to the fewer downregulated proteins detected in the 40 dpf testes, indicating distinct effects of the *zbtb48* knockout between differentiating and differentiated gonads, regardless of age or sex.

In total, 15 other proteins, in addition to Mtfp1 and Chia.6, were commonly downregulated in both 40 dpf ovaries and 10.5-month-old testes. This list includes Gbp3, previously identified as downregulated in 5 dpf *zbtb48*^{-/-} larvae. Two additional mitochondrial proteins (Mipep and Mrpl45) were also downregulated, although their relationship remains unclear. The remaining commonly downregulated proteins include Dhx57, Gtf2h4, Hira, Mdn1, Nbas, Parp14rs1, Parp9, Slc29a1a, Tdrd12, Zgc:161973, Zmym4, and Zpr1.

Tdrd12, a protein crucial for gametogenesis, is noteworthy. Deficiency in Tdrd12 is known to cause infertility in males¹⁶², yet in *zbtb48*^{-/-} mutants, its downregulation was not severe enough to induce infertility, as both male and female *zbtb48*^{-/-} fish were fertile, even up to second generation (Fish ID: 872) at 1.5 years of age.

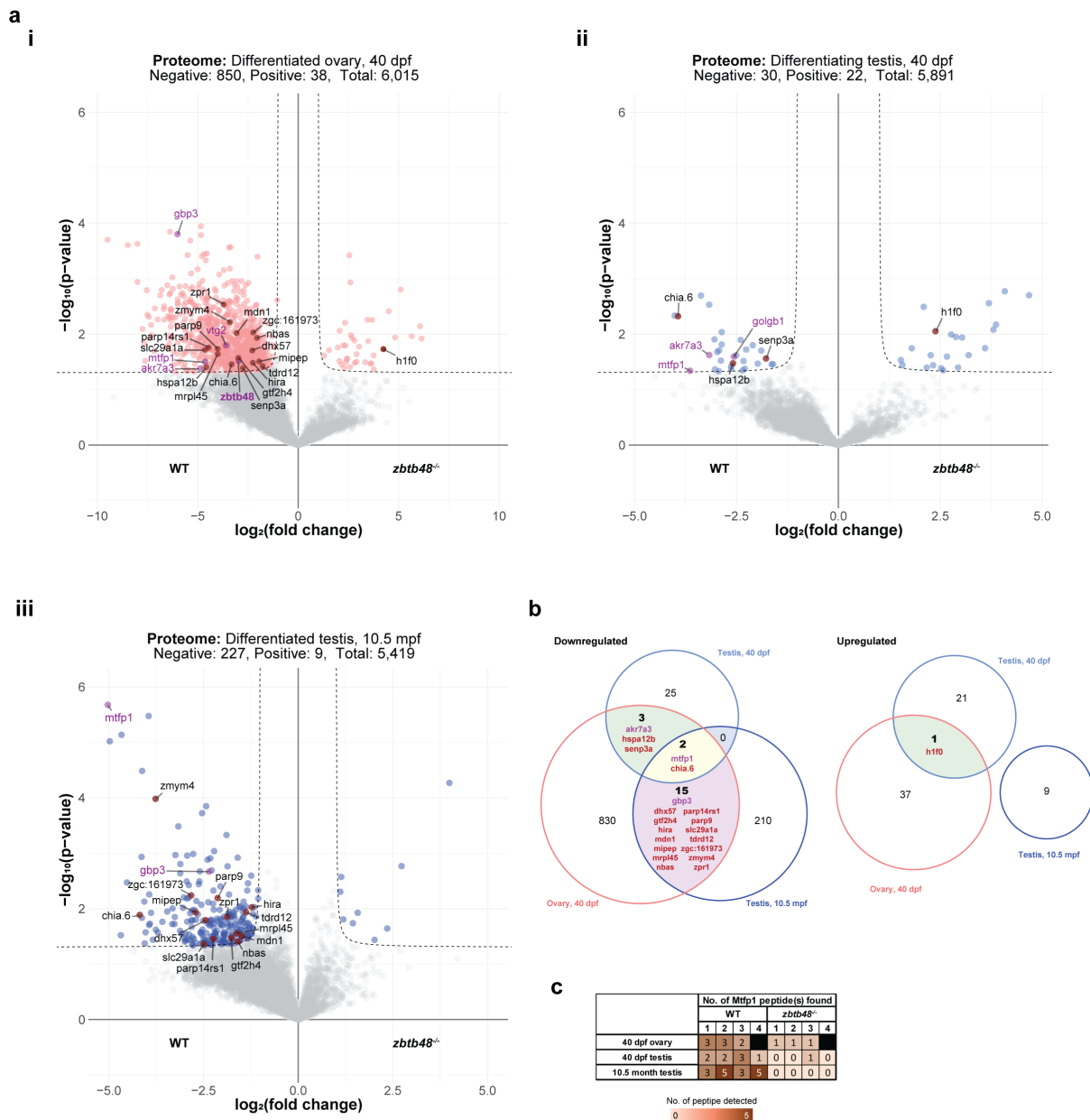


Figure 48. Proteome analysis of gonads from *zbtb48*^{-/-} mutants

a) Volcano plots illustrate the results for **i)** 40 dpf differentiated ovaries, **ii)** 40 dpf differentiating testes, and **iii)** 10.5-month-old differentiated testes. Biological replicates were used for each genotype, with 40 dpf ovaries (n=3) and both 40 dpf and 10.5-month-old testes (n=4). The results were log-transformed and plotted on the x-axis as $\log_2(\text{fold change})$ and on the y-axis as $-\log_{10}(\text{p-value})$ (Welch's t-test). The protein enrichment threshold was set at a fold change $> |2|$ and a p-value < 0.05 . Proteins identified in Figure 42 are highlighted in purple, while commonly dysregulated proteins across gonads are marked in red.

b) displays Venn diagrams representing downregulated (left) and upregulated (right) proteins in different gonadal tissues of *zbtb48*^{-/-} mutants compared to wild-type.

c) Table showing the number of Mtfp1 peptides identified in the proteome analysis of gonads using MS.

Conversely, only three proteins were commonly downregulated between 40 dpf ovaries and testes: *Ark7a3*, *Hspa12b*, and *Senp3a*. *Ark7a3* had been previously identified as downregulated in 5 dpf *zbtb48*^{-/-} larvae, though its transcript downregulation in 40 dpf sibling testes was not significant according to qRT-PCR. Despite the high number of downregulated proteins, only one protein—H1.0 linker histone (H1f0)—was commonly upregulated in both 40 dpf ovaries and testes. Additionally, *Golgb1*, which was downregulated in the 5 dpf *zbtb48*^{-/-} larvae, was only found to be downregulated in the 40 dpf testes but not in the ovaries or 10.5-month-old testes.

Comparing the proteomic results from adult testes with the qRT-PCR data, the significant downregulation of *ark7a3* and *atox1*, as well as the upregulation of *zgc:153284* transcripts in qRT-PCR of one-year-old testes, were not reflected in the proteome of the 10.5-month-old testes.

Overall, the proteome analysis of gonads has yielded intriguing results. The differentiated ovaries at 40 dpf surprisingly showed protein dysregulation in the *zbtb48*^{-/-} mutant that was more similar to the 10.5-month-old testes than to the 40 dpf differentiating testes of its sibling. This suggests that gonad maturity may influence *Zbtb48* expression levels, which in turn could affect the transcriptional differentiation of certain genes. Although the precise role of *Zbtb48* remains unclear in this study, the proteome analysis across various gonads offers valuable insights into its potential function as a transcriptional regulator. Additionally, it highlights *Mtfp1* as an evolutionarily conserved target of *Zbtb48*, as *Mtfp1* was consistently downregulated at both the transcript and protein levels in the *zbtb48*^{-/-} mutant.

Table 3. Summary of dysregulated proteins identified in gonads of *zbtb48*^{-/-} mutant fish

	Gene name	Full gene name	Zin description
Up	<i>h1f0</i>	H1.0 linker histone	Predicted to have double-stranded DNA binding activity and nucleosomal DNA binding activity. Predicted to be involved in DNA packaging and regulation of nucleobase-containing compound metabolic process. Predicted to localize to nucleus. Orthologous to human H1-0 (H1.0 linker histone).
	<i>akr7a3</i>	aldo-keto reductase family 7, member A3 (afatoxin aldehyde reductase)	Is expressed in gut, liver, myotome, and optic vesicle. Orthologous to several human genes including AKR7A3 (aldo-keto reductase family 7 member A3).
Downregulated	<i>chia.6</i>	chitinase, acidic.6	Predicted to have chitin binding activity and chitinase activity. Predicted to be involved in chitin catabolic process. Predicted to localize to extracellular region.
	<i>dhx57</i>	DEAH (Asp-Glu-Ala-Asp/His) box polypeptide 57	Predicted to have RNA binding activity. Predicted to localize to intracellular. Orthologous to human DHX57 (DEAH-box helicase 57).
	<i>gbp3</i>	guanylate binding protein 3	Predicted to have GTP binding activity and GTPase activity. Predicted to be involved in inflammatory response, innate immune response, and regulation of apoptotic process.
	<i>gtf2h4</i>	general transcription factor IIH, polypeptide 4	Predicted to have double-stranded DNA binding activity. Predicted to be involved in nucleotide-excision repair and phosphorylation of RNA polymerase II C-terminal domain. Predicted to localize to transcription factor TFIID core complex and transcription factor TFIID holo complex. Orthologous to human GTF2H4 (general transcription factor IIH subunit 4).
	<i>hira</i>	histone cell cycle regulator a	Predicted to contribute to DNA binding activity and nucleosome binding activity. Predicted to be involved in DNA replication-independent nucleosome assembly, mitotic sister chromatid segregation, and regulation of chromatin silencing. Predicted to localize to HIR complex and nuclear chromatin. Orthologous to human HIRA (histone cell cycle regulator).
	<i>hspa12b</i>	heat shock protein 12B	Involved in sprouting angiogenesis. Is expressed in several structures, including brain; cardiovascular system; cranium; somite 1; and ventral mesoderm. Orthologous to human HSPA12B (heat shock protein family A (Hsp70) member 12B).
	<i>mdn1</i>	midasin AAA ATPase 1	Predicted to have ATP binding activity and ATPase activity. Predicted to be involved in rRNA processing and ribosomal large subunit assembly. Predicted to localize to pre-ribosome, large subunit precursor. Orthologous to human MDN1 (midasin AAA ATPase 1).
	<i>mipep</i>	mitochondrial intermediate peptidase	Predicted to have metalloendopeptidase activity. Predicted to be involved in peptide metabolic process and protein processing involved in protein targeting to mitochondrion. Predicted to localize to mitochondrion. Human ortholog(s) of this gene implicated in combined oxidative phosphorylation deficiency 31. Orthologous to human MIPEP (mitochondrial intermediate peptidase).
	<i>mrpl45</i>	mitochondrial ribosomal protein L45	Predicted to localize to mitochondrion. Orthologous to human MRPL45 (mitochondrial ribosomal protein L45).
	<i>mtfp1</i>	mitochondrial fission process 1	Predicted to be involved in mitochondrial fission. Predicted to localize to mitochondrial inner membrane. Is expressed in visceral fat. Orthologous to human MTFP1 (mitochondrial fission process 1).
	<i>ribas</i>	NBAS subunit of NRZ tethering complex	Involved in embryonic viscerocranium morphogenesis; negative regulation of nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; and nuclear-transcribed mRNA catabolic process, nonsense-mediated decay. Predicted to localize to endoplasmic reticulum. Orthologous to human NBAS (NBAS subunit of NRZ tethering complex).
	<i>parp14rs1</i>	poly(ADP-ribose) polymerase family member 14-related sequence 1	Predicted to have NAD+ ADP-ribosyltransferase activity; protein ADP-ribosylase activity; and transcription corepressor activity. Predicted to be involved in negative regulation of gene expression; protein mono-ADP-ribosylation; and protein poly-ADP-ribosylation. Predicted to localize to cytoplasm and nucleus. Orthologous to several human genes including PARP14 (poly(ADP-ribose) polymerase family member 14).
	<i>parp9</i>	poly(ADP-ribose) polymerase family member 9	Predicted to have transcription corepressor activity; transferase activity, transferring pentosyl groups; and ubiquitin-like protein ligase binding activity. Predicted to be involved in negative regulation of gene expression; positive regulation of interferon-gamma-mediated signaling pathway; and protein ADP-ribosylation. Predicted to localize to cytoplasm and nucleus. Orthologous to human PARP9 (poly(ADP-ribose) polymerase family member 9).
	<i>senp3a</i>	SUMO specific peptidase 3a	Predicted to have SUMO-specific protease activity. Predicted to be involved in protein desumoylation. Predicted to localize to nucleus. Orthologous to human SENP3 (SUMO specific peptidase 3).
	<i>slc29a1a</i>	solute carrier family 29 member 1a	Predicted to have nucleoside transmembrane transporter activity. Predicted to be involved in uridine transport. Predicted to localize to integral component of membrane and plasma membrane. Is expressed in several structures, including central nervous system; eye; immature eye; neural plate; and pectoral fin. Orthologous to human SLC29A1 (solute carrier family 29 member 1 (Augustine blood group)).
	<i>tldr12</i>	tudor domain containing 12	Predicted to have ATP binding activity and nucleic acid binding activity. Involved in germ cell development. Predicted to localize to nucleus. Is expressed in gonad; ovary; and testis. Orthologous to human TDRD12 (tudor domain containing 12).
	<i>zgc:161973</i>		Orthologous to human FBXO15 (F-box protein 15).
	<i>zmym4</i>	zinc finger MYM-type containing 4	Predicted to have sequence-specific DNA binding activity and zinc ion binding activity. Predicted to be involved in cell cycle. Predicted to localize to nucleus. Orthologous to human ZMYM4 (zinc finger MYM-type containing 4).
	<i>zpr1</i>	ZPR1 zinc finger	Predicted to have zinc ion binding activity. Predicted to localize to cytoplasm and nucleus. Orthologous to human ZPR1 (ZPR1 zinc finger).

Discussion

Overview

This study has successfully established CRISPR-Cas9 knockout zebrafish lines for *hmbox1a*, *hmbox1b*, and *zbtb48*, with all the deletions resulting in frameshift mutations. For Hmbox1a/b study, two CRISPR knockout lines were generated for each gene, with most of the experiments performed on *hmbox1a* and *hmbox1b* mutants with larger deletions, as single mutants (*hmbox1a*^{LD/LD} and *hmbox1b*^{LD/LD}) and/or (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutant. On the other hand, only one knockout line has been established for Zbtb48 study.

Previously, HMBOX1 was reported as a positive regulator of telomere length in mammalian cells⁷⁷, whereas ZBTB48 was reported as a negative regulator^{73,74}. This study has investigated whether this role is also conserved in the zebrafish ortholog. However, no significant difference in telomere length was found between the first generation homozygous (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutants and *zbtb48*^{-/-} mutants when compared to their respective wild-type counterparts. One of the plausible explanations may be that it requires several generations before the telomere length change became evident in the mutants. Furthermore, telomere length was primarily measured in adult fins or embryo tails, rather than in tissues known for high telomere elongation activity or where high expression of Hmbox1a, Hmbox1b, and Zbtb48 has been reported.

This study also did not find any difference in the lifespan of *hmbox1a*^{LD/LD}, *hmbox1b*^{LD/LD} and (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) mutants from their wild-type counterpart in the first three inbred generation. Conversely, the first generation *zbtb48*^{-/-} mutants preliminarily showed a better survival rate than their wild-type counterparts in the first year of age. Since this result was based on two crosses, with one terminating at one year, further investigation with a larger sample size and more crosses would be required to draw a definitive conclusion.

All homozygous mutants also appeared healthy and did not exhibit any visible physical difference from wild-type fish. Although unusual swimming behavior was observed in the third generation of inbred (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutants, such abnormality was not observed in subsequent breeding attempt. Thus, suggesting that it may have resulted from environmental factors rather than the knockouts.

Since no physical differences were observed at the organismal level, nor any changes in telomere length, attempts were made to study these mutants at the cellular level, focusing on cellular oxidative stress, senescence and telomere dysfunction-induced foci (TIF). However, these studies were hindered by technical challenges, as the protocol required further optimization, and the suitable reagents for zebrafish applications need to be explored.

Overall, this study did not provide any evidence to clarify their role in telomere biology in zebrafish. It is very likely that these proteins are not as critical for fish survival as compared to other telomere-binding proteins like Terfa, Pot1, and telomerase^{122,124,131,132}.

In addition to the CRISPR knockout lines, this study has also generated antibodies to gain further insights into these proteins. In total, two Hmbox1a, one Hmbox1b, and four Zbtb48 antibodies have been purified and validated. A series of experiments have revealed that Hmbox1a and Hmbox1b antibodies were more effective in IF applications in embryos than in Western blotting or IP. The Hmbox1a-T1 and Hmbox1b-T2 antibodies demonstrated specificity for Hmbox1a and Hmbox1b, respectively, while the Hmbox1a-T3 antibody appeared to recognize both paralogues, serving as a pan-Hmbox1a/b antibody.

In contrast, the Zbtb48 antibodies showed suboptimal performance across all tested applications, leading to inconclusive results. While these antibodies could detect endogenous Zbtb48 in embryo protein lysates during IP applications, it also has a false-positive detection in *zbtb48*^{-/-} mutants. All the Zbtb48 peptides detected in *zbtb48*^{-/-} mutants have sequences that correspond to the epitope used for antibody generation and purification. Additionally, these epitope sequences were detected along with the MBP tag. Therefore, it is likely that the purified Zbtb48 antibody still contains the epitope peptide from the purification and resulted in a false positive detection of Zbtb48 in the mass spectrometry analysis. This issue was also observed with Hmbox1a/b antibodies, which raises concerns and cautions should be taken when interpreting such IP results.

In addition to the challenges faced in IP, detecting Hmbox1a, Hmbox1b, and Zbtb48 via Western blot has been difficult due to nonspecific bands from the polyclonal antibodies. While not ruling out that the antibodies generated may be sub-optimal, this issue seems to be further exacerbated by the low abundance of these proteins in the sampled tissues. None of these proteins were detected in the proteomic analyses conducted on 4 hpf and 5 dpf embryos, despite the detection of Hmbox1a expression at 4 hpf via IF and Zbtb48 at 5 dpf using telomere pull-down. Furthermore, even with the enrichment methods deployed like telomere pull-down, the Hmbox1a/b paralogues were not detected, and while Zbtb48 was detected, it was often below the enrichment threshold when compared to the control probe in the embryos. These results highlight the low abundance of these proteins in the whole embryo, making it difficult to detect them on Western blot using the antibodies raised.

According to the IF results, Hmbox1a was expressed during the blastula and gastrula stages (3–10 hpf), while Hmbox1b was expressed at 10 hpf to 3 dpf, but confined to the notochord and pronephros, which represent only a small population in the whole embryo. Therefore, this might explain the absence of Hmbox1a/b paralogues in telomere pull-downs. However, the telomere pull-down was not repeated on embryos of earlier stages due to technical challenges. Embryos at blastula stage have a larger yolk-to-cell ratio and removing yolk proteins would be crucial to avoid competition with proteins of interest during the pull-down process¹³⁶. However, to account for the limited number of cells in each embryo, and the inevitable loss of samples during de-yolking and washing steps, a larger sample size will be required to obtain sufficient protein concentrations for pull-downs. As a result, such pull-down was not performed on the embryos at blastula stage. On the other hand, as Hmbox1b was expressed only in a small group of cells, its protein level may have been diluted in the whole-embryo protein extract. Nevertheless, the IF experiments have successfully detected the expression of Hmbox1a/b paralogues in the embryos, revealing that they exhibit transient expression patterns. These results further underscore the importance of selecting samples at the appropriate developmental stage for analysis.

The telomere pull-down experiment in BRF41 has shown significant enrichment for Hmbox1b, Zbtb48, and all six shelterin components. However, in contrast to the telomere pull-down experiments in the embryos, most of these proteins were not detected in the embryo. These discrepancies could be attributed to various biological and/or technical factors.

Firstly, tissue differences may play a role. As BRF41 cell line was derived from fin fibroblasts, it likely exhibits a distinct protein expression profile compared to whole embryos, which contain multiple cell types. Secondly, protein extraction methods are generally more optimized and simpler for cell lines than for embryos. Processing embryos may require the incorporation of different additional methods depending on the developmental stages. Early stages typically need de-yolking, while later stages often require additional enzymatic digestion or more intensive lysis techniques, such as sonication, to break down the skin barrier and extracellular matrix. Having an optimal lysis method is crucial, as suboptimal processing can lead to low protein yields and biased results, while excessive sonication may alter protein conformation, potentially compromising protein integrity and affecting telomere-binding activity^{163,164}. Thirdly, proteins with very high affinity for telomeres may remain bound to DNA during cell lysis, causing them to pellet and be discarded, which could explain their absence from the lysate used in the pull-down assay. Alternatively, lower affinity binders may not have been pulled down due to competition. Lastly, low abundance proteins may not be detected if they fall below the detection limit of mass spectrometer.

Overall, although this study did not uncover the telomeric roles of these proteins in zebrafish, it has gained a lot of insight on the spatiotemporal expression patterns of Hmbox1a/b paralogues and Zbtb48 using IF staining and existing single-cell RNA sequencing (sc-RNA-seq) data. In the following discussion, I will delve deeper into their potential roles in various tissues based on the expression patterns of these proteins.

Hmbox1a/b paralogues

Through IF staining, Hmbox1a was found expressed in the nuclei of embryonic cells as early as 1k-cell stage (3 hpf) and persisting at least until the bud stage (10 hpf). Since Hmbox1a is maternally deposited and not expressed by the zygotes after the zygotic genome activation, its expression likely began in the oocytes.

It was speculated that HMBOX1 elongates telomere through facilitating telomerase recruitment⁷⁷. This was based on a few pieces of evidence like co-localization of HMBOX1 signal with telomere foci, localization of HMBOX1 signal to Cajal bodies, and enrichment of telomerase active component using IP. Like the previous report, Hmbox1a was found co-localizing with telomere foci in 4 hpf embryos. This frequency was higher than its ortholog in human HeLa cells, but lower than in mouse spermatocytes⁷⁷, which the latter has a higher telomerase activity⁸⁸. Unfortunately, IF staining of the Cajal bodies was not performed in this study, hence, it is unclear if Hmbox1a also localized to it to participate in telomerase activity in zebrafish. Furthermore, the IP experiments performed on the 1 to 3 dpf embryos and the PAC2 cell line lysates failed to enrich for Hmbox1a, and no IP was not performed on the tissues where Hmbox1a expressed, such as oocytes or early embryonic stages. Although some of these IPs enriched for Hmbox1b, they did not enrich for the active component of the telomerase.

Despite lacking these pieces of evidence to support Hmbox1a's role as a telomerase recruiter, this study found that the expression level of Hmbox1a in the nucleus appears to correlate with changes in telomerase activity level throughout the cell cycle, which was not reported in the previous study. Hmbox1a expression was lowest during mitosis, which coincides to when telomerase activity is also at its lowest in the cell cycle¹⁶⁵. Based on the co-localization of the Hmbox1a signal with telomere foci and its expression level coinciding with telomerase activity, it is likely that Hmbox1a may also help recruit telomerase in zebrafish. However, further investigations at molecular level are needed to strengthen this hypothesis.

The expression of Hmbox1b begins later than Hmbox1a. However, the absence of Hmbox1a during early embryonic development does not appear to influence the spatiotemporal expression of Hmbox1b. In *hmbox1a^{LD/LD}* mutants, Hmbox1b expression remains detectable like in the wild-type embryos. Hmbox1b expression begins in the notochord at around 90% epiboly (9 hpf) when it forms and diminishes after 2 dpf. Additionally, Hmbox1b was expressed in the pronephros shortly after forming at the 14-somite stage (16 hpf) and fades by 3 dpf. Hmbox1b expression was observed in the nucleus of both the notochord and pronephros, which are both derived from the mesoderm layer. Notably, Hmbox1b expression in the pronephros aligns with microarray data showing that the human HMBOX1 ortholog is highly expressed in human kidney tubules among various tissues analyzed⁸¹.

Generally, most of the genes from the homeobox family play a critical role in orchestrating embryonic development, which may explain the early expression of Hmbox1a/b paralogues. To speculate on the role of Hmbox1b, its expression pattern was compared to other homeobox genes during early embryo development. Hmbox1b expression pattern appears to be different from the well-characterized *hox* genes, which exhibit transient and regionally restricted expression to coordinate body segment identity¹⁶⁶. On the other hand, although Hmbox1b also shows transient expression but spans the entire notochord in the somite region.

The notochord is a transient rod-like structure that extends along the embryo's anterior-posterior axis, serving as a signaling center to guide surrounding tissue development in chordates¹⁶⁷. *noto* (also known as *floating head*, *flh*), another homeobox gene, is important for proper formation of notochord¹⁶⁸. Similar to *noto*, Hmbox1b is expressed throughout the notochord in somite segments but is absent in rhombomeres (brain regions)^{168,169}. Although Hmbox1b was expressed during notochord formation, *hmbox1b^{LD/LD}* mutants showed no apparent defects, unlike *noto* mutations, which result in notochord loss¹⁶⁸, or other mutations that result in notochord deformities^{170,171}. Similarly, adult *hmbox1b^{LD/LD}* mutants displayed no signs of spine deformity, despite the notochord transforming into the spine during later development^{172,173}. Therefore, it is unlikely that Hmbox1b plays an important role in notochord or spine formation, and its function in the notochord remains unresolved in this study.

It is plausible that Hmbox1b functions as a signaling molecule in the notochord, potentially associated with the pronephros. While the specific signals it mediates were unclear, Hmbox1b expression spans the entire pronephros (excluding the podocytes) and aligned with the timing of pronephros extension and segmentation. Although the HMBOX1 ortholog has been shown to aid stem cell differentiation into vascular endothelial cells in rodents^{92,93}, it appears unlikely that it influences stem cell differentiation in the pronephros in zebrafish. The *hmbox1b^{LD/LD}* mutant embryos showed no apparent defect on pronephros morphology or positioning, as

visualized by the RNA-FISH staining of pronephros marker *cdh17*. Therefore, it seems unlikely that Hmbox1b plays a critical role in pronephros differentiation, elongation, or positioning along the body axis.

The localization of Hmbox1b expression in the pronephros closely mirrors *cdh17*, a pan-pronephros marker, which also excludes the podocytes. However, the *cdh17* marker alone does not provide information about nephron segmentation. The zebrafish pronephros is divided into eight distinct segments and each segment being characterized by different solute carriers, which are responsible for nutrient, electrolyte, and water reabsorption¹⁷⁴. Differential gene expression in the pronephros begins as early as the 8-somite stage (13 hpf), with all segments formed by 24 hpf and some continuing to elongate until 48 hpf. This segmentation is regulated by various signaling factors such as *hnf1b*, retinoic acid, and Notch^{141,175–180}.

Intriguingly, the *hnf1b* paralogues, which are among the segmentation regulators, also belong to the HNF class of the homeobox gene family, like *hmbox1a/b*. Given that they are from the same class of genes, this raised concerns that the Hmbox1b antibody generated might be recognizing Hnf1b and staining the pronephros. However, further investigation of the amino acid sequence identity revealed that these Hnf1b paralogues share less than 30% identity with Hmbox1b. Furthermore, according to the previously published RNA *in situ* hybridization of *hnf1ba* and *hnf1bb*¹⁷⁶, it showed that *hnf1ba* is expressed throughout the pronephros except for the podocytes and cloaca, while *hnf1bb* is mainly expressed in the proximal part of the pronephros. In contrast, Hmbox1b expression spans the entire pronephros that except for the podocytes. The Hmbox1b antibody that also showed notochord staining, which neither of the *hnf1b* paralogue expression was found in the notochord¹⁷⁶. Lastly, the absence of Hmbox1b signal in the immunofluorescence of *hmbox1b^{LD/LD}* mutants further confirms that the generated Hmbox1b antibody was recognizing Hmbox1b in the pronephros and not Hnf1b. These results not only validate Hmbox1b expression in the pronephros but also suggest a distinct role for Hmbox1b compared to Hnf1b in pronephros development.

To assess Hmbox1b's role in pronephros segmentation, attempts were made to perform RNA *in situ* hybridization for segment labelling. However, these efforts were hindered by technical challenges. Despite this, the uniform expression of Hmbox1b, which lacks the gradient or restricted patterns typical of segmentation factors, suggests it is less likely to play a significant role in segmentation. While nephron tubule segmentation and reabsorption remain unexamined, glomerular filtration by podocytes appears unaffected in (*hmbox1a^{LD/LD}*; *hmbox1b^{LD/LD}*) mutants as it is where Hmbox1b was not expressed.

The early onset of Hmbox1b expression in the pronephros during its formation raises the question of whether Hmbox1b is also expressed in the newly formed nephrons as the kidney transitions into the mesonephros. Understanding this could reveal Hmbox1b's role in nephrogenesis beyond embryonic development and its potential involvement in nephron regeneration, given that the zebrafish kidney's is known to have regeneration capability¹⁸¹. Speculatively, Hmbox1b could also play a role in regulating telomere length in these nephrons during nephrogenesis or regeneration. However, there was insufficient evidence to support this hypothesis, as IP failed to enrich any active components of telomerase in embryos, and no IF staining for Cajal bodies was performed. Additionally, the Hmbox1b staining in the

notochord and pronephros does not form foci, which hinders the verification of Hmbox1b localization to the telomeres.

Apart from its renal function, the zebrafish adult kidney also acts as a hematopoietic marrow, similar to the bone marrow in mammals, and is sometimes referred to as the kidney marrow or head kidney^{182,183}. Notably, the telomere length of kidney marrow is similar to the testis, which is longer than those in the fins¹²⁵. These long telomeres are essential for maintaining hematopoietic function. This also raises the question of whether Hmbox1b contributes to maintaining telomere length in hematopoietic stem cells or plays a role in hematopoiesis, as the human HMBOX1 ortholog has been shown to be highly expressed in resting natural killer (NK) cells¹⁸⁴.

However, there was a lack of evidence to support that Hmbox1b participates in hematopoietic function during early development. The spatial expression of Hmbox1b, in the pronephros and notochord, does not coincide with the sites of blood cell production during early embryogenesis. Prior to the formation of the kidney marrow, hematopoiesis in zebrafish occurs at distinct locations. Primitive erythropoiesis takes place in the intermediate cell mass (ICM) before 24 hpf, while primitive myelopoiesis occurs in the anterior lateral mesoderm (ALM)^{183,185}. After circulation begins at 24 hpf, hematopoiesis shifts to the posterior blood island (PBI) briefly before transitioning to the aorta-gonad-mesonephros (AGM) region. After 2 dpf, hematopoietic stem cells (HSCs) migrate to various sites, including the caudal hematopoietic tissue (CHT) for blood development, the thymus for lymphocyte maturation, and the kidney marrow, which becomes the primary site of hematopoiesis¹⁸³.

Importantly, IF staining did not detect Hmbox1b expression at any of these hematopoietic sites. Although the ICM is located near the pronephros, the two do not overlap spatially. IF-FISH experiments conducted in this study confirmed that Hmbox1b consistently localizes to the pronephros, as indicated by co-localization with the pronephros marker *cdh17*. Furthermore, a previously published *in situ* hybridization staining¹⁸⁶ has confirmed that the ICM and pronephros are two adjacent tissues, and they do not overlap with each other.

So far, all studies have been conducted on embryos, leaving the roles of Hmbox1a and Hmbox1b in adults unaddressed. This study has focused on the spatiotemporal expression patterns of Hmbox1a and Hmbox1b, showing distinct expressions during embryonic development. Notably, their expressions were mainly linked to tissues with telomerase activity, such as oocytes, blastula-stage embryos, and the future pronephros, kidney. While it remains speculative whether Hmbox1a/b paralogues have a role in telomere length regulation, their expressions were found in tissues associated with telomerase activity. Unfortunately, this study failed to gather any strong evidence to support this hypothesis.

Although attempts were made to conduct deeper analysis or functional tests specific to each tissue of interest, they were mainly unsuccessful due to technical limitations. Future work should focus on targeted adult tissues. For instance, the spatiotemporal expression of Hmbox1a in the ovary and Hmbox1b in the kidney should be validated in adult zebrafish. To assess their influence on telomeres, future experiments could include measuring telomere length, investigating co-localization of Hmbox1a and Hmbox1b with Cajal bodies or APBs, and evaluating their association with active telomerase components through IP. ChIP-seq could

also be used to decode the regulatory aspects of these paralogues. Additional studies, such as kidney function tests, blood panel analyses, and regeneration assays, would help clarify whether *Hmbox1b* plays a role in both renal and hematopoietic functions in the adult kidney.

Zbtb48

In comparison to the *Hmbox1a/b* study, which utilized IF to investigate spatiotemporal expression, the spatiotemporal expression of *Zbtb48* was primarily explored using publicly available scRNA-seq data. Analysis of these data indicated that PGCs exhibited the highest *zbtb48* transcript levels in 5 dpf larvae^{151,152}, while the gonads showed notable expression among adult tissues¹⁵³. Further analysis of gonad scRNA-seq datasets—including 40 dpf ovaries¹⁵⁵ and testes aged 5 to 22 months¹⁵⁴—consistently demonstrated that *zbtb48* transcript is highly expressed in specific gonad cell subsets, such as germ stem cells and gametocytes.

Gonads, such as the testes, possess longer telomeres compared to most somatic tissues¹²⁵. While telomere length in somatic cells declines progressively over time, germline stem cells depend on telomerase activity to maintain their telomeres and sustain mitotic cell division necessary for reproduction²⁹. High *Zbtb48* expression in the stem cells may help reinforce the upper limit of telomere length, consistent with its proposed role as a telomere trimmer⁷⁴. Beyond germ stem cells, high *zbtb48* transcript levels were also reported in gametocytes, which undergo meiosis, a process that halves the chromosome number. However, the role of *Zbtb48* in regulating telomere length during meiosis remains unclear, as there are insufficient studies to confirm whether telomere length changes during meiosis¹⁸⁷.

Given that scRNA-seq profiling indicated *zbtb48* transcript expression in PGCs at 5 dpf and adult gonads, IF staining was attempted using the *Zbtb48* antibodies developed in this study to validate *Zbtb48* presence in germ cells. Before the specialization of cells into PGCs, germline materials are localized within germplasms at cleavage furrows during early cleavage stages¹⁵⁸, identifiable as early as the 4-cell stage. When IF staining was performed at the 4-cell stage and at 1 dpf using different *Zbtb48* antibodies, the results were inconclusive. In wild-type samples, *Zbtb48*-T2-BTB antibody stained the germplasm at the 4-cell stage and PGCs at 1 dpf, but this staining was still present in *zbtb48*^{-/-} mutants. In contrast, *Zbtb48*-T3-BTB antibody did not stain the germplasm or PGCs but instead labeled the epidermal layer of wild-type embryos at 1 dpf, with this signal absent in *zbtb48*^{-/-} mutants, supporting the specificity of the *Zbtb48*-T3-BTB staining. However, the *Zbtb48*-T2-BTB antibody did not produce this epidermal signal. Unfortunately, as IF was not performed on 5 dpf larvae or adult gonads, where *Zbtb48* expression was predicted based on RNA-seq data, it remains inconclusive which of the two *Zbtb48* antibodies reliably represents the protein's expression.

Proceeding with the inconclusive IF result, proteome analyses were performed on various tissues that potentially express *Zbtb48*. This includes the 5 dpf whole larvae, and whole gonad tissues from 40 dpf ovaries and testes, and 10.5-month-old testes. Despite scRNA-seq data indicating high *zbtb48* transcript levels in these tissues, *Zbtb48* protein was barely detected in any proteome except for 40 dpf ovaries. The detection in 40 dpf ovaries likely reflects the higher population of germ stem cells and early oocytes in the just differentiated ovaries.

Gonads remain bipotential until ~21 dpf and start differentiation around 25 dpf^{160,161}. Ovary differentiation completes earlier (~35 dpf) than testis differentiation (~45 dpf), as testes undergo an additional step involving apoptosis of oocyte-like germ cells before differentiating into testes¹⁸⁸. Thus, the gonads analyzed likely represent different developmental stages: 40 dpf ovaries had recently completed differentiation, 40 dpf testes were still differentiating, and 10.5-month-old testes were fully matured.

The differing cell population compositions in the gonads likely influence Zbtb48 protein levels. Zbtb48 expression appears higher in 40 dpf ovaries, which are primarily undergoing germ stem cell expansion and meiosis, compared to 40 dpf testes, which are dominated by ongoing differentiation and may be lacking the cell types that are highly expressing Zbtb48. However, Zbtb48 was also undetected in the proteome analysis of 10.5-month-old testes, despite the presence of germ stem cells and spermatocytes population, which supposedly expresses Zbtb48¹⁵⁴. It is likely that Zbtb48 was not detected due to the higher proportion of spermatids in the testes, which have minimal Zbtb48 expression. Thus, this is likely to reduce the overall Zbtb48 levels in 10.5-month-old matured testes¹⁵⁴. However, this low Zbtb48 level also conflicts with previous studies indicating that porcine mesenchymal stem cells (pMSC) exhibit higher ZBTB48 levels in later passages¹¹⁰. Nevertheless, this finding suggests that Zbtb48 expression is very low and confined to specific cell types.

Despite the limited detection of Zbtb48, the proteome analyses revealed a skewed proteome distribution, with a significant number of proteins downregulated in the differentiated gonads of *zbtb48*^{-/-} mutants. This supports its role as a transcriptional activator, in addition to its previously reported function in regulating telomere length in human studies⁷³. In 40 dpf ovaries, 850 proteins were found downregulated and 38 were upregulated, while in 10.5-month-old testes, 227 proteins were downregulated and 9 were upregulated. This trend was less pronounced in 40 dpf testes (30 downregulated, 22 upregulated) and absent in 5 dpf proteomic data (80 and 99 downregulated, 109 and 152 upregulated, respectively). In contrast to the proteomic data, the transcriptomic analysis of 5 dpf samples showed a higher proportion of downregulated (n=25) than upregulated (n=7) genes in *zbtb48*^{-/-} mutants. Notably, human U2OS and HeLa cells also showed this similar level of genes downregulation (23 and 11, respectively)⁷³. Overall, these results suggest that the protein downregulation intensifies in the gonad as the development progresses.

Among the dysregulated genes, the nuclear-encoded mitochondrial protein Mtfp1 stood out due to its consistent downregulation across all proteome analyses, with the strongest effect observed in 10.5-month-old testes. This finding aligns with human studies where *MTFP1* was downregulated in *ZBTB48* knockout cancer cell lines⁷³, indicating it is an evolutionarily conserved target of ZBTB48. Although the transcriptomic analysis of 5 dpf samples did not reflect a significant downregulation of *mtfp1* with the threshold used, qRT-PCR confirmed its significant reduction at 5 dpf. Similarly, downregulation of *mtfp1* was also observed via qRT-PCR in 33 dpf differentiating gonads, 40 dpf differentiating testes, and 1-year-old differentiated testes. These results highlight *mtfp1* as a consistent target of Zbtb48 and underscore its critical role across developmental stages, even in the absence of detectable Zbtb48 expression.

MTFP1 (also known as MTP18) is a protein essential for mitochondrial fission in human cells¹⁸⁹. The knockout of *MTFP1* leads to mitochondrial elongation¹⁸⁹, a phenotype also observed

in human *ZBTB48* knockout cell lines ⁷³. However, this mitochondrial elongation was not observed in *zbtb48*^{-/-} zebrafish mutants at the blastula stage (~3 hpf). Instead, mitochondria appeared clustered in both wild-type and *zbtb48*^{-/-} mutant embryos. Such clustered mitochondrial morphology has also been previously observed in wild-type zebrafish blastulae in another study ¹⁹⁰. This suggests that zebrafish mitochondrial morphology may inherently differ from the tubular network observed in human cell lines. Moreover, studies on *Mtfp1* knockout mice have shown no significant morphological changes in the mitochondria of cardiomyocytes ¹⁹¹ or hepatocytes ¹⁹², and concluded its role in mitochondrial fission being dispensable. These findings highlight a contrast between the *MTFP1* knockout phenotypes between human cell lines and tissues of model organisms.

Beyond mitochondrial morphology, MTFP1 also influences mitochondrial function. For instance, knockdown of *MTFP1* in oral squamous cell carcinoma caused cell cycle arrest, increased apoptosis, and reduced reactive oxygen species (ROS) production ¹⁹³. In mice, *Mtfp1* knockout resulted in impaired mitochondrial function in cardiomyocytes, leading to heart failure ¹⁹¹. Interestingly, in contrast, the knockout of *Mtfp1* in hepatocytes increased mitochondrial respiration and provided protection against apoptotic liver damage ¹⁹². These findings indicate that MTFP1 exerts varying effects on cellular and tissue functions depending on the tissues. However, in this study, the impact of *zbtb48* knockout on mitochondrial function in zebrafish was not explored beyond structural changes.

Apart from being structurally dynamic organelles that continuously undergo fission and fusion ¹⁹⁴, mitochondria are also multifunctional. Mitochondria are the main site of energy production, and they are also the primary source of ROS generation, as they utilize the electron transport chain to produce ATP ¹⁹⁵. Importantly, they also regulate redox homeostasis and apoptosis ¹⁹⁵⁻¹⁹⁹. Dysregulation of ROS can cause cellular oxidative stress, which can lead to damage of cellular components, including DNA ^{200,201}. Telomeres, in particular, are vulnerable, as they are primarily composed of guanine, the nucleotide most susceptible to oxidative damage ²⁰². Although the oxidative stress levels and oxidative DNA damage in the cells of *zbtb48*^{-/-} mutants were not measured, the omics analysis revealed some interesting findings.

Both transcriptomic and proteomic analyses of *zbtb48*^{-/-} mutants revealed a consistent trend of dysregulation in genes associated with redox homeostasis. However, none of these candidates were identified in the previous human cell line study. They are *aldo-keto reductase (akr7a3)* and *guanylate-binding protein (gbp3)*, which both were frequently downregulated in the transcriptome and proteome of 5 dpf and gonads in *zbtb48*^{-/-} mutants. In contrast, *zbtb48* knockout significantly upregulated *zgc:153284* (predicted to have electron transfer and protein disulfide oxidoreductase activities) at both transcript and protein levels in the 5 dpf larvae. However, this upregulation was not observed in gonad proteomes. It remains unclear whether these genes are directly regulated by *Zbtb48* at their promoters or a consequence of *Mtfp1* dysregulation. Further studies are needed to explore the mechanistic connections between these candidates and *Zbtb48*, as well as their association with mitochondrial functions.

Apart from *Zbtb48*'s regulation of *Mtfp1* at the transcriptional level, the mouse MTFP1 ortholog is also regulated at the translational level by the nutrient-sensing mTORC1 pathway in mouse embryonic fibroblasts (MEF) cell line ²⁰³. The mTORC1 pathway has drawn attention recently due to its role in promoting longevity through calorie restriction (CR) ²⁰⁴. CR has been shown

to enhance antioxidant activity, reduce ROS levels, and protect against oxidative stress^{205,206}. However, it remains unclear whether ZBTB48 also participates in the mTORC1 pathway through MTFP1 regulation.

Taken together, numerous connections between Zbtb48 and mitochondrial functions have been drawn in this study. Concomitantly, its high expression in zebrafish gonad tissues where telomerase activity is elevated, suggests a deeper relationship between Zbtb48's role in mitochondrial function, telomerase activity, and telomere maintenance. Further exploration of Zbtb48 on the role of mitochondria and telomerase activity may help to better understand Zbtb48's function at telomere. So far, interconnections between these factors have been briefly studied²⁰⁷. For example, telomere dysfunction has been shown to impair mitochondrial biogenesis and function through p53 activation²⁰⁸. Additionally, under oxidative stress, TERT is increasingly imported into the mitochondria, where it helps reduce mitochondrial ROS levels and protect both mitochondrial and nuclear DNA from damage^{209,210}.

To gain a deeper understanding of how mitochondrial dysregulation and high telomerase activity affect the gonads in *zbtb48*^{-/-} mutants, further studies should focus on various functional assays in gonads from different developmental stages. These should include, but not be limited to, the assessment of oxidative stress, TIF, localization of TERT in mitochondria, senescence markers, fertility, telomere length, proliferative capability, and differentiation potential to address questions that remain unresolved in this study.

Conclusion

The initial aim of this study was to investigate the telomeric effects of Hmbox1a/b and Zbtb48 at an organismal level. However, no significant changes in telomere length were observed in the fins of first-generation homozygous mutant animals, despite successfully generating the CRISPR knockout lines. Similarly, lifespan and fertility appeared unaffected in these mutants for at least the first two inbred generations. It is possible that a longer study duration is required to observe significant effects on telomere length in animal models.

On the other hand, this study successfully examined the spatiotemporal expression of these proteins during early zebrafish development. For the first time, the physiological spatiotemporal expression of these proteins during development has been reported, as previous studies were mostly conducted on cell lines, focusing primarily on cancerous cells. Using IF, it was found that Hmbox1a is maternally inherited and expressed during early embryonic development. In contrast, Hmbox1b is expressed later in development, with its expression unaffected by Hmbox1a. Specifically, Hmbox1b was found in the pronephros and notochord, with the signal in the pronephros gradually fading as segmentation completes. Meanwhile, scRNA-seq data have revealed that Zbtb48 is highly expressed in the germ stem cells and gametocytes of the gonads.

Unfortunately, due to time and technical limitations, the functional roles of Hmbox1a, Hmbox1b, and Zbtb48 were not uncovered in this study. However, the proteomic analyses in the Zbtb48 study successfully recapitulated its regulatory role on Mtfp1 in zebrafish. This suggests an important relationship between them, as MTFP1 remains a conserved target of ZBTB48 throughout evolution.

As this study primarily focused on the spatiotemporal expression of these proteins during early development, future research should expand the investigation to specific adult organs where they are expressed. This would involve conducting organ-specific functional tests, molecular assays, and telomere length measurements.

Materials

Biological materials

Cell lines

Cell line	Source
BRF41	Zebrafish fibroblast (CVCL_4131)
PAC2	Zebrafish fibroblast (CVCL_5853)
HEK293T	

E. coli strains

Strain	Source
NEB® 5-alpha Competent <i>E. coli</i> (High Efficiency)	NEB, C2987
One Shot™ TOP10 Chemically Competent <i>E. coli</i>	Invitrogen, C404010
BL21 Competent <i>E. coli</i>	NEB, C2530
NEB® Stable Competent <i>E. coli</i>	NEB, C3040

Zebrafish lines

Fish line	Type	Source
AB/Tübingen strain	Wild-type	KIT - European Zebrafish Resource Center (EZRC)
<i>Tg(vasa:eGFP)</i>	Transgenic line	KIT - European Zebrafish Resource Center (EZRC) (from Krøvel, & Olsen ¹⁵⁶)
<i>Tg(wt1b:eGFP)li1</i>	Transgenic line	KIT - European Zebrafish Resource Center (EZRC) (from Zhou et al. ¹⁴²)
<i>hmbox1a^{LD/LD}</i>	CRISPR-Cas mutant	Created during this study
<i>hmbox1a^{SD/SD}</i>		
<i>hmbox1b^{LD/LD}</i>		
<i>hmbox1b^{SD/SD}</i>		
<i>hmbox1a^{LD/LD}; hmbox1b^{LD/LD}</i>		
<i>hmbox1a^{SD/SD}; hmbox1b^{SD/SD}</i>		
<i>zbtb48^{-/-}</i>		

Plasmids

Plasmid number	Glycerol number	Plasmid Insert	Backbone	Description/ Purpose
P1157	G515	Full length <i>hmbox1a</i> -Main (with stop codon)	pCR8/GW/TOPO	Holding vector
P1155	G513	Full length cDNA of <i>hmbox1b</i> -X1 (with stop codon)		
P1139	G507	Full length cDNA of <i>hmbox1b</i> -X2 (with stop codon)		
P1156	G514	Full length cDNA of <i>hmbox1b</i> -X3 (with stop codon)		
P1142	G509	Full length cDNA of <i>zbtb48</i> (with stop codon)		
P1137	G504	C-terminal of Hmbox1a	pCoofy4	Epitope expression (N-terminal His ₆ -MBP)
P1133	G506	C-terminal of Hmbox1b		
P1234	G516	BTB domain of Zbtb48		
P1136	G501	ZNF domain of Zbtb48		
P1019		(Empty)	PX459-V2	CRISPR-Cas9 For transfection in PAC2 cells
P1247	G534	<i>hmbox1a</i> , gRNA1		
P1249	G535	<i>hmbox1a</i> , gRNA2		
P1248	G536	<i>hmbox1b</i> , gRNA1		
P1250	G537	<i>hmbox1b</i> , gRNA2		
P1285	G550	(Empty)	plentiCRISPRV2-Puro	CRISPR-Cas9 For transduction in PAC2 cells
P1281	G546	<i>hmbox1b</i> , gRNA1		
P1282	G547	<i>hmbox1b</i> , gRNA2		
P1237	G524	<i>hmbox1a</i> , gRNA1	pDR274 (SP6 promoter)	CRISPR gRNA synthesis for embryo injection
P1238	G525	<i>hmbox1a</i> , gRNA2		
P1239	G526	<i>hmbox1b</i> , gRNA1		
P1240	G527	<i>hmbox1b</i> , gRNA2		
P1243	G530	<i>zbtb48</i> , gRNA1		
P1244	G531	<i>zbtb48</i> , gRNA2		
		(Empty)	pCS2-eGFP	For transfection test in PAC2 cells
		(Empty)	pMDLg/pRRE	Lentivirus packaging plasmids
		(Empty)	pRSV-Rev	
		(Empty)	pMD2.G	

Oligonucleotides

Primers for amplification and genotyping (1/2)

Oligo number	Sequence	Description/ Purpose
4211	ATGTCCCATTATACAGATGAG	Full-length cDNA of <i>hmbox1a</i> for pCR8/GW/TOPO cloning; Forward primer Genotyping- T7E1 for <i>hmbox1a</i>
4212	TCACTCATCATCCAGCAC	Full-length cDNA of <i>hmbox1a</i> for pCR8/GW/TOPO cloning; Reverse primer with stop codon
4214	ATGTCTCACTACACAGACGAG	Full-length cDNA of <i>hmbox1b</i> for pCR8/GW/TOPO cloning; Forward primer Genotyping- T7E1 for <i>hmbox1b</i>
4215	TCAGTCGTCATCTATCGCTTC	Full-length cDNA of <i>hmbox1b</i> for pCR8/GW/TOPO cloning; Reverse primer with stop codon
4220	ATGGAGAAACTTGCAGAA	Full-length cDNA of <i>zbtb48</i> for pCR8/GW/TOPO cloning; Forward primer Genotyping- <i>zbtb48</i>
4221	TTAGTCTTTTCTATCAGCCTC	Full-length cDNA of <i>zbtb48</i> for pCR8/GW/TOPO cloning; Reverse primer with stop codon
4223	AAGTTCTGTTCCAGGGGCC CATGGGATTGATGTCCAAGC	Epitope sequence for Hmbox1a for pCoofy4 cloning; Forward primer Genotyping- Identifying <i>hmbox1a</i> variants
4224	CCCAGAACATCAGGTTAATGGCG CATCTCTACTGCCAGAGC	Epitope sequence for Hmbox1a for pCoofy4 cloning; Reverse primer Genotyping- Identifying <i>hmbox1a</i> variants
4286	AAGTTCTGTTCCAGGGGCC GGTTGCGATGTTGCCTATTTT	Epitope sequence for Hmbox1b for pCoofy4 cloning; Forward primer
4287	CCCAGAACATCAGGTTAATGGCG TAGGGCGAGGATGCTGTG	Epitope sequence for Hmbox1b for pCoofy4 cloning; Reverse primer
4229	AAGTTCTGTTCCAGGGGCC GTGCTGTCTTTGCTCAACA	Epitope sequence for Zbtb48-BTB for pCoofy4 cloning; Forward primer

Primers for amplification and genotyping (2/2)

Oligo number	Sequence	Description/ Purpose
4719	CCCCAGAACATCAGGTTAATGGCG GGTTTTGATGGATCACCTTCA	Epitope sequence for Zbtb48-BTB for pCoofy4 cloning; Reverse primer Genotyping- <i>zbtb48</i>
4231	AAGTTCTGTTCCAGGGGCC ACTCAAGCCAGTCTGGACA	Epitope sequence for Zbtb48-ZNF for pCoofy4 cloning; Forward primer
4232	CCCCAGAACATCAGGTTAATGGCG AATTTCTCATGCCCTTGTG	Epitope sequence for Zbtb48-ZNF for pCoofy4 cloning; Reverse primer
4767	CTCGACTCGGTCATCGA	Genotyping- T7E1 for <i>hmbox1a</i> and Genotyping- <i>hmbox1a^{LD}</i> and <i>hmbox1a^{SD}</i>
4774	CTCATGAGCTCCTCCAC	Genotyping- T7E1 for <i>hmbox1b</i>
4835	GCAACAGTCTGGCATTGAGA	Genotyping- <i>hmbox1a^{LD}</i>
4820	AAACTGAGCATGCCGATGGCCAAAC	Genotyping- <i>hmbox1a^{LD}</i> (also <i>hmbox1a</i> gRNA1)
4768	GTGGAGGAGCTCATGAG	Genotyping- <i>hmbox1b^{LD}</i>
4976	GTGTTGTGCTGTGTGTCAGTGTGTTT	Genotyping- <i>hmbox1b^{LD}</i>

Primers for plasmid sequencing

Oligo number	Sequence	Description/ Purpose
404	GTTGCAACAAATTGATGAGCAATGC	pCR8/GW/TOPO
405	GTTGCAACAAATTGATGAGCAATTA	pCR8/GW/TOPO
1421	GATGAAGCCCTGAAAGACG	pCoofy4
2191	GAGGGCCTATTTCCCATGATTCC	PX459-V2 & plentiCRISPRV2-Puro
3725	TGAAAACGACGGCCAGT	pDR274

gRNA oligonucleotides for PX459 & pLentiCRISPRV2-Puro

Oligo number	Sequence	Description/ Purpose
4819	CACCG TTTGGCCATCGGCATGCTCA	<i>hmbox1a</i> gRNA1-Top
4820	AAAC TGAGCATGCCGATGGCCAAA C	<i>hmbox1a</i> gRNA1-Bottom
4821	CACCG GTGCTACAACCGGCATAG	<i>hmbox1a</i> gRNA2-Top
4822	AAAC CTATGCCGTTGTAGCAC C	<i>hmbox1a</i> gRNA2-Bottom
4823	CACCG TGAGGGCGTGCACGATCT	<i>hmbox1b</i> gRNA1-Top
4824	AAAC AGATCGTGCACGCCCTCGA C	<i>hmbox1b</i> gRNA1-Bottom
4825	CACCG AACTGCGTCTGAGTTGCTG	<i>hmbox1b</i> gRNA2-Top
4826	AAAC CAGCAACTCAGACGCAGTT C	<i>hmbox1b</i> gRNA2-Bottom

gRNA oligonucleotides for pDR274

Oligo number	Sequence	Description/ Purpose
4732	TAGG TTTGGCCATCGGCATGCTCA	<i>hmbox1a</i> gRNA1-Top
4733	AAAC TGAGCATGCCGATGGCCAAA	<i>hmbox1a</i> gRNA1-Bottom
4734	TAGG GTGCTACAACCGGCATAG	<i>hmbox1a</i> gRNA2-Top
4735	AAAC CTATGCCGGTTGTAGCAC	<i>hmbox1a</i> gRNA2-Bottom
4736	TAGG TCGAGGGCGTGCACGATCT	<i>hmbox1b</i> gRNA1-Top
4737	AAAC AGATCGTGCACGCCCTCGA	<i>hmbox1b</i> gRNA1-Bottom
4738	TAGG AACTGCGTCTGAGTTGCTG	<i>hmbox1b</i> gRNA2-Top
4739	AAAC CAGCAACTCAGACGCAGTT	<i>hmbox1b</i> gRNA2-Bottom
4744	TAGG CAGGCCAGGACATTGCGAT	<i>zbtb48</i> gRNA1-Top
4745	AAAC ATCGCAATGTCCTGGCCTG	<i>zbtb48</i> gRNA1-Bottom
4746	AGG CATCGGGACACTCTAAAGGG	<i>zbtb48</i> gRNA2-Top
4747	AAAC CCCTTTAGAGTGTCCCGATG	<i>zbtb48</i> gRNA2-Bottom

qRT-PCR primers

Oligo number	Sequence	Description/ Purpose
NW542	CCTCACTTTGAGCTCCTCCAC	<i>β-actin</i> - Forward
NW543	GACCCACGATGGATGGGAAG	<i>β-actin</i> - Reverse
5655	ACATTTTCGCACTCAAGCCAG	<i>zbtb48</i> - Forward
5656	GTGTGGTTCGACCTTCTTGGT	<i>zbtb48</i> - Reverse
5679	CTGGGTCAGATTCCTGGGTT	<i>mtfp1</i> - Forward
5680	GCCTTTATCGATGGCGTCTG	<i>mtfp1</i> - Reverse
5685	TTCCTGGTCCTTGAGTTGGC	<i>zgc:153284</i> - Forward
5686	ACCCTGATGTTGCTTTACCTCT	<i>zgc:153284</i> - Reverse
5687	CACCGCGCTGATGTACAATG	<i>akr7a3</i> - Forward
5688	GGTCGCGATTTCGAAGTGT	<i>akr7a3</i> - Reverse
5714	AACTGATGTCCTTCTGGAAACA	<i>atox1</i> - Forward
5715	GCAGCATGATTCTGACCTTATTGT	<i>atox1</i> - Reverse
5716	TGAATGTGGACACGAAGCCT	<i>col28a2a</i> - Forward
5717	CCAGGGGAGCCGATTCTTTTA	<i>col28a2a</i> - Reverse
5718	GGGGACCCAGAAGGTGAAAT	<i>hdac10</i> - Forward
5719	TTTTTCCTGCAGCCAAAGGC	<i>hdac10</i> - Reverse
5681	TGTTTCGAGACCCCTCTTTGC	<i>pxmp2</i> - Forward
5682	AGCAGACAGAATACCACTCGT	<i>pxmp2</i> - Reverse
5675	CTCCGATGGCTCGCTGTTTA	<i>vwa5a</i> - Forward
5676	GTGCAGACTTTGCTTCACGG	<i>vwa5a</i> - Reverse
5677	GAAAGGACGACTAAAAGCGGTG	<i>ppp3cb</i> - Forward
5678	CAAACACCTCTCGCATGCTC	<i>ppp3cb</i> - Reverse

Telomere pull down baits

Oligo number	Sequence	Description/ Purpose
4031	TTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG GTTAGGGTTAGGGTTAGGGTTAGGG	TTAGGG- Forward
4032	AACCCTAACCCCTAACCCCTAACCCCTAACCC TAACCCTAACCCCTAACCCCTAACCCCT	TTAGGG- Reverse
4033	GTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGA GTGTGAGTGTGAGTGTGAGTGTGAGT	GTGAGT- Forward
4034	ACACTCACACTCACACTCACACTCACACTCACACT CACACTCACACTCACACTCACACTC	GTGAGT- Reverse

RNA-FISH (smFISH) *cdh17* probe set sequences (generated by Stellaris)

Probe #	Probe (5'-> 3')	Probe position *	Percent GC
1	CCTGCTCAAGTCAAAGTAGG	18	50.00%
2	GTAAAACCTGTCAGCAGAGC	81	50.00%
3	CAAAGTGCACCTCTCATCG	110	50.00%
4	CATGACCAATGCTGACGAGA	144	50.00%
5	TCTGGCACATCTAAAACCGT	202	45.00%
6	CGTCTTCAACTGCAGATGTA	255	45.00%
7	CCATCTTCTGAAAGTGCTTC	394	45.00%
8	ATCACTTGCAGTACAACCTG	436	45.00%
9	AATGCTCTCTAATACTGCCA	501	40.00%
10	TCTGAAGCAAACACCTGCAC	541	50.00%
11	ATCTGTCCATTGTTTGGATT	652	35.00%
12	TGTATGTTACGCTAGGTCTG	708	45.00%
13	CTTTTTCTTCAGGACGTCAG	752	45.00%
14	GTGCAATGTTGTTCTTTGGA	789	40.00%
15	TTCTTTCTGCACGTTCAACA	834	40.00%
16	TAGTCAGGATCCTGGAGAAC	865	50.00%
17	CTCAATGAGTTCACAGCGTT	919	45.00%
18	AGAGGTTTTGAAGGATGGCT	960	45.00%
19	TCTAATGGTGATTGGTCCAG	989	45.00%
20	GGTATTCCTCATCTGTATTT	1011	35.00%
21	CATTTCTTTCTCTTCTCTGT	1157	35.00%
22	TTCTGCTATGACTACCAGTA	1181	40.00%
23	CAATCGAATTACCTACGGGC	1314	50.00%
24	GTGTAGGTTAAAGCTGAGCT	1375	45.00%
25	TGATTGGCCACTTTGATTTTC	1459	40.00%
26	GCACTTGTTTTCTCTGGAAG	1482	45.00%
27	AGTCACTTCAAAGGTGAGCT	1511	45.00%
28	GATCACCTTGATGATTGCTT	1556	40.00%
29	GGGAACACTGTAAGTTCCAT	1616	45.00%

Probe #	Probe (5'-> 3')	Probe position *	Percent GC
30	CTGTGGCTTTAATGTTGAGC	1662	45.00%
31	TGTGATGTGATACTCCACTC	1712	45.00%
32	CAACCTCAATGGCGAAGAGG	1749	55.00%
33	AGATTGTAGACGCTCTGGAG	1822	50.00%
34	GTCTACCAGCTCAATGACAA	1913	45.00%
35	TGTTTTCTGGCACATTGACA	1971	40.00%
36	TTGGCTTCAGCTTTCATTAG	2008	40.00%
37	TGTCAACATTCAGCTCCAAC	2082	45.00%
38	AGATGAATGTCCACCTTCAT	2206	40.00%
39	AGGGTAGTTGTCGTTTACAT	2231	40.00%
40	TTGTGATTTTCGGCTGATGG	2357	45.00%
41	ATCCACAGGTTTGATCTCAA	2381	40.00%
42	CTTCTTCAGGATCAGTTTCG	2411	45.00%
43	AGTGACGTTCTGTTCACTTG	2438	45.00%
44	CAGCGTTGTCTAAGACGTTG	2466	50.00%
45	GAGAAGCTGGCTCGATGTAA	2547	50.00%
46	AGGCCACATGGCAAATAAGT	2862	45.00%
47	CATTTGACCTCAGCAAGACA	3185	45.00%
48	TGCTTTTAACGCCTTTTGTA	3240	35.00%

Antibodies

Primary antibodies

Antibody	Supplier	Catalogue number
Mouse anti-GFP	Roche	11814460001
Mouse anti-human HMBOX1	Obtained from Kappei et al. ⁷⁷	
Rabbit anti-Actin	Sigma	A2066
Rabbit anti-Cas9	Diagenode	C15310258
Rabbit anti-pan Histone 3	Diagenode	C15310135
Rabbit anti-human MTFP1	Sigma	SAB4301167
Rabbit anti-Hmbox1a-T1	Raised in this study Purified by myself	
Rabbit anti-Hmbox1a-T3		
Rabbit anti-Hmbox1b-T2		
Rabbit anti-Zbtb48-BTB-T2		
Rabbit anti-Zbtb48-ZNF-T2		
Rabbit anti-Zbtb48-BTB-T3		
Rabbit anti-Zbtb48-ZNF-T3		
Rabbit anti-Hmbox1b-T2 (M)	Raised in this study Purified by IMB Protein Production Core Facility	
Rabbit anti-Zbtb48-BTB-T3 (M)		
Rat anti-zebrafish Ziwi	Raised by Ketting group	396
Rabbit IgG	Sigma	I5006

Secondary antibodies

Antibody	Supplier	Catalogue number
Anti-mouse IgG, HRP-linked	Cell Signaling Technology	7076
Anti-rabbit IgG, HRP-linked	GE Healthcare	NA934V
Anti-mouse IgG (H+L), AF488	Invitrogen	A32723
Anti-rabbit IgG (H+L), AF488	Invitrogen	A11034
Anti-rabbit IgG (H+L), AF568	Invitrogen	A11011
Anti-rat IgG, IRDye 680RD	Li-COR	926-68076

Enzymes and Polymerase

Enzymes and Polymerase	Supplier	Catalogue number
BbsI	NEB	R0539
BsaI	NEB	R3733
BsmBI-V2	NEB	R0580
DNase I (RNase-free)	NEB	M0303
EnGen® Spy Cas9 NLS	NEB	M0646
HF-DNA Polymerase	IMB Protein Production Core Facility	
Pronase	Sigma	10165921001
Proteinase K	Sigma	P2308
Q5® High-Fidelity DNA Polymerase	NEB	M0491
RecA	NEB	M0249
T4 DNA Ligase	NEB	M0202
T4 Polynucleotide Kinase (PNK)	Thermo Scientific	EK0032
T7 Endonuclease I (T7E1)	NEB	M0302
Taq DNA Polymerase	IMB Protein Production Core Facility	
Trypsin-EDTA (0.25%), phenol red	Gibco	25200056
Trypsin, MS approved	Serva	37286.03

Chemical/ Reagent

Chemical/ Reagent	Supplier	Catalogue number
1-phenyl 2-thiourea (PTU)	Sigma-Aldrich	P7629
100 bp DNA Ladder	Jena Bioscience	M214L
Acetic acid	Sigma-Aldrich	33209
Acetone	Carl-Roth	9372
Acetonitrile	VWR	20048.32
Agarose gel	Sigma-Aldrich	A9539
Amersham ECL Detection Reagents	cytiva	RPN2105
Ammonium bicarbonate	Sigma-Aldrich	A6141
Ampicillin	IMB Core Facility Media lab	
Benzoylated dialysis tubing	Sigma-Aldrich	D7884
Blocking reagent	Roche	11096176001
BlueEye Prestained Protein Marker	Jena Bioscience	PS-104
Bovine Serum Albumin	Sigma-Aldrich	A3294
C18 material	Dr Maisch	r119.aq.0001
Calcium chloride, dihydrate	Carl-Roth	5239.1
Chloroform	Carl-Roth	3313.4
cOmplete Mini EDTA-free protease inhibitor tablets	Roche	4693159001
Coomassie Blue G-250	Biozym	902120 (HS-605)
DAPI dihydrochloride powder	Sigma-Aldrich	D9542
Dextran sulfate	Sigma-Aldrich	D8906
Dextran, Alexa Fluor™ 647, 10,000 MW	Invitrogen	D22914
Dextran, Fluorescein, 500,000 MW	Invitrogen	D7136
DL-Dithiothreitol (DTT)	Sigma-Aldrich	D0632-25G
DMSO	Sigma-Aldrich	D2650
dNTPs (4x 100 mM)	Jena Bioscience	NU-1005S
DPBS, no calcium, no magnesium	Gibco	14190094
Dynabeads™ MyOne™ Streptavidin C1	Invitrogen	65002
Dynabeads™ Protein G	Invitrogen	10004D
EDTA	IMB Core Facility Media lab	
Ethanol	Carl-Roth	9065.3
Empore™ SPE Disks	Supelco	66886-U
Fetal Bovine Serum	Gibco	10270106
Formamide	Carl-Roth	6749.1
Formic acid	Merck	1.00264.1000
FuGENE® HD Transfection Reagent	Promega	E2311

Gelatin	Sigma-Aldrich	G7041
GeneRuler 1 Kb	Thermo Fisher Scientific	SM0312
Glycerol	Sigma-Aldrich	G7757
GlycoBlue	Invitrogen	10301575
Heparin	Sigma-Aldrich	H3393
Hepes	Carl-Roth	HN78.2
HisTrap™ HP	GE Healthcare	17-5247-01
Igepal CA-630	Sigma	I8896
Imidazole	Sigma-Aldrich	56750
Iodoacetamide	Sigma-Aldrich	I6125
Isopropyl-β-D-thiogalactopyranosid (IPTG)	AppliChem	A1008,0005
Ispopropanol	Carl-Roth	9866.6
Kanamycin	IMB Core Facility Media lab	
Leupeptin	Serva	51867.02
Magnesium sulfate, anhydrous	Sigma-Aldrich	M7506
Maleic acid	Sigma-Aldrich	M0375
MES buffer	Invitrogen	NP0002
Methanol MS grade	VWR	20864.32
MitoTracker™ Deep Red FM	Invitrogen	M22426
Mounting Medium	ibidi	50001
Amersham™ Protran™ 0.45 μm Nitrocellulose membranes	Cytiva	15259794
NuPAGE Novex 4-12% Bis-Tris Protein Gels, 1.0 mm, 10-well	Invitrogen	NP0321BOX
NuPAGE® Novex® 10% Bis-Tris Gels, 1.0 mm, 10 well	Invitrogen	NP0301BOX
NuPAGE™ LDS Sample Buffer, 4X	Invitrogen	11559166
Paraformaldehyde	Sigma-Aldrich	P6148
Penicillin-Streptomycin (Sterile-Filtered)	Sigma-Aldrich	P0781
Pepstatin A	Serva	52682.02
Phenol red	Gibco	B21710.30
Phenol:Chloroform:isoamyl alcohol (25:24:1)	Invitrogen	1559309
Phenylmethylsulfonyl fluoride (PMSF)	Serva	32395.02
Polybrene	Santa Cruz	sc-134220
Ponceau S	Appllichem	A2935
Potassium chloride	Carl-Roth	6781.1
Potassium phosphate monobasic	Carl-Roth	P018.2
Protein Assay Dye Reagent, 5x	Bio-rad	5000006
Puromycin dihydrochloride	Santa Cruz	sc-108071B
Salmon sperm	Invitrogen	10605543
Skim milk powder	Sigma-Aldrich	70166

Sodium chloride	Thermo Fisher Scientific	15626770
Sodium citrate tribasic hydrate	Sigma-Aldrich	25114
Sodium deoxycholate monohydrate	Sigma-Aldrich	D5670
Sodium dodecyl sulfate	Carl-Roth	4360.1
Sodium phosphate dibasic	Carl-Roth	4984.2
Spectinomycin	IMB Core Facility Media lab	
SuperSignal™ West Pico PLUS Chemiluminescent Substrate	Thermo Scientific	34580
SYBR Safe DNA Gel Stain	Invitrogen	S33102
SYBR™ Power SYBR™ Green PCR Master Mix	Applied Biosystems	4367659
TAMRA-labeled C-rich telomere probe	Eurogentec Deutschland GmbH	507207
Tricane	Sigma-Aldrich	E10521
Tris(2-carboxyethyl)phosphine hydrochloride	Sigma-Aldrich	C4706
Triton X-100	Sigma-Aldrich	X100
TRIzol™ Reagent	Invitrogen	15596026
Tween-20	Sigma	P7949
Vanadyl-ribonucleoside complex	NEB	S1402S
Whatman paper	GE Healthcare	WHA10426892
Yeast tRNA	Invitrogen	54016

Media

Medium	Composition	Supplier	Catalogue number
E3 (Zebrafish embryo medium)	5 mM NaCl 0.17 mM KCl 0.33 mM CaCl ₂ 0.33 mM MgSO ₄	Prepared by IMB Animal Facility	
DMEM, high glucose, pyruvate, no glutamine	Supplemented with 10% FBS 1% Penicillin-streptomycin	Gibco	21969035
L-15 (Leibovitz) Medium	Supplemented with 15% FBS 1% Penicillin-streptomycin	Gibco	11415049
LB Luria Medium	0.17 M NaCl 1 % (w/v) Bacto-Tryptone 0.5 % (w/v) Bacto-Yeast extract Adjusted to pH 7.0 ± 0.2	IMB Core Facility Media Lab	
LB Agar plates	0.17 M NaCl 1 % (w/v) Bacto-Tryptone 0.5 % (w/v) Bacto-Yeast extract 1.5% (w/v) Agar Adjusted to pH 7.0 ± 0.2 Supplemented with respective antibiotics 100 µg/mL Ampicillin 50 µg/mL Kanamycin 100 µg/mL Spectinomycin	IMB Core Facility Media Lab	

Buffers and Solutions

Buffer/ Solution	Composition	Description/ Purpose
HBS, 2X	280 mM NaCl 10 mM KCl 1.5 mM Na ₂ HPO ₄ 12 mM D-glucose 20 mM Hepes	Lentivirus production
IPTG	1 M Isopropyl β- d-1-thiogalactopyranoside (IPTG)	To induce protein expression in BL21 cells
Bacteria Lysis buffer	25 mM Tris-HCl, pH7.5 300 mM NaCl Freshly supplement: 1 mM DTT 20 mM Imidazole 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin 100 μg of DNaseI	To harvest protein expressed from IPTG-induction in BL21 cells
Column binding buffer	20 mM Tris-HCl, pH7.5 500 mM NaCl Freshly supplement: 20 mM Imidazole 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin	Ni Column for protein purification
Column wash buffer	20 mM Tris-HCl, pH7.5 500 mM NaCl Freshly supplement: 50 mM Imidazole 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin	Ni Column for protein purification
Column Elution buffer	20 mM Tris-HCl, pH7.5 500 mM NaCl Freshly supplement: 500 mM Imidazole 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin	Ni Column for protein purification

Buffer/ Solution	Composition	Description/ Purpose
Dialysis buffer	25 mM Tris-HCl, pH7.5 100 mM NaCl 5% Glycerol Freshly supplement: 1mM DTT	Salt exchange for eluant
Tricane	0.02% tricane in E3 water	Zebrafish anaesthesia
Fin-clipping lysis buffer	50 mM KCl 2.5 mM MgCl ₂ 10 mM Tris-HCl, pH 8.0 0.45% Igepal 0.45% Tween-20 0.01% Gelatin Freshly supplement: 100 µg/mL of Proteinase K	Tissue lysis buffer for genotyping
TAE, 10x	0.4 M Tris 0.01 M Na ₂ EDTA*2H ₂ O 0.2 M Glacial acetic acid	Supplied by IMB Core Facility Media Lab (To be used at 1X)
TBE, 10x	0.9 M Tris base 0.9 M Boric acid 0.02 M Na ₂ EDTA*2H ₂ O	Supplied by IMB Core Facility Media Lab (To be used at 1X)
Pronase solution	1 mg/ml Pronase	Zebrafish embryo de-chorion buffer
Ringer solution (calcium-free)	116 mM NaCl 2.9 mM KCl 5.0 mM Hepes, pH 7.2	Zebrafish embryo wash buffer
Ringer solution with PMSF and EDTA	Calcium-free Ringer's solution Freshly supplement: 0.3 mM PMSF 1 mM EDTA, pH7.0	Zebrafish embryo de-yolking buffer
RIPA protein lysis buffer	150 mM NaCl 1% Triton X-100 0.5% Sodium deoxycholate 50 mM Tris, pH 8.0 5 mM EDTA 0.1 % SDS Freshly supplement: 1x cOmplete protease inhibitors (Roche)	Whole cell lysis buffer

Buffer/ Solution	Composition	Description/ Purpose
Zebrafish embryo lysis buffer	25 mM Tris pH7.5 150 mM NaCl 1.5 mM MgCl ₂ 1% Triton Freshly supplement: 1 mM DTT 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin	Whole cell lysis buffer for IP and telomere pull-down
Nuclear extraction buffer A	10 mM HEPES-KOH pH 7.4 1.5 mM MgCl ₂ 10 mM KCl Freshly supplement: 1 mM DTT 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin	Nuclear extraction
Nuclear extraction buffer C	20 mM HEPES-KOH pH 7.4 420 mM NaCl 2 mM MgCl ₂ 0.2 mM EDTA pH8 20% Glycerol Freshly supplement: 1 mM DTT 0.1% Igepal 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin	Nuclear extraction
PBB buffer	150 mM NaCl 50 mM Tris-HCl, pH 7.5 0.5% Igepal 5 mM MgCl ₂ Freshly supplement: 1 mM DTT 1 mM PMSF 1 μM Pepstatin A 2 μM Leupeptin	Telomere pull down and immunoprecipitation buffer
10X Transfer Buffer	1.92 M Glycine 0.25 M Tris	Supplied by IMB Core Facility Media Lab
Transfer buffer	1X Transfer buffer 20% Ethanol	Transfer buffer for Western blot

Buffer/ Solution	Composition	Description/ Purpose
5X PBS	685 mM NaCl 13.5 mM KCl 50 mM Na ₂ HPO ₄ 9 mM KH ₂ PO ₄ Adjust to pH 7.4 ± 0.1	Supplied by IMB Core Facility Media Lab
PBST	0.05% Tween-20 in 1X PBS	Wash buffer for Western blot and microscopy.
Skim milk	5% Skim milk in PBST	Blocking buffer for Western blot
Coomassie staining solution	0.25% Coomassie Blue G-250 10% Acetic acid 43% Ethanol	Mass spectrometry sample preparation
MS ABC	50 mM Ammonium bicarbonate buffer (ABC) (pH 8.0)	Mass spectrometry sample preparation
MS destaining buffer	50% (v/v) 50 mM ABC 50% (v/v) Ethanol 99.9% p.a	Mass spectrometry sample preparation
MS reduction buffer	50 mM ABC 10 mM DTT	Mass spectrometry sample preparation
MS alkylation buffer	50 mM ABC 50 mM Iodoacetamide (IAA)	Mass spectrometry sample preparation
MS trypsin solution	50 mM ABC 1 µg Trypsin (per sample)	Mass spectrometry sample preparation
MS extraction buffer	30% Acetonitrile	Mass spectrometry sample preparation
MS Buffer A	0.1% (v/v) Formic acid in HPLC grade H ₂ O	Mass spectrometry sample preparation
MS Buffer B	80% (v/v) Acetonitrile 0.1% (v/v) Formic acid HPLC grade H ₂ O	Mass spectrometry sample preparation
0.003% PTU	0.003% of 1-phenyl 2-thiourea (PTU) in E3	To prevent pigment formation in embryo development
4% PFA	4% Paraformaldehyde (PFA) in PBS	Sample fixing
10% BSA	10% Bovine serum albumin (BSA) in PBS	Blocking buffer for IF staining
20X SSC	3 M NaCl 0.3 M Sodium citrate tribasic Adjust to pH 7.0	Supplied by IMB Core Facility Media Lab
smFISH wash buffer	10% Formamide 2X SSC	smFISH wash buffer

Buffer/ Solution	Composition	Description/ Purpose
smFISH hybridization buffer	10% Dextran sulfate 10% Formamide 1 mg/mL Yeast tRNA 2X SSC 0.02% BSA 2 mM Vanadyl-ribonucleoside complex	smFISH staining
Telomere-FISH hybridization buffer	1X Blocking reagent (Roche) 70% Formamide 100 µg/mL Yeast tRNA 100 µg/mL Salmon sperm DNA 10 mM Tris-HCl, pH 7.2 Prepared in Maleic acid buffer	Telomere-FISH staining
Maleic acid buffer	100 mM Maleic acid 150 mM NaCl, pH 7.5	To prepare telomere-FISH hybridization buffer
Telomere-FISH wash buffer	2X SSC 0.1% Tween-20	Telomere-FISH wash buffer

Commercial kits

Kit	Supplier	Catalogue number
AllPrep DNA/RNA/Protein Mini Kit	Qiagen	80004
Direct-zol RNA Microprep Kit	Zymo	R2062
DNeasy Blood and Tissue Kit	Qiagen	69504
First Strand cDNA Synthesis Kit	Thermo Scientific	K1612
GenElute™ Plasmid Miniprep Kit	Sigma-Aldrich	PLN350
HiScribe™ SP6 RNA Synthesis Kit	NEB	E2070
Monarch® DNA Gel Extraction Kit	NEB	T1020
Nucleofector® Kit V	Lonza	VCA-1003
NucleoSpin Gel and PCR Clean-up	MACHEREY-NAGEL	740609.50
pCR™8/GW/TOPO™ TA Cloning Kit	Invitrogen	K250020
ProtoScript® II First Strand cDNA Synthesis Kit	NEB	E6560
SulfoLink™ Immobilization Kit for Peptides	Thermo Scientific	44999
TruSeq Stranded mRNA LT Sample Prep Kit	Illumina	
QIAGEN Plasmid Midi Kit	Qiagen	12143
QIAGEN Plasmid Mega Kit	Qiagen	12181
Qubit dsDNA HS Assay Kit	Invitrogen	Q32851
RNeasy Mini Kit	Qiagen	74104

Instruments

Instruments	Supplier
2100 Bioanalyzer	Agilent Technologies
Aurora Ultimate CSI 25×75 C18 UHPLC column	Ionopticks
BC43 Benchtop Microscope	Andor
Bioruptor Plus UCD-300I	Diagenode
Easy-nLC 1200	Thermo Fisher Scientific
Exploris 480	Thermo Fisher Scientific
NanoElute2 HPLC system	Bruker
NextSeq 500 High Output FC	Illumina
Optima XE-100 Ultracentrifuge	Beckman-Coulter
QuantStudio™ 5 Real-Time PCR machine	Applied Biosystems
Qubit 2.0 Fluorometer	Life Technologies
QExactive Plus mass spectrometer	Thermo Fisher Scientific
Sonifier® Cell Disrupters	Branson
TimsTOF HT mass spectrometer	Bruker
VisiScope Confocal Microscope	Visitron

Softwares

Software	Distributor
Adobe Illustrator 2024	Adobe
ApE plasmid editor	
Image J	Image J
Image Lab	BioRad
MaxQuant (version 1.6.10.43 and 2.4.2.0)	MaxQuant
QuantStudio™ Design and Analysis Desktop Software	Thermo Fisher Scientific
R	The R foundation
R-studio	R Studio Inc

Methods

Orthologous comparison

Nucleotide sequences were aligned in A plasmid Editor software (ApE). The genomic sequences were obtained from NCBI with the following gene IDs: *hmbox1a* (571508), *hmbox1b* (100317353), and *zbtb48* (567773). The transcript ID were as follows: *hmbox1a*-Main (NM_001030264.1), *hmbox1a*-X1 (XM_005169832.4), *hmbox1a*-X2 (XM_009292894.3), *hmbox1a*-X3 (XM_021467220.1), *hmbox1a*-X4 (XM_021467221.1), *hmbox1a*-X5 (XM_005169831.4), *hmbox1b*-X1 (XM_005173819.4), *hmbox1b*-X2 (XM_009301469.3), *hmbox1b*-X3 (XM_009301470.3), and *zbtb48* (XM_691074).

Amino acid sequences and domain annotations of proteins were obtained from UniProt and the sequences were aligned using UniProt alignment tool. The percentage of domain and ortholog identity was obtained through NCBI BLAST search. The UniProt accession numbers for the proteins used in the amino acid sequence alignments are as follows: human HMBOX1 (Q6NT76) and zebrafish Hmbox1a (Q4V904) and Hmbox1b (A0A8M2BL53), human ZBTB48 (P10074), zebrafish Zbtb48 (E7FDZ5), human MTFP1 (Q9UDX5), and zebrafish Mtfp1 (Q6PCS6).

Methods for *E. coli*

Plasmid transformation

Respective *E. coli* competent cells (DH5 α , Stable, or BL21) were thawed on ice and mixed with 250 - 1000 ng plasmid for 30 min on ice. Cells were heat shocked at 42 °C for 45 s and immediately rested on ice for 2 min. Cells were then diluted with 250 μ L of SOC medium and incubated at 37 °C for 1 h, with shaking at 800 rpm on thermomixer. Afterward, the cells were plated on LB agar plates with respective antibiotics and incubated overnight at 37 °C.

Bacterial culturing

Individual colonies were selected from the plate or inoculated from glycerol stock and cultured overnight in LB medium containing the respective antibiotics at 37 °C, shaking at 180 rpm. For cultures larger than 500 mL, a smaller volume of starter culture was prepared and then subcultured for the overnight growth. The antibiotic concentration used for culture was 50 μ g/ml for kanamycin or 100 μ g/ml for ampicillin or spectinomycin.

Plasmid purification

Cultures were centrifuged at max speed (or, 4000 xg) for 15min at 4 °C. Depending on the culture volume and plasmid copy number, plasmids were purified using different kits like GenElute Plasmid Miniprep Kit (Sigma-Aldrich), Plasmid Midi Kit (Qiagen), or Plasmid Megaprep kit (Qiagen), as directed by the manufacturers.

Glycerol stock preparation

For long-term storage of bacterial possessing plasmids of interest, 300 μ L of the overnight culture was added to 100 μ L of 80% glycerol and mixed homogeneously, bringing the final glycerol concentration to 20%. Stored at -80 °C.

Cloning

Cloning gene of interest into pCR8/GW/TOPO holding vector

RNA was isolated from BRF41 cells using RNeasy Mini Kit (Qiagen) and reverse transcribed into cDNA with First Strand cDNA Synthesis Kit (Thermo Scientific). Full-length cDNAs of interest were amplified using Q5 High-Fidelity DNA Polymerase (NEB). The PCR was performed at an annealing temperature of 58 °C for *hmbox1a* (primer #4211 and #4212, or #4283) and *zbtb48* (primer #4220 and #4221), and 64 °C for *hmbox1b* (primer #4214 and #4215), with an elongation time of 90 s. The amplicons were cloned into holding vector pCR8 using the pCR8/GW/TOPO TA Cloning Kit (Invitrogen) as directed by the manufacturer. The plasmids were transformed into DH5 α competent cells (NEB) and selected using spectinomycin. Six colonies were picked and sequenced with primers #404 and #405. Those that possessed the correct inserts and orientation were cultured and plasmids were extracted using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich).

Cloning epitope sequence into pCoofy4 vector

The epitopes used to raise the antibodies were from a small fragment of each protein rather than their full-length forms. These sequences were amplified from the respective holding vectors using Q5 High-Fidelity DNA Polymerase (NEB) and primers with added 5' overhangs: Forward - AAGTTCTGTTCCAGGGGCC, and Reverse - CCCCCAGAACATCAGGTTAATGGCG; Primer #4223 and #4224 for Hmbox1a, #4286 and #4287 for Hmbox1b, #4229 and #4719 for Zbtb48-BTB, and #4231 and #4232 for Zbtb48-ZNF. They were then cloned into the expression vector pCoofy4 (N-terminal His₆-MBP) following the Sequence and Ligation Independent Cloning (SLIC) method²¹¹. All amplifications were conducted at an annealing temperature of 64 °C with a 90 s elongation time. The amplicons were cloned into a linearized pCoofy4 vector at a 3:1 weight ratio, utilizing RecA enzyme (NEB) at a 1:1000 dilution. The plasmids were then transformed into DH5 α competent cells (NEB), except for Zbtb48-BTB in TOP10 (Invitrogen), and selected with kanamycin. Colonies were picked, and their plasmids were sequenced using primer #1421. Those containing the correct inserts were cultured, and plasmid DNA was extracted using the Plasmid Midi Kit (Qiagen).

Cloning gRNA into PX459

Guide RNAs (gRNAs) were designed using CRISPRscan software (<https://www.crisprscan.org>). Two gRNAs were selected for each gene to target the most upstream coding exon(s), refer to Table M1 for the oligo number and sequence. The cloning method was adapted from Ran et al.²¹². Firstly, pSpCas9(BB)-2A-Puro (PX459) vector was linearised by BbSI enzyme (NEB) and purified using Monarch® DNA Gel Extraction Kit (NEB). Next, the complementary strands of the synthesized gRNA oligos were annealed, and then phosphorylated by T4 PNK (ThermoFisher). Lastly, the phosphorylated gRNA oligos were ligated to the digested-PX459 vector using T4 ligase (NEB). The plasmids were transformed into DH5 α competent cells (NEB) and selected using ampicillin. Colonies were picked and their plasmids were sequenced with primer #2191. Those that possessed the correct inserts were cultured and the plasmids were extracted using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich).

Cloning gRNA into plentiCRISPRV2-Puro

The cloning follows the method from “Cloning gRNA into PX459” with the below-mentioned adjustments. plentiCRISPRV2-Puro vector was linearized by Bsmbl-V2 enzyme (NEB) and the ligated plasmids were transformed into Stable Competent E. coli (NEB). The plasmids were extracted using the Plasmid Midi Kit (Qiagen).

Cloning gRNA into pDR274

The cloning follows the method from “Cloning gRNA into PX459” with the below-mentioned adjustments. Due to the differences in the overhang sequence of pDR274, new oligos were synthesized with compatible overhangs whilst the gRNA target sequences remained the same, refer to Table M1 for the oligo number and sequence. pDR274 vector was linearized by BsaI enzyme (NEB). Colonies were selected using kanamycin and plasmids were sequenced with primer #3725.

Table M1. Details of *hmbox1a/b* and *zbtb48* CRISPR-Cas9 plasmids.

Primer no.	Oligo name	Sequence 5' to 3'	PX459		pLentiCRISPRV2-Puro	
			Plasmid no.	Glycerol no.	Plasmid no.	Glycerol no.
4819	<i>hmbox1a</i> , gRNA1- Top	CACCg TTT GGC CAT CGG CAT GCT CA	P1247	G534	P1279	G544
4820	<i>hmbox1a</i> , gRNA1- Bottom	aaac TG AGC ATG CCG ATG GCC AAA c				
4821	<i>hmbox1a</i> , gRNA2- Top	CACCg GTG CTA CAA CCG GCA TAG	P1248	G535	P1280	G545
4822	<i>hmbox1a</i> , gRNA2- Bottom	aaac CTA TGC CGG TTG TAG CAC c				
4823	<i>hmbox1b</i> , gRNA1- Top	CACCg TCG AGG GCG TGC ACG ATC T	P1249	G536	P1281	G546
4824	<i>hmbox1b</i> , gRNA1R- Bottom	aaac A GAT CGT GCA CGC CCT CGA				
4825	<i>hmbox1b</i> , gRNA2- Top	CACCg AAC TGC GTC TGA GTT GCT G	P1250	G537	P1282	G547
4826	<i>hmbox1b</i> , gRNA2- Bottom	aaac C AGC AAC TCA GAC GCA GTT				

Primer no.	Oligo name	Sequence 5' to 3'	pDR274	
			Plasmid no.	Glycerol no.
4732	<i>hmbox1a</i> , gRNA1- Top	tagg TTT GGC CAT CGG CAT GCT CA	P1237	G524
4733	<i>hmbox1a</i> , gRNA1- Bottom	aaac TG AGC ATG CCG ATG GCC AAA		
4734	<i>hmbox1a</i> , gRNA2- Top	tagg GTG CTA CAA CCG GCA TAG	P1238	G525
4735	<i>hmbox1a</i> , gRNA2- Bottom	aaac CTA TGC CGG TTG TAG CAC		
4736	<i>hmbox1b</i> , gRNA1- Top	tagG TCG AGG GCG TGC ACG ATC T	P1239	G526
4737	<i>hmbox1b</i> , gRNA1R- Bottom	aaac A GAT CGT GCA CGC CCT CGA		
4738	<i>hmbox1b</i> , gRNA2- Top	tagG AAC TGC GTC TGA GTT GCT G	P1240	G527
4739	<i>hmbox1b</i> , gRNA2- Bottom	aaac C AGC AAC TCA GAC GCA GTT		
4744	<i>zbtb48</i> , gRNA1- Top	tagG CAG GCC AGG ACA TTG CGA T	P1243	G530
4745	<i>zbtb48</i> , gRNA1R- Bottom	aaac A TCG CAA TGT CCT GGC CTG		
4746	<i>zbtb48</i> , gRNA2- Top	agg CAT CGG GAC ACT CTA AAG GG	P1244	G531
4747	<i>zbtb48</i> , gRNA2- Bottom	aaac CC CTT TAG AGT GTC CCG ATG		

Polymerase chain reaction (PCR)

PCR amplification

For cDNA amplification, either Q5 (NEB) or in-house HF polymerase was used to amplify cDNA. Reactions and amplifications were set up as directed by the manufacturer. A 50 µL reaction composed of 1X reaction buffer, 200 µM dNTPs, 0.5 µM of each primer, < 1 µg of cDNA, and 0.5 µL of polymerase.

For genotyping, In-house Taq polymerase was used for genotyping and reactions were set up in 10 µL reaction volume. Each reaction comprised of 1X reaction buffer, 250 µM dNTPs, 0.1 µM of each primer, 0.05 µg BSA, 0.5 µL to 1 µL of DNA lysate, and 0.5 µL of Taq polymerase. Universal amplification protocol was used for all genotyping and the setting on the thermocycler was as follows: denaturing at 95 °C for 5 min, a touch-down format of amplification with 5 cycles of 95 °C for 15 s, 64 °C for 30 s, 68 °C for 30 s, followed by another 5 cycles of 95 °C

for 15 s, 60 °C for 30 s, 68 °C for 30 s, and lastly 27 cycles of 95 °C for 15 s, 55 °C for 30 s, 68 °C for 30 s, and final extension at 68 °C for 5 min.

DNA electrophoresis

2% agarose was used for genotyping (product size < 1kb) while 0.8% to 1.5% agarose was used for cDNA amplification, depending on the product size. Agarose powder was dissolved in TBE or TAE buffer by heating through microwaving and SYBR Safe (Invitrogen) was added to the gel before casting. PCR products were mixed with loading dye before loading it on the agarose for DNA gel electrophoresis. Electrophoresis was performed at 120V for 25 min, or up to 1 h depending on the gel size and product size to resolve.

PCR product purification

To precipitate DNA using ethanol, 500 µL of 100% ethanol was added to 25-30 µL of PCR product, and 0.5 µL of GlycoBlue (Ambion). Then, incubated at -80 °C for 10 min before centrifuging at max speed for 1 h, at 4 °C. Remove ethanol and air dry the pellet, followed by resuspending it in an appropriate volume of nuclease-free water.

To purify PCR products from the agarose gel, band of the desired base pair was excised under the UV-light. DNA gel purification kits like Monarch® DNA Gel Extraction Kit (NEB) or NucleoSpin Gel and PCR Clean-up (MACHEREY-NAGEL) were used as directed by the manufacturer.

DNA or plasmid sequencing

400 µg of DNA, or 7.5 µL of DNA (if < 400 µg DNA available), supplemented with 2.5 µL of 10 µM sequencing primer was sent to GATC Biotech for sequencing.

Generating antibodies

Protein production

Cloned pCoofy4 plasmids were transformed into protein-expressing BL21 competent cells (NEB) and selected using kanamycin. For each epitope, the BL21 was cultured in 2 L of LB medium with kanamycin, with volume split equally into four 2 L flasks. For epitopes that have low yield like His₆-MBP-Hmbox1b and His₆-MBP-Zbtb48-BTB, LB medium used was supplemented with 3% ethanol to enhance the protein expression¹³⁷. The cultures were incubated at 37 °C, 180 rpm until the OD₆₀₀ of 0.6-0.8 was reached. When the optimum OD₆₀₀ value was reached, the cultures were cooled on ice before inducing with 1 mM IPTG (Carl Roth). The cultures were incubated overnight at 18 °C, with 180 rpm shaking for protein expression.

Protein extraction and purification

To maintain the integrity of the expressed protein, all the steps were performed either on ice or at 4 °C. The cultures were pelleted by centrifuging at 4000 xg for 30 min at 4 °C. BL21 were lysed on ice in bacteria lysis buffer using sonication (Branson sonifier) with 9mm tip at 30% duty cycle, output 4 for 3 min, twice. The lysates were transferred to the appropriate tubes for ultracentrifugation. Lysates were ultracentrifuged in Beckman-Coulter OPTIMA XE-100 using a 70Ti rotor at 27, 000 rpm for 30 min at 4 °C. The supernatant was collected and proceeded with affinity purification.

Each epitope was purified using a fresh HisTrap HP column (GE Healthcare). The column was equilibrated with column binding buffer prior to passing the supernatant through it. The column was washed with column binding buffer followed by column wash buffer. The epitope was subsequently eluted from the column using column elution buffer, the eluant was collected in fractions, at approximately 1 mL per tube. The column was prepared for storage by cleaning it with 1M imidazole solution (Sigma), followed by MQ water, and finally with 20% ethanol. Throughout the process, the flow rate for all buffers entering the column was maintained at 1 mL/min, except for the supernatant and elution buffer, which were set at 0.5 mL/min.

All eluant fractions were run on SDS-PAGE gel to check for the quality and quantity of the protein. Dialysis was performed on four of the best elution fractions using the benzoylated dialysis tubing (Sigma) in the dialysis buffer to remove salts.

Epitope immunization and antibody purification

Between 400 µg and 1.6 mg of each epitope was sent to Pineda for immunization. Three rabbits were immunized per protein, and sera were collected 61 days later. The sera were then purified using the SulfoLink™ Immobilization Kit for Peptides (ThermoFisher) in a 5 mL polypropylene column (Thermo Scientific) following the manufacturer's instructions. The columns were coupled with 800 µg of the respective epitopes, and 5 mL of sera were purified each time. Antibodies were collected across 10 elution fractions (1 mL per fraction). The flow-throughs and elutions were also collected and analyzed on SDS-PAGE.

Methods for cell culture

Cell culture and maintenance

Zebrafish-derived fibroblast cell line BRF41 (CVCL_4131) and PAC2 (CVCL_5853) were cultured in Leibovitz's L-15 medium (Gibco), supplemented with 15% fetal bovine serum (Gibco) and 1% penicillin-streptomycin (Sigma). Cells were maintained at 33 °C for BRF41 whilst 28 °C for PAC2 without carbon dioxide supply and passaged through trypsinization every three to four days.

HEK293T were cultured in Dulbecco's Modified Eagle Medium (DMEM) medium (Gibco), supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin-streptomycin (Sigma). Cells were maintained at 37 °C with 5% CO₂ and passaged through trypsinization every three to four days.

Transfection with FuGENE® HD Transfection Reagent (Promega)

PAC2 cells were seeded in a 6-well plate at a density of 4.5×10^5 cells per well the night prior to transfection. The medium in each well was replaced with 1800 µL of L-15 medium without FBS and antibiotic supplements, and 200 µL of the transfection mixture was added dropwise to the cells. The transfection mixture was prepared by combining FuGENE® HD Transfection Reagent (Promega) with either 2 or 4 µg of the empty pCS2-eGFP plasmid at a ratio ranging from 2:1 to 4:1, then incubated at room temperature for 20 min before adding it to the cells¹³⁸. The cells were incubated in the transfection mixture for 24 h, after which the medium was replaced with L-15 medium supplemented with antibiotics and FBS.

Transfection with Nucleofector® solution V (Lonza)

As described by He et al. ¹³⁹, 2×10^6 PAC2 cells were resuspended in 100 μ l of Nucleofector® solution V containing 5 μ g of total plasmids. The cells were either transfected with 5 μ g of empty PX459 plasmid or with 2.5 μ g of each gRNA-insert PX459 plasmids (P1247 and P1248 for *hmbox1a*, or P1249 and P1250 for *hmbox1b*). The cell suspension was transferred into a kit-provided cuvette and placed into the Lonza™ Nucleofector™ Transfection 2b Device. The cells were pulsed with the pre-programmed nucleofection settings T027 or X001. Afterwards, the cells were transferred into a 10 cm dish containing L-15 medium supplemented with 15% FBS and the medium was refreshed at 36 hours post-transfection (hpt). For selection, 0.5 μ g/mL of puromycin was added to the cells at 3 days post-transfection (dpt).

Lentivirus production

Five plates of 3×10^6 HEK293T cells were seeded in 10 cm dishes the day before transfection for each *hmbox1b* gRNA for lentivirus production. The transfection mixture was prepared as follows: 25 μ g of each packaging plasmid (pMDLg/pRRE, pRSV-Rev, and pMD2.G) was combined with 50 μ g of either the empty lentiCRISPR-V2 plasmid or each of the *hmbox1b* gRNA-insert (P1281 or P1282). To this mixture, 275 μ L of 2M CaCl₂ was added and the volume was topped up to 2250 μ L with water. This plasmid mixture was then combined with 2250 μ L of 2X HBS (containing 280 mM NaCl, 10 mM KCl, 1.5 mM Na₂HPO₄, 12 mM D-glucose, and 20 mM HEPES) and incubated at room temperature for 5 min. To each plate of HEK293T cells containing 7 mL of freshly replaced DMEM, 900 μ L of the transfection mixture was added in a dropwise manner. The cells were incubated at 37 °C with 5% CO₂ for 8 to 10 h before refreshing the medium with 7 mL of DMEM medium supplemented with 10% FBS. At 48 hpt, the medium was collected and centrifuged at 300 xg for 5 min at 4 °C. The supernatant was filtered through a 0.45 μ m filter before an overnight centrifugation at 4000 rpm at 4 °C. The supernatant from the overnight centrifugation was discarded while the virus pellet was resuspended in DMEM at 1:100 of the initial volume and stored at -80 °C.

Plaque assay

To measure the virus titer, the produced lentivirus was diluted in a series of ten-fold dilutions, ranging from 10⁻⁴ to 10⁻⁸, using DMEM supplemented with 5 μ g/mL of polybrene (Santa Cruz). The diluted virus was then added to 1×10^5 HEK293T cells per well in a 6-well plate, seeded the day before. The medium was replaced with DMEM supplemented with 10% FBS and 2 μ g/mL of puromycin for selection at 2 dpt, and then reduced to 1 μ g/mL of puromycin at 5 dpt. The plaque assay was performed on 17 dpt. The wells were washed once with PBS before fixing the cells with 2 mL of a 3:1 mixture of methanol to acetic acid for 10 min. The cells were then stained with 1.5% crystal violet in ethanol for 20 min, followed by washing the wells with water. The number of colonies were counted, and the virus titer was represented in Infectious Unit (IU)/mL.

Lentivirus transduction

Each well of 1×10^5 PAC2 cells, seeded in a 6-well plate the night before, was transduced with 2 ml of L-15 medium containing viruses diluted at 1:100 and 5 μ g/ml of polybrene. The medium was replaced with L-15 supplemented with 10% FBS at 2 dpt. The transduced cells were not subjected to puromycin selection.

T7 Endonuclease I assay

PAC2 cells were trypsinized from wells, and the genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's instructions. The target genes

were amplified using HF-DNA polymerase (in-house) with an annealing temperature of 63 °C. Primers used for amplifying *hmbox1a* were #4211 and #4767, while primers for *hmbox1b* were #4214 and #4774. The amplified DNA was purified using the ethanol precipitation method and the DNA pellet resuspended in 30 µL of water.

A T7 Endonuclease I (T7E1) assay was performed on these purified DNA samples according to the manufacturer's instructions (NEB). The amplicons were melted at 95 °C for 5 min before allowing the DNA to slowly re-hybridize with a gradual temperature reduction (from 95 °C to 85 °C at -2 °C/s, followed by 85 °C to 25 °C at -0.1 °C/s). As a non-T7E1 treatment control, 5 µL of the hybridized amplicons were withdrawn prior to T7E1 treatment. The remaining 15 µL were treated with 0.5 µL of T7E1 and incubated at 37 °C for 40 min. After incubation, 5 µL of the T7E1-treated DNA and 5 µL of the untreated control were loaded onto a 1.5% agarose gel for analysis.

Methods for zebrafish

Zebrafish husbandry

Zebrafish were housed at the Institute of Molecular Biology in Mainz and husbandry was carried out following the standards described ²¹³. Room and water temperature were maintained at 26-28 °C with a 14:10 hours light: dark cycle. Adult fish were mated in the morning after an overnight male and female separation. Embryos were kept in a 10-cm Petri dish filled with E3 medium and raised in a 28.5 °C incubator until 5 days post fertilization (dpf). Afterward, they were transferred to the aquarium facility. All experiments were approved and licensed by the local authorities of Rhineland-Palatinate and were conducted in accordance with European animal welfare law. The zebrafish strain used in this study was AB/Tübingen and transgenic lines used include *Tg(vasa:eGFP)* ¹⁵⁶ and *Tg(wt1b:eGFP)li1* ¹⁴².

Creating CRISPR-Cas9 knockout zebrafish line

The gRNA transcripts were produced as previously described ²¹⁴. Briefly, the pDR274 plasmids with gRNA insert were linearized with Dral enzyme (NEB). As the pDR274 vector used in this experiment harbours an SP6 promoter, the gRNAs were transcribed with HiScribe™ SP6 RNA Synthesis Kit (NEB). The transcription products were treated with DNaseI (NEB) and purified using the ethanol precipitation method ²¹⁵. The gRNA pellets were resuspended in 5 µL of nuclease-free water and 0.25 µL was loaded on 2% agarose gel for quality and quantity check.

Each CRISPR knockout line was generated separately. Cas9-injection mix was prepared as follows: 0.5 µL NEB 3.1 buffer, 3 µL of both gRNA transcripts, 0.5 µL phenol red (Gibco), and 1 µL EnGen® Spy Cas9 NLS (NEB). The mixture was injected into embryos at one-cell stage (F0) and were raised to adulthood. They were out-crossed to AB/Tübingen wild-type fish, and the offspring (F1) were screened for deletion through genotyping. Those that possessed desired deletion were selected. For *hmbox1a/b*, *hmbox1a*^{+/-} mutants were crossed with *hmbox1b*^{+/-} mutants to create *hmbox1a*^{+/-}; *hmbox1b*^{+/-} F2 mutants. For *zbtb48*, the *zbtb48*^{+/-} were out-crossed with wild-type. The F2 heterozygotes were then in-crossed to create F3 generation of mutants, or first homozygous mutant generation (G1).

Fin clipping and genotyping

Zebrafish were anesthetized with 0.02% tricaine (Sigma) for fin-clipping. Upon clipping the caudal tail fins, the fish were put into individual tanks filled with fresh system water for recovery.

For those fish to be used for organ harvesting, they were sacrificed in an ice water bath instead. Similarly, embryos before 5 dpf were also sacrificed with an ice water bath, and the whole embryo was used for genotyping. All samples were lysed in 20 µl of lysis buffer at 60 °C for 60 min, followed by 90 °C for 15 min for proteinase K inactivation. Lysates were diluted with 30 µl of water and PCR was conducted as mentioned under “PCR methods”. *hmbox1a^{LD}* was genotyped with #4835, #4820 and #4767, while *hmbox1a^{SD}* with #4211 and #4767. *hmbox1b^{LD}* was genotyped with #4214, #4768 and #4976, while *hmbox1b^{SD}* with #4214 and #4774. *zbtb48* mutants were genotyped with #4220 and #4719. Fish were then sorted based on genotyping results.

Protein extraction

Sample collection

BRF41 and PAC2 cells were harvested by trypsinization and neutralized with L-15 medium (with FBS). Then, pelleted by centrifugation at 400 xg for 5 min and washed with PBS (Gibco). U2OS protein lysate was obtained from a colleague in the lab.

The chorions of zebrafish embryos younger than 3 dpf were removed by incubating them in a 1 mg/mL solution of pronase for 10 to 30 min with gentle agitation. After decanting the debris, the embryos were rinsed with cold calcium-free Ringer’s solution. De-yolking of embryo and larvae was carried out using a P200 pipette tip in cold Ringer’s solution supplemented with 0.3 mM PMSF and 1 mM EDTA. The solution was then removed, and the embryos were washed twice with cold Ringer’s solution before being kept on ice.

Juvenile and adult zebrafish were sacrificed in an ice bath and dissected under a stereomicroscope to harvest the respective organs. For transgenic lines, dissections were performed under a fluorescence stereomicroscope. The harvested organs were immediately transferred into a tube and kept on ice.

Whole-cell (total) protein lysate with RIPA buffer

Samples were lysed in RIPA buffer for downstream applications such as Western blotting or mass spectrometry proteome measurement. Cell pellets were resuspended with 50 µL of RIPA buffer and vortexed thoroughly. Embryos and organs were snap-freeze in liquid nitrogen and grinded with a micro-pestle before adding 50 µL of RIPA lysis buffer. For smaller organs, like 40 dpf gonads, they were sonicated in 25 µL of RIPA using a Bioruptor Plus UCD-300I at a low setting with three cycles of 30 s on and 30 s off, or until the samples were disintegrated, at 4 °C. All lysates were incubated on ice for at least 30 min before centrifuging at maximum speed for 30 min at 4 °C to collect the supernatant.

Whole-cell (total) protein lysate with IP buffer

Samples were lysed in IP buffer for downstream applications like telomere pull-down or immunoprecipitation (IP). Sonication was performed using Bioruptor Plus UCD-300I at a low setting with three cycles of 30 s on and 30 s off, or until the samples were disintegrated, at 4 °C. The lysate was centrifuged at maximum speed for 30 min, at 4 °C and the supernatant was collected.

Nuclear extraction

To enrich for nuclear proteins, the cell pellets were resuspended in 5 volumes of cold nuclear extraction buffer A, or 50 embryos in 1 mL of cold nuclear extraction buffer A, and incubated

on ice for 30 min. They were transferred to a pre-cooled glass douncer and sheared with either 50 strokes of tight-fitting pestle B for cells, or 20 strokes with loose-fitting pestle A followed by 60 strokes with a tight-fitting pestle B for embryos. Samples were rested on ice after every 20 strokes for 5 min. The lysates were centrifuged at 2000 xg for 10 min at 4 °C. The pellets were washed by resuspending in cold 1X PBS and centrifuging at 2000 xg for 5 min at 4 °C. After washing, the pellets were resuspended in cold Buffer C and rotated at 4 °C for 1h. The lysates were centrifuged at maximum speed for 1h at 4 °C. The supernatant was collected as the nuclear fraction.

For the second telomere pull-down experiments with 5 dpf nuclear extract, an additional step of passing the lysate through a 20 µM cell strainer was performed before the centrifugation to separate the nucleus from the cytoplasmic extract, and an extra round of cold 1X PBS wash to the nuclear pellet was added.

Protein concentration measurement

Protein concentration was measured using Bradford assay (Bio-rad) and prepared into the desired concentration for downstream applications. The protein extracts were kept at 4 °C for immediate or within the week use. For long-term storage, they were supplemented with 10% glycerol and snap-freeze for -80 °C storage.

Protein analysis

Western blotting

Approximately 50 µg of protein lysates were prepared in 1X NuPAGE™ LDS Sample Buffer (Thermo Scientific) and 0.1 M DTT. The lysates were boiled at 70 °C for 10 min with shaking at 500 rpm. Proteins were resolved in a 4-12% NuPAGE Bis-Tris precast gel (Thermo Scientific) using MES running buffer at 180V for 1 h. The proteins were transferred to a nitrocellulose membrane (Cytiva) using transfer buffer (IMB CF), supplemented with 20% ethanol, at 300 mA for 90 min.

The membrane was stained with Ponceau S (Applichem) to confirm protein transfer and imaged. The Ponceau stain was then washed off with PBST (PBS containing 0.05% Tween-20) before blocking in 5% skim milk for 1 h at room temperature with shaking. Then, incubated with primary antibodies overnight at 4 °C with shaking. The membrane was washed three times with PBST for 10 min before incubating with secondary antibodies at room temperature for 1 h. The membrane was then washed again for three more times with PBST for 10 min before developing with chemiluminescent reagents (Thermo Fisher or Amersham) and imaged using ChemiDoc (Bio-Rad).

The primary and secondary antibodies used in this study were prepared in 5 mL of 5% skim milk. The dilution of primary antibodies were used as follows: rabbit antibodies raised in this study, such as Hmbox1a (T1 and T3), Hmbox1b (T2), Zbtb48-BTB (T2 and T3), and Zbtb48-ZNF (T2 and T3) antibodies, were used at a dilution of 1:200 of elution 3, whereas Hmbox1b-T2 (M) and Zbtb48-BTB-T3 (M) antibodies, purified by the protein production core facility, was used at 1:1000. Mouse anti-GFP antibody (Roche), rabbit anti-human MTFP1 antibody (Sigma), and rabbit anti-pan Histone 3 antibody (Diagenode) were used at 1:1000. Rabbit anti-Cas9 (Diagenode) was used at 1:5000, and rabbit anti-actin (Sigma) was used at 1:8000. For secondary antibodies, anti-mouse IgG, HRP-linked antibody (Cell Signaling Technology) was

used at 1:10,000, while anti-rabbit IgG, HRP-linked antibody (GE Healthcare) was used at 1:2500.

Telomere pull down

Telomere pull-down was performed as previously described⁷⁹. The oligonucleotide baits were formed by chemically synthesizing DNA oligonucleotides containing either the telomeric TTAGGG sequence or a shuffled control sequence, which were then concatenated through *in vitro* ligation. These concatenated telomeric and control oligonucleotides were biotinylated *in vitro* and immobilized on Dynabeads MyOne C1 streptavidin beads (Thermo Scientific). The experiments were conducted in quadruplicate, with each replicate involving 450 µg of beads incubated with 300-750 µg of protein lysate. Both telomere (TTAGGG) and scrambled (GTGAGT; control) baits were employed in the BRF41 and zebrafish embryo pull-downs, while only telomere bait was used in the wild-type versus *zbtb48*^{-/-} mutant larvae pull-downs. Samples were incubated on a rotating wheel at 4 °C for 1 h, followed by three washes with PBB buffer. Samples were eluted from the beads with 20 µL of 0.1 M DTT and 1X NuPAGE™ LDS Sample Buffer (Thermo Scientific) by boiling at 70 °C for 10 min.

Immunoprecipitation

All experiments were conducted with four replicates. For each replicate, 750 µg of Protein G beads (Invitrogen) was coupled with antibody on a rotating wheel for 1 h at 4 °C. Antibodies were used at either a 1:10 dilution of rabbit antibodies raised in this study, such as Hmbox1a (T1 and T3), Hmbox1b (T2), Zbtb48-BTB (T2 and T3), and Zbtb48-ZNF (T2 and T3) antibodies, or with 5 µg of rabbit-IgG control (Sigma), Hmbox1b-T2 (M), and Zbtb48-BTB-T3 (M) purified by the protein production core facility. After coupling, the beads were washed three times with PBB buffer. The coupled beads were incubated with 200-800 µg of protein lysates on a rotating wheel for 2 h at 4 °C, followed by three washes with PBB buffer. Samples were eluted from the beads with 20 µL of 0.1 M DTT and 1X NuPAGE™ LDS Sample Buffer (Thermo Scientific) by boiling at 70 °C for 10 min.

Mass spectrometry sample preparations

Samples eluted from telomere pull-down and immunoprecipitation were loaded onto 10% NuPAGE Bis-Tris precast gel (Thermo Scientific) and resolved at 180V for 8 min in MES running buffer (Life Technology). In-gel digestion was performed as previously described⁷⁹. Briefly, the gels were stained with Coomassie blue and washed overnight in distilled water with shaking. The next day, the protein bands were excised from the gel, cut into small pieces, and destained in a 96-well filter plate with 50% ethanol. Gel pieces were dehydrated using acetonitrile (VWR) before reducing the samples with 10 mM DTT (Sigma) for 1 h at 56 °C, followed by alkylation with 55 mM iodoacetamide (Sigma) for 45 min in the dark. Gel pieces were dehydrated again using acetonitrile (VWR) before tryptic digestion with 1 µg of MS-grade trypsin (Sigma) in 50 mM ammonium bicarbonate buffer (pH 8.0) at 37 °C overnight. Afterwards, the peptides were extracted using 30% acetonitrile and stored on activated C18 material (Empore).

MS measurement and analysis (telomere pull-down and immunoprecipitation)

Measurements of telomere pull-down and immunoprecipitation were conducted on either Exploris 480 (Thermo Fisher Scientific) or Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) coupled online to an Easy-nLC 1200 (Thermo Fisher Scientific) on an in-house

packed C18 (Dr. Maisch) column. The data were acquired via data-dependent acquisition (DDA) with a top15 method. The data were analyzed with MaxQuant ²¹⁶ version 1.6.10.43 with standard settings except label-free quantification (LFQ) and the match between run option were activated. UniProt zebrafish database (39,559 entries), was supplemented with the MBP and epitope sequences used in this study. Protein quantification was performed as previously described ²¹⁷. In brief, contaminants, reverse database hits, protein groups identified only by site, and protein groups with fewer than two peptides (with at least one classified as unique) were filtered out from the MaxQuant proteinGroups.txt file. Missing values were imputed by shifting a beta distribution, based on the LFQ intensity values, to the limit of quantitation. Volcano plots, heatmap, and Venn diagram were generated from Rstudio using packages including ggplot2 and more. The threshold for protein enrichment was set at a fold change > |2| and a p-value < 0.05 (Welch's t test). The peptide sequences detected by the mass spectrometer were obtained from the peptide.txt file.

MS measurement and analysis (proteome)

For proteome analysis, 4 hpf embryos of various *hmbox1a/b* genotypes (wild-type, (*hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD}) double mutant, *hmbox1a*^{LD/LD} and *hmbox1b*^{LD/LD} single mutants) and 5 dpf of wild-type and *zbtb48*^{-/-} mutants were respectively pooled and lysed in RIPA. Proteome analyses were performed in technical quadruplicate, utilizing 50 µg of protein per replicate, with two independent collections of biological replicates for *zbtb48* study. The proteins were loaded onto a 10% NuPAGE Bis-Tris precast gel (Thermo Scientific) and resolved at 180 V for 30 min using MES running buffer (Life Technology). The gels were subsequently divided into three parts at the 41 kDa and 93 kDa ladder bands for in-gel digestion, following the procedure mentioned above. The proteins were measured on Exploris 480 mass spectrometer (Thermo Fisher Scientific) coupled online to an in-house packed C18 column (Dr. Maisch GmbH).

For the proteome analysis involving juvenile and adult gonads in the *zbtb48* study, three to four biological replicates were conducted, each consisting of individual fish. Sample preparation for mass spectrometry was conducted by research assistant Rachel Mullner, with 400 ng of protein used for each replicate. Proteomes were analyzed using a TimsTOF HT mass spectrometer (Bruker), interfaced with a NanoElute2 HPLC system (Bruker). An Aurora Ultimate CSI 25×75 C18 UHPLC column (Ionopticks) was utilized at 50 °C.

The peptides were eluted along a 95-min gradient from 4% to 32% acetonitrile (VWR) and measured with a DDA acquisition scheme. The measurement files were processed with MaxQuant (2.4.2.0) and a combined Swissprot/Trembl database (47,082 entries) with the standard settings except for peptides for quantification, which was switched to unique and LFQ quantitation activated. Protein analysis was performed as previously described ²¹⁷. In brief, contaminants, reverse database hits, protein groups identified only by site, and protein groups with fewer than two peptides (with at least one classified as unique) were filtered out from the MaxQuant proteinGroups.txt file. Missing values were imputed by shifting a beta distribution, based on the LFQ intensity values, to the limit of quantitation. Volcano plots, heatmap, and Venn diagram were generated from Rstudio using packages including ggplot2 and more. The threshold for protein enrichment was set at a fold change > |2| and a p-value < 0.05 (Welch's t test).

RNA-related works

Quantitative real-time PCR (qRT-PCR)

RNA was extracted from 5 dpf larvae, 33 dpf gonads, 40 dpf differentiating testes, and 1 year-old differentiated testes for both wild-type and *zbtb48*^{-/-} genotypes. Each biological replicate consisted of either a pool of 25 larvae or the gonad from a single individual. Samples were lysed in 500 µL of TRIzol Reagent (Invitrogen) and sonicated for 3 cycles of 30s ON and 30s OFF at 4 °C. RNA purification was performed using the Direct-zol RNA Microprep kit protocol (Zymo Research), and 1 µg of RNA was converted to cDNA with the First Strand cDNA Synthesis Kit (ThermoFisher) for larvae and ProtoScript II First Strand cDNA Synthesis Kit (NEB) for gonads.

Quantitative real-time PCR (qRT-PCR) was conducted using SYBR™ Power SYBR™ Green PCR Master Mix (Applied Biosystems), with three technical replicates. The qRT-PCR was carried out on a QuantStudio™ 5 Real-Time PCR machine under the following conditions: 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min, ending with a melting curve analysis. β -actin was used as the reference gene, and data were analyzed using the 2^{- $\Delta\Delta$ Ct} method.

RNA-seq

From a single mating of *zbtb48*^{+/-} parents, 22 sibling larvae were picked at 5 dpf and processed individually. RNA and DNA were extracted using an AllPrep DNA/RNA/Protein Mini Kit (Qiagen) according to the manufacturer's instructions. Larvae were lysed in the provided RLT buffer by the sonication. Genotyping was performed on the extracted DNA, and subsequently, the respective RNA of the 4 selected wild-type and *zbtb48*^{-/-} mutants were submitted for RNA sequencing (RNA-seq), with each larva representing a biological replicate.

RNA-seq preparations and sequencing were handled by the IMB Genomic Core Facility. NGS library preparation was performed with Illumina's TruSeq Stranded mRNA LT Sample Prep Kit following the TruSeq Stranded mRNA Reference Guide (Oct. 2017) (Document # 1000000040498v00) using ¼ of the reagents. Libraries were prepared with a starting amount of 250 ng and amplified in 10 PCR cycles. Libraries were profiled on a high-sensitivity DNA chip on a 2100 Bioanalyzer (Agilent Technologies) and quantified using the Qubit dsDNA HS Assay Kit in a Qubit 2.0 Fluorometer (Life Technologies). All 8 samples were pooled together with 20 samples from another project in equimolar ratio and sequenced on a NextSeq 500 High Output FC, SR for 1x 75 cycles plus 2x 8 cycles for the dual index read.

The library quality was assessed with FastQC (version 0.11.9) and FastQscreen (version 0.15.2) before alignment against the *D. rerio* genome assembly GRCz11 and its canonical gene annotations (Danio_rerio.GRCz11.100.chr.gtf). Alignment was performed with STAR aligner²¹⁸ version 2.7. Reads mapped to annotated features in the GTF file were counted with featureCounts²¹⁹ version 1.6.2 using the featureCounts functionality. Coverage tracks were generated with deepTools²²⁰ version 3.5.1 and plotted using Gviz²²¹ on an R framework. The threshold for transcript enrichment was set at a fold change > |2| and a p-value < 0.01 (Welch's t test).

Analysis of published RNA-seq data

Metadata from RNA-seq analysis of embryo development ^{135,151,152} and adult tissues ¹⁵³ were downloaded and processed in R.

The data and images of the scRNA measurements of 40 dpf ovaries ¹⁵⁵ and adult testes ¹⁵⁴ were obtained from the following websites: https://singlecell.broadinstitute.org/single_cell/study/SCP928/40dpf-ovary-all-cells and https://sposato.shinyapps.io/testis_shiny_app/.

Microscopy imaging and analysis

Breeding

Embryos and larvae were bred from in-crossing parents of the same homozygous genotype to ensure 100% offspring of the same genotype, unless stated otherwise. In some *hmbox1a/b* staining experiments, larvae with *Tg(wt1b:eGFP)li1* transgenic background lines were used to label the podocytes.

Fixing and Dehydration

Embryos younger than 6 hpf (50% epiboly) were fixed in 4% paraformaldehyde (PFA) at room temperature for 1 h, while older embryos were fixed overnight at 4 °C. After fixation, the PFA was removed, and embryos were washed twice with PBST (PBS with 0.05% Tween-20) before manually removing their chorions.

For 24 hpf (1 dpf) embryos, chorions were removed before fixing in 4% PFA at room temperature for 1 h. To prevent pigmentation in larvae older than 1 dpf, they were treated with 0.003% PTU in E3 medium at 1 dpf, refreshed at 3 dpf. These larvae were also dechorionated before fixing in 4% PFA at room temperature for 2 h.

All samples were washed once in PBST, dehydrated through a methanol/H₂O series (25%, 50%, 75%, 100%, and 100% methanol). They were then stored overnight in 100% methanol at -20 °C. Before proceeding to the staining steps each time, samples were rehydrated in a reverse methanol/H₂O series (75%, 50%, and 25% methanol, followed by twice in PBST).

Immunofluorescence (IF)

All incubations and washes for immunofluorescence (IF) were performed on a shaker at room temperature, unless specified otherwise, with each wash in PBST lasting for 10 min. Rehydrated embryos (1 dpf and younger) were permeabilized in PBST for 30min while larvae (2 dpf and older) were permeabilized with ice cold 100% acetone in -20 °C for 7 min. Samples were blocked in 10% BSA (Sigma) for at least 2 h, followed by an overnight incubation with primary antibody at 4 °C. The next day, the samples were washed three times before incubating with secondary antibody and DAPI for 1 h. Samples were washed three times before mounting onto slides or 8-well chambers (ibidi) with mounting medium (ibidi). Samples were imaged on the same day whenever possible using either the VisiScope spinning disk confocal microscope or the BC43 benchtop confocal microscope (Andor).

The primary and secondary antibodies used in this study were prepared in 10% BSA. The primary antibodies were diluted as follows: rabbit antibodies raised in this study, such as Hmbox1a (T1 and T3), Hmbox1b (T2), Zbtb48-BTB (T2 and T3), and Zbtb48-ZNF (T2 and T3) antibodies, were used at dilutions between 1:100 to 1:400 of elution 3, whereas Hmbox1b-T2

(M) and Zbtb48-T3 (M) antibodies, purified by the protein production core facility, was used at 1:2000. Mouse anti-human HMBOX1 antibody was used at 1:100, and rat anti-zebrafish Ziwi antibody (Ketting's self-raised antibody, 396) was used at 1:250. For secondary antibodies, anti-mouse IgG (H+L) AF488 (Invitrogen), anti-rabbit IgG (H+L) AF488 (Invitrogen), anti-rabbit IgG (H+L) AF568 (Invitrogen), and anti-rat IgG IRDye 680RD (Li-COR) were used at a 1:250 dilution.

IF with single-molecule FISH (smFISH) *cdh17* probe set

A Stellaris™ RNA FISH probe set comprising 48 unique probes labelled with Quasar® 670 dye was synthesized to target zebrafish *cdh17* mRNA (OTTDART00000026378) to mark the pronephros.

All incubations and washes for IF-single-molecule FISH (smFISH) were performed on a thermomixer at 30 °C with shaking at 300 rpm, with each wash lasting for 10 min. The smFISH hybridization buffer was prepared as follows: 10% dextran sulfate (Sigma), 10% formamide (Roth), 1 mg/mL yeast tRNA (Invitrogen), 2X SSC, 0.02% BSA (Sigma), and 2 mM vanadyl-ribonucleoside complex (NEB).

Rehydrated samples were pre-hybridized in smFISH hybridization buffer for 2 h. The samples were hybridized overnight with a 1:50 dilution of smFISH-*cdh17*-Quasar670 probe, added with either a 1:400 dilution of Hmbox1a-T1 antibody or a 1:2000 dilution of Hmbox1b (M) antibody. The next day, samples were washed twice with pre-warmed smFISH wash buffer (10% formamide with 2X SSC), followed by a 1 h incubation with a 1:250 dilution of anti-rabbit-AF568 secondary antibody (Invitrogen) and 1:500 DAPI prepared in the smFISH wash buffer. Afterward, samples were washed three times with smFISH wash buffer and once with PBST at room temperature before mounting on chamber slides. They were imaged on the same day using a BC43 benchtop confocal microscope (Andor) with a 20X objective.

Telomere staining

All incubations and washes for telomere staining were conducted on a thermomixer with shaking at 300 rpm for 10 min, unless stated otherwise. The hybridization buffer for the TAMRA-labeled C-rich telomere probe (Eurogentec) was prepared using the following components: 1X Blocking Reagent (Roche), 70% formamide (Sigma), 100 µg/mL yeast tRNA (Invitrogen), 100 µg/mL salmon sperm DNA (Invitrogen), and 10 mM Tris-HCl (pH 7.2), all diluted in maleic acid buffer (100 mM maleic acid (Sigma), 150 mM NaCl, pH 7.5).

The telomere probe was diluted 1:100 in the hybridization buffer and preheated to 90 °C prior to use. For staining telomere staining coupled with IF on 4 hpf embryos, samples were rehydrated and incubated with the preheated hybridization mixture at 85 °C, followed by 1 h of hybridization at room temperature. The samples were washed once with 2X SSC/0.1% Tween-20 at 30 °C before rinsing with PBST. They were then blocked with 10% BSA for subsequent IF staining with the Hmbox1a antibody.

For telomere staining coupled with IF-smFISH on 1 dpf embryos, the telomere staining was performed as described above for 4 hpf, but the room temperature hybridization was extended to 4 h, and samples were washed three times with smFISH wash buffer before proceeding to the pre-hybridization step of smFISH. The samples were subsequently probed with the smFISH-*cdh17* and Hmbox1b antibody with the methods mentioned above.

Detection, colocalization, and quantification of telomere foci and Hmbox1a signals were carried out using the ComDet v.0.5.5 plugin for ImageJ (<https://github.com/UU-cellbiology/ComDet>). The threshold for the maximum distance between colocalized spots was set at 1 pixel, with the approximate particle size for both telomere foci and Hmbox1a signals set at 2 pixels. The intensity threshold (in SD) was established at 3.

Glomerulus filtration test

The glomerulus filtration assay was conducted following the methods previously described^{146,147}. The injection mixture was prepared by combining equal parts of 20 mg/mL Alexa Fluor 647-conjugated 10-kDa dextran (Invitrogen) and 10 mg/mL FITC-conjugated 500-kDa dextran (Invitrogen) in PBS. Wild-type and *hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD} double mutant larvae at 4 dpf were anesthetized using 0.02% tricaine (Sigma) before the dextran mixture was injected into their hearts, after which they were placed in E3 medium for recovery.

At 3 hours post-injection (hpi), the larvae were re-anesthetized for imaging with a VisiScope spinning disk confocal microscope. Following imaging, the larvae were returned to E3 medium at 28 °C until 24 hpi, when they were anesthetized and imaged again. To evaluate dextran clearance, images from the 10-kDa and 500-kDa dextran channels were overlaid, and fluorescence intensity was compared between the wild-type and *hmbox1a*^{LD/LD}; *hmbox1b*^{LD/LD} larvae at both the 3 hpi and 24 hpi time points.

Mitotracker

Wild-type and *zbtb48*^{-/-} embryos were bred from parents of the respective genotype. Embryos were collected at around the 32-cell to 64-cell stage (1.5 hpf) and incubated with 1 mL of 1:2000 MitoTracker™ Deep Red FM Dye (Invitrogen) diluted in E3 medium for 2 h at 28 °C, in the dark. The embryos were washed once with E3 medium for 5 min and fixed with 4% PFA at room temperature for 1.5 h. The embryos were washed with PBST for 5 min with shaking and manually dechorionated. The embryos were stained with 1:500 DAPI at room temperature for 5 min before being mounted onto a slide and imaged using the BC43 benchtop confocal microscope (Andor) with a 100X objective.

Quantitative fluorescence in situ hybridization (q-FISH)

Telomere quantification using q-FISH method was only performed on larva in the *zbtb48* study. Thirty larvae bred from the single mating of *zbtb48*^{+/-} parents were collected at 30 hpf and 5 dpf. The larvae were fixed in 4% PFA for 1 h at room temperature, followed by two washes with PBS/0.05% Tween-20. The samples were then dehydrated through a methanol/H₂O series and stored overnight in 100% methanol at -20 °C prior to hybridization. Samples were rehydrated in 2X SSC/0.1% Tween-20 twice for 10 min each.

To ensure uniform staining, all larvae of the same stage were processed together in the same tube. Similar to the telomere staining mentioned earlier, a TAMRA-labeled C-rich telomere probe was used at a 1:100 dilution and combined with 1 µg/mL DAPI in 250 µL of hybridization buffer. The hybridization mixture was preheated to 90 °C before incubating it with the samples at 85 °C. The samples were hybridized at room temperature for 2 h, followed by three washes with 2X SSC/0.1% Tween-20 for 10 min each at 55 °C with 300 rpm shaking. The larvae were then cut at the trunk, with the tails mounted on glass slides for telomere length analysis and the other half used for genotyping.

The tail biopsies were imaged using the BC43 benchtop confocal microscope (Andor) with a 100X objective. The images were processed using ImageJ software. To measure telomere fluorescence intensity, a difference of Gaussian blur filter (sigma = 1 subtract sigma = 4) was used to set the threshold for telomere foci identification. The fluorescence intensity in the telomere channel was normalized to the fluorescence intensity of the corresponding area in the DAPI channel. Measurements were performed throughout the z-stack images, and the normalized intensity values were averaged to represent each individual. Medians were then calculated for each genotype, and p-values were obtained using Welch's t-test.

Telomere length measurement

Telomere restriction fragment (TRF) analysis by Southern blot

Telomere length was measured on the caudal fins of three 8-month-old males of each genotype. For *hmbox1a/b* study, *hmbox1a^{LD/LD}*; *hmbox1b^{LD/LD}* and the wild-type counterpart fish were bred from a group mating of *hmbox1a^{+LD}*; *hmbox1b^{+LD}* parents. For *zbtb48* study, *zbtb48^{-/-}* and the wild-type counterpart fish bred from a group mating of *zbtb48^{+/-}* parents. The TRF experiment and analysis was performed by our collaborator, Miguel Godinho Ferreira from Institute for Research on Cancer and Aging, Nice (IRCAN).

Fish genotype distribution and survival curve

Genotype distribution

The genotype distribution was analysed on the first generation mutants, offsprings bred from heterozygous parents (*hmbox1a^{+LD}*; *hmbox1b^{+LD}* or *zbtb48^{+/-}*), at the indicated ages. Chi-square tests were performed on individual and the overall crosses to evaluate whether the distribution follows Mendelian inheritance pattern.

Survival curve

The number of fish per homozygous genotype was monitored by routinely counting those left in each tank. Dead or sick fish were removed by caretakers, but escapes (assumed dead) could not be tracked. As a result, Kaplan-Meier survival curves were not used. Instead, simpler graphs displayed the percentage of fish remaining in each tank by genotype over time, with 100% based on counts from the time of genotyping (for the first generation) or when the larvae were placed in the aquarium (for the subsequent generation).

References

1. Goodenough U, Heitman J. Origins of eukaryotic sexual reproduction. *Cold Spring Harb Perspect Biol.* 2014;6(3). doi:10.1101/cshperspect.a016154
2. Meyne J, Ratliff RL, Moyzis RK. Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proceedings of the National Academy of Sciences.* 1989;86(18):7049-7053. doi:10.1073/pnas.86.18.7049
3. Gomes NMV, Shay JW, Wright WE. Telomere biology in Metazoa. *FEBS Lett.* 2010;584(17):3741-3751. doi:10.1016/j.febslet.2010.07.031
4. Fulnečková J, Ševčíková T, Fajkus J, et al. A Broad Phylogenetic Survey Unveils the Diversity and Evolution of Telomeres in Eukaryotes. *Genome Biol Evol.* 2013;5(3):468-483. doi:10.1093/gbe/evt019
5. Červenák F, Sepšiová R, Nosek J, Tomáška L. Step-by-Step Evolution of Telomeres: Lessons from Yeasts. *Genome Biol Evol.* 2021;13(2). doi:10.1093/gbe/evaa268
6. Richards EJ, Ausubel FM. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell.* 1988;53(1):127-136. doi:10.1016/0092-8674(88)90494-1
7. Cangiano G, Volpe A La. Repetitive DNA sequences located in the terminal portion of the *Caenorhabditis elegans* chromosomes. *Nucleic Acids Res.* 1993;21(5):1133-1139. doi:10.1093/nar/21.5.1133
8. Blackburn EH, Gall JG. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J Mol Biol.* 1978;120(1):33-53. doi:10.1016/0022-2836(78)90294-2
9. Butler MG, Tilburt J, DeVries A, et al. Comparison of Chromosome Telomere Integrity in Multiple Tissues from Subjects at Different Ages. *Cancer Genet Cytogenet.* 1998;105(2):138-144. doi:10.1016/S0165-4608(98)00029-6
10. Daniali L, Benetos A, Susser E, et al. Telomeres shorten at equivalent rates in somatic tissues of adults. *Nat Commun.* 2013;4(1):1597. doi:10.1038/ncomms2602
11. Demanelis K, Jasmine F, Chen LS, et al. Determinants of telomere length across human tissues. *Science (1979).* 2020;369(6509). doi:10.1126/science.aaz6876
12. Graakjaer J, Pascoe L, Der-Sarkissian H, et al. The relative lengths of individual telomeres are defined in the zygote and strictly maintained during life. *Aging Cell.* 2004;3(3):97-102. doi:10.1111/j.1474-9728.2004.00093.x
13. Londoño-Vallejo JA, DerSarkissian H, Cazes L, Thomas G. Differences in telomere length between homologous chromosomes in humans. *Nucleic Acids Res.* 2001;29(15):3164-3171. doi:10.1093/nar/29.15.3164
14. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature.* 1990;345(6274):458-460. doi:10.1038/345458a0
15. Huffman KE, Levene SD, Tesmer VM, Shay JW, Wright WE. Telomere Shortening Is Proportional to the Size of the G-rich Telomeric 3'-Overhang. *Journal of Biological Chemistry.* 2000;275(26):19719-19722. doi:10.1074/jbc.M002843200
16. Griffith JD, Comeau L, Rosenfield S, et al. Mammalian Telomeres End in a Large Duplex Loop. *Cell.* 1999;97(4):503-514. doi:10.1016/S0092-8674(00)80760-6
17. de Lange T. How Telomeres Solve the End-Protection Problem. *Science (1979).* 2009;326(5955):948-952. doi:10.1126/science.1170633
18. Fagagna F d'Adda di, Reaper PM, Clay-Farrace L, et al. A DNA damage checkpoint response in telomere-initiated senescence. *Nature.* 2003;426(6963):194-198. doi:10.1038/nature02118
19. Herbig U, Jobling WA, Chen BPC, Chen DJ, Sedivy JM. Telomere Shortening Triggers Senescence of Human Cells through a Pathway Involving ATM, p53, and p21CIP1, but Not p16INK4a. *Mol Cell.* 2004;14(4):501-513. doi:10.1016/S1097-2765(04)00256-4
20. Hayflick L, Moorhead PS. The serial cultivation of human diploid cell strains. *Exp Cell Res.* 1961;25(3):585-621. doi:10.1016/0014-4827(61)90192-6
21. Campisi J. Cellular senescence as a tumor-suppressor mechanism. *Trends Cell Biol.* 2001;11:S27-S31. doi:10.1016/S0962-8924(01)82148-6
22. Counter CM, Avilion AA, LeFeuvre CE, et al. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* 1992;11(5):1921-1929. doi:10.1002/j.1460-2075.1992.tb05245.x
23. van Deursen JM. The role of senescent cells in ageing. *Nature.* 2014;509(7501):439-446. doi:10.1038/nature13193
24. Wilkinson HN, Hardman MJ. Senescence in Wound Repair: Emerging Strategies to Target Chronic Healing Wounds. *Front Cell Dev Biol.* 2020;8. doi:10.3389/fcell.2020.00773
25. Childs BG, Durik M, Baker DJ, van Deursen JM. Cellular senescence in aging and age-related disease: from mechanisms to therapy. *Nat Med.* 2015;21(12):1424-1435. doi:10.1038/nm.4000

26. Di Micco R, Krizhanovsky V, Baker D, d'Adda di Fagagna F. Cellular senescence in ageing: from mechanisms to therapeutic opportunities. *Nat Rev Mol Cell Biol.* 2021;22(2):75-95. doi:10.1038/s41580-020-00314-w
27. Coppé JP, Desprez PY, Krtolica A, Campisi J. The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annu Rev Pathol.* 2010;5:99-118. doi:10.1146/annurev-pathol-121808-102144
28. Schmitt CA, Wang B, Demaria M. Senescence and cancer — role and therapeutic opportunities. *Nat Rev Clin Oncol.* 2022;19(10):619-636. doi:10.1038/s41571-022-00668-4
29. Hiyama E, Hiyama K. Telomere and telomerase in stem cells. *Br J Cancer.* 2007;96(7):1020-1024. doi:10.1038/sj.bjc.6603671
30. Wu RA, Upton HE, Vogan JM, Collins K. Telomerase Mechanism of Telomere Synthesis. *Annu Rev Biochem.* 2017;86:439-460. doi:10.1146/annurev-biochem-061516-045019
31. Zhong FL, Batista LFZ, Freund A, Pech MF, Venteicher AS, Artandi SE. TPP1 OB-Fold Domain Controls Telomere Maintenance by Recruiting Telomerase to Chromosome Ends. *Cell.* 2012;150(3):481-494. doi:10.1016/j.cell.2012.07.012
32. Sandhu R, Sharma M, Wei D, Xu L. The structurally conserved TELR region on shelterin protein TPP1 is essential for telomerase processivity but not recruitment. *Proceedings of the National Academy of Sciences.* 2021;118(30). doi:10.1073/pnas.2024889118
33. Cai SW, de Lange T. CST-Pol α /Primase: the second telomere maintenance machine. *Genes Dev.* 2023;37(13-14):555-569. doi:10.1101/gad.350479.123
34. Bodnar AG, Ouellette M, Frolkis M, et al. Extension of Life-Span by Introduction of Telomerase into Normal Human Cells. *Science (1979).* 1998;279(5349):349-352. doi:10.1126/science.279.5349.349
35. Kim NW, Piatyszek MA, Prowse KR, et al. Specific Association of Human Telomerase Activity with Immortal Cells and Cancer. *Science (1979).* 1994;266(5193):2011-2015. doi:10.1126/science.7605428
36. Shay JW, Bacchetti S. A survey of telomerase activity in human cancer. *Eur J Cancer.* 1997;33(5):787-791. doi:10.1016/S0959-8049(97)00062-2
37. Bryan TM, Englezou A, Gupta J, Bacchetti S, Reddel RR. Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* 1995;14(17):4240-4248. doi:10.1002/j.1460-2075.1995.tb00098.x
38. Dunham MA, Neumann AA, Fasching CL, Reddel RR. Telomere maintenance by recombination in human cells. *Nat Genet.* 2000;26(4):447-450. doi:10.1038/82586
39. Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, Kerber RA. Association between telomere length in blood and mortality in people aged 60 years or older. *The Lancet.* 2003;361(9355):393-395. doi:10.1016/S0140-6736(03)12384-7
40. Vera E, Bernardes de Jesus B, Foronda M, Flores JM, Blasco MA. The Rate of Increase of Short Telomeres Predicts Longevity in Mammals. *Cell Rep.* 2012;2(4):732-737. doi:10.1016/j.celrep.2012.08.023
41. Whittmore K, Vera E, Martínez-Nevado E, Sanpera C, Blasco MA. Telomere shortening rate predicts species life span. *Proceedings of the National Academy of Sciences.* 2019;116(30):15122-15127. doi:10.1073/pnas.1902452116
42. de Lange T. Shelterin-Mediated Telomere Protection. *Annu Rev Genet.* 2018;52(1):223-247. doi:10.1146/annurev-genet-032918-021921
43. Broccoli D, Smogorzewska A, Chong L, de Lange T. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat Genet.* 1997;17(2):231-235. doi:10.1038/ng1097-231
44. Lei M, Podell ER, Cech TR. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol.* 2004;11(12):1223-1229. doi:10.1038/nsmb867
45. Erdel F, Kratz K, Willcox S, Griffith JD, Greene EC, de Lange T. Telomere Recognition and Assembly Mechanism of Mammalian Shelterin. *Cell Rep.* 2017;18(1):41-53. doi:10.1016/j.celrep.2016.12.005
46. Wang F, Podell ER, Zaug AJ, et al. The POT1–TPP1 telomere complex is a telomerase processivity factor. *Nature.* 2007;445(7127):506-510. doi:10.1038/nature05454
47. Hanaoka S, Nagadoi A, Yoshimura S, et al. NMR structure of the hrap1 myb motif reveals a canonical three-helix bundle lacking the positive surface charge typical of myb DNA-binding domains. *J Mol Biol.* 2001;312(1):167-175. doi:10.1006/jmbi.2001.4924
48. Arat NÖ, Griffith JD. Human Rap1 Interacts Directly with Telomeric DNA and Regulates TRF2 Localization at the Telomere. *Journal of Biological Chemistry.* 2012;287(50):41583-41594. doi:10.1074/jbc.M112.415984
49. Denchi EL, de Lange T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature.* 2007;448(7157):1068-1071. doi:10.1038/nature06065

50. Barrientos KS, Kendellen MF, Freibaum BD, Armbruster BN, Etheridge KT, Counter CM. Distinct Functions of POT1 at Telomeres. *Mol Cell Biol*. 2008;28(17):5251-5264. doi:10.1128/MCB.00048-08
51. Takai KK, Kibe T, Donigian JR, Frescas D, de Lange T. Telomere Protection by TPP1/POT1 Requires Tethering to TIN2. *Mol Cell*. 2011;44(4):647-659. doi:10.1016/j.molcel.2011.08.043
52. Benarroch-Popivker D, Pisano S, Mendez-Bermudez A, et al. TRF2-Mediated Control of Telomere DNA Topology as a Mechanism for Chromosome-End Protection. *Mol Cell*. 2016;61(2):274-286. doi:10.1016/j.molcel.2015.12.009
53. Timashev LA, De Lange T. Characterization of t-loop formation by TRF2. *Nucleus*. 2020;11(1):164-177. doi:10.1080/19491034.2020.1783782
54. Wu P, van Overbeek M, Rooney S, de Lange T. Apollo Contributes to G Overhang Maintenance and Protects Leading-End Telomeres. *Mol Cell*. 2010;39(4):606-617. doi:10.1016/j.molcel.2010.06.031
55. Lenain C, Bauwens S, Amiard S, Brunori M, Giraud-Panis MJ, Gilson E. The Apollo 5' Exonuclease Functions Together with TRF2 to Protect Telomeres from DNA Repair. *Current Biology*. 2006;16(13):1303-1310. doi:10.1016/j.cub.2006.05.021
56. Lam YC, Akhter S, Gu P, et al. SNMIB/Apollo protects leading-strand telomeres against NHEJ-mediated repair. *EMBO J*. 2010;29(13):2230-2241. doi:10.1038/emboj.2010.58
57. Ribes-Zamora A, Indiviglio SM, Mihalek I, Williams CL, Bertuch AA. TRF2 Interaction with Ku Heterotetramerization Interface Gives Insight into c-NHEJ Prevention at Human Telomeres. *Cell Rep*. 2013;5(1):194-206. doi:10.1016/j.celrep.2013.08.040
58. Bae NS, Baumann P. A RAP1/TRF2 Complex Inhibits Nonhomologous End-Joining at Human Telomeric DNA Ends. *Mol Cell*. 2007;26(3):323-334. doi:10.1016/j.molcel.2007.03.023
59. Rai R, Chen Y, Lei M, Chang S. TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions. *Nat Commun*. 2016;7(1):10881. doi:10.1038/ncomms10881
60. van Steensel B, Smogorzewska A, de Lange T. TRF2 Protects Human Telomeres from End-to-End Fusions. *Cell*. 1998;92(3):401-413. doi:10.1016/S0092-8674(00)80932-0
61. Giraud-Panis MJ, Teixeira MT, Géli V, Gilson E. CST Meets Shelterin to Keep Telomeres in Check. *Mol Cell*. 2010;39(5):665-676. doi:10.1016/j.molcel.2010.08.024
62. Price C, Boltz KA, Chaiken MF, Stewart JA, Beilstein MA, Shippen DE. Evolution of CST function in telomere maintenance. *Cell Cycle*. 2010;9(16):3177-3185. doi:10.4161/cc.9.16.12547
63. Sun J, Yu EY, Yang Y, et al. Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes Dev*. 2009;23(24):2900-2914. doi:10.1101/gad.1851909
64. Gao H, Cervantes RB, Mandell EK, Otero JH, Lundblad V. RPA-like proteins mediate yeast telomere function. *Nat Struct Mol Biol*. 2007;14(3):208-214. doi:10.1038/nsmb1205
65. Huang C, Dai X, Chai W. Human Stn1 protects telomere integrity by promoting efficient lagging-strand synthesis at telomeres and mediating C-strand fill-in. *Cell Res*. 2012;22(12):1681-1695. doi:10.1038/cr.2012.132
66. Feng X, Hsu SJ, Kasbek C, Chaiken M, Price CM. CTC1-mediated C-strand fill-in is an essential step in telomere length maintenance. *Nucleic Acids Res*. 2017;45(8):4281-4293. doi:10.1093/nar/gkx125
67. Miyake Y, Nakamura M, Nabetani A, et al. RPA-like Mammalian Ctc1-Stn1-Ten1 Complex Binds to Single-Stranded DNA and Protects Telomeres Independently of the Pot1 Pathway. *Mol Cell*. 2009;36(2):193-206. doi:10.1016/j.molcel.2009.08.009
68. Survtseva Y V, Churikov D, Boltz KA, et al. Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes. *Mol Cell*. 2009;36(2):207-218. doi:10.1016/j.molcel.2009.09.017
69. Lyu X, Sang PB, Chai W. CST in maintaining genome stability: Beyond telomeres. *DNA Repair (Amst)*. 2021;102:103104. doi:10.1016/j.dnarep.2021.103104
70. Lim CJ, Cech TR. Shaping human telomeres: from shelterin and CST complexes to telomeric chromatin organization. *Nat Rev Mol Cell Biol*. 2021;22(4):283-298. doi:10.1038/s41580-021-00328-y
71. Bluhm A, Viceconte N, Li F, et al. ZBTB10 binds the telomeric variant repeat TTGGGG and interacts with TRF2. *Nucleic Acids Res*. 2019;47(4):1896-1907. doi:10.1093/nar/gky1289
72. Zhou M, Cui Y, Zuo S, et al. ZBTB40 is a telomere-associated protein and protects telomeres in human ALT cells. *J Biol Chem*. 2023;299(9):105053. doi:10.1016/j.jbc.2023.105053
73. Jahn A, Rane G, Paszkowski-Rogacz M, et al. ZBTB48 is both a vertebrate telomere-binding protein and a transcriptional activator. *EMBO Rep*. 2017;18(6):929-946. doi:10.15252/embr.201744095
74. Li JSZ, Miralles Fusté J, Simavorian T, et al. TZAP: A telomere-associated protein involved in telomere length control. *Science (1979)*. 2017;355(6325):638-641. doi:10.1126/science.aah6752
75. Braun H, Xu Z, Chang F, et al. ZNF524 directly interacts with telomeric DNA and supports telomere integrity. *Nat Commun*. 2023;14(1):8252. doi:10.1038/s41467-023-43397-7

76. Chung I, Osterwald S, Deeg KI, Rippe K. PML body meets telomere. *Nucleus*. 2012;3(3):263-275. doi:10.4161/nucl.20326
77. Kappei D, Butter F, Benda C, et al. HOT1 is a mammalian direct telomere repeat-binding protein contributing to telomerase recruitment. *EMBO J*. 2013;32(12):1681-1701. doi:10.1038/emboj.2013.105
78. Feng X, Luo Z, Jiang S, et al. The telomere-associated homeobox-containing protein TAH1 participates in telomere maintenance in ALT Cells. *J Cell Sci*. Published online January 1, 2013. doi:10.1242/jcs.128512
79. Kappei D, Scheibe M, Paszkowski-Rogacz M, et al. Phylointeractomics reconstructs functional evolution of protein binding. *Nat Commun*. 2017;8(1):14334. doi:10.1038/ncomms14334
80. Chen S, Saiyin H, Zeng X, et al. Isolation and functional analysis of human HMBOX1, a homeobox containing protein with transcriptional repressor activity. *Cytogenet Genome Res*. 2006;114(2):131-136. doi:10.1159/000093328
81. Dai J, Zhang C, Tian Z, Zhang J. Expression profile of HMBOX1, a novel transcription factor, in human cancers using highly specific monoclonal antibodies. *Exp Ther Med*. 2011;2(3):487-490. doi:10.3892/etm.2011.240
82. Dai J, Wu L, Zhang C, Zheng X, Tian Z, Zhang J. Recombinant Expression of a Novel Human Transcriptional Repressor HMBOX1 and Preparation of Anti-HMBOX1 Monoclonal Antibody. *Cell Mol Immunol*. 2009;6(4):261-268. doi:10.1038/cmi.2009.35
83. Jády BE, Bertrand E, Kiss T. Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J Cell Biol*. 2004;164(5):647-652. doi:10.1083/jcb.200310138
84. Zhu Y, Tomlinson RL, Lukowiak AA, Terns RM, Terns MP. Telomerase RNA accumulates in Cajal bodies in human cancer cells. *Mol Biol Cell*. 2004;15(1):81-90. doi:10.1091/mbc.e03-07-0525
85. Jády BE, Richard P, Bertrand E, Kiss T. Cell Cycle-dependent Recruitment of Telomerase RNA and Cajal Bodies to Human Telomeres. *Mol Biol Cell*. 2006;17(2):944-954. doi:10.1091/mbc.e05-09-0904
86. Wang C, Meier UT. Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins. *EMBO J*. 2004;23(8):1857-1867. doi:10.1038/sj.emboj.7600181
87. Lee JH, Hong J, Zhang Z, et al. Regulation of telomere homeostasis and genomic stability in cancer by N6-adenosine methylation (m6A). *Sci Adv*. 2021;7(31). doi:10.1126/sciadv.abg7073
88. Riou L, Bastos H, Lassalle B, et al. The Telomerase Activity of Adult Mouse Testis Resides in the Spermatogonial α 6-Integrin-Positive Side Population Enriched in Germinal Stem Cells. *Endocrinology*. 2005;146(9):3926-3932. doi:10.1210/en.2005-0502
89. Henson JD, Cao Y, Huschtscha LI, et al. DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat Biotechnol*. 2009;27(12):1181-1185. doi:10.1038/nbt.1587
90. Diao N, Li Y, Yang J, et al. High expression of HMBOX1 contributes to poor prognosis of gastric cancer by promoting cell proliferation and migration. *Biomedicine & Pharmacotherapy*. 2019;115:108867. doi:10.1016/j.biopha.2019.108867
91. Yu YL, Diao NN, Li YZ, et al. Low expression level of HMBOX1 in high-grade serous ovarian cancer accelerates cell proliferation by inhibiting cell apoptosis. *Biochem Biophys Res Commun*. 2018;501(2):380-386. doi:10.1016/j.bbrc.2018.04.203
92. Su L, Zhao H, Sun C, et al. Role of Hmbox1 in Endothelial Differentiation of Bone-Marrow Stromal Cells by a Small Molecule. *ACS Chem Biol*. 2010;5(11):1035-1043. doi:10.1021/cb100153r
93. Han L, Shao J, Su L, et al. A Chemical Small Molecule Induces Mouse Embryonic Stem Cell Differentiation into Functional Vascular Endothelial Cells via Hmbox1. *Stem Cells Dev*. 2012;21(15):2762-2769. doi:10.1089/scd.2012.0055
94. Ma H, Su L, Yue H, et al. HMBOX1 interacts with MT2A to regulate autophagy and apoptosis in vascular endothelial cells. *Sci Rep*. 2015;5(1):15121. doi:10.1038/srep15121
95. Jiang Y, Mu H, Zhao H. HMBOX1, a member of the homeobox family: current research progress. *Central European Journal of Immunology*. 2023;48(1):63-69. doi:10.5114/ceji.2023.126615
96. Zhao Y, Zhang G, He C, Mei Y, Shi Y, Li F. The 11th C2H2 zinc finger and an adjacent C-terminal arm are responsible for TZAP recognition of telomeric DNA. *Cell Res*. 2018;28(1):130-134. doi:10.1038/cr.2017.141
97. Wang S, Xu Z, Li M, et al. Structural insights into the recognition of telomeric variant repeat TTGGGG by broad-complex, tramtrack and bric-à-brac - zinc finger protein ZBTB10. *Journal of Biological Chemistry*. 2023;299(3):102918. doi:10.1016/j.jbc.2023.102918
98. Moreno SP, Fusté JM, Kaiser M, et al. TZAP overexpression induces telomere dysfunction and ALT-like activity in ATRX/DAXX-deficient cells. *iScience*. 2023;26(4):106405. doi:10.1016/j.isci.2023.106405
99. Heo YR, Lee MH, Kwon SY, Cho J, Lee JH. TZAP Mutation Leads to Poor Prognosis of Patients with Breast Cancer. *Medicina (Kaunas)*. 2019;55(11). doi:10.3390/medicina55110748

100. Park WJ, Park JH, Shin HY, Lee JH. Clinical and Prognostic Significance of TZAP Expression in Cervical Cancer. *Medicina (B Aires)*. 2020;56(5):207. doi:10.3390/medicina56050207
101. Jung SJ, Seo YR, Park WJ, et al. Clinicopathological Characteristics of TZAP Expression in Colorectal Cancers. *Onco Targets Ther*. 2020;Volume 13:12933-12942. doi:10.2147/OTT.S274394
102. Kim GJ, Lee JH, Chae M, Lee DH. Prognostic Value of Telomeric Zinc Finger-Associated Protein Expression in Adenocarcinoma and Squamous Cell Carcinoma of Lung. *Medicina (B Aires)*. 2021;57(11):1223. doi:10.3390/medicina57111223
103. Jung SJ, Kil SH, Lee HW, et al. Clinical Characteristics of TZAP (ZBTB48) in Hepatocellular Carcinomas from Tissue, Cell Line, and TCGA. *Medicina (B Aires)*. 2022;58(12):1778. doi:10.3390/medicina58121778
104. Dos Santos GA, Viana NI, Pimenta R, et al. Telomeric zinc-finger associated protein (TZAP) in cancer biology: friend or foe? *Mol Biol Res Commun*. 2021;10(3):121-129. doi:10.22099/mbr.2021.40106.1607
105. Yoon JH, Choi WI, Jeon BN, et al. Human Krüppel-related 3 (HKR3) Is a Novel Transcription Activator of Alternate Reading Frame (ARF) Gene. *Journal of Biological Chemistry*. 2014;289(7):4018-4031. doi:10.1074/jbc.M113.526855
106. Gil J, Peters G. Regulation of the INK4b–ARF–INK4a tumour suppressor locus: all for one or one for all. *Nat Rev Mol Cell Biol*. 2006;7(9):667-677. doi:10.1038/nrm1987
107. Gallagher SJ, Kefford RF, Rizos H. The ARF tumour suppressor. *Int J Biochem Cell Biol*. 2006;38(10):1637-1641. doi:10.1016/j.biocel.2006.02.008
108. Harland M, Taylor CF, Chambers PA, et al. A mutation hotspot at the p14ARF splice site. *Oncogene*. 2005;24(28):4604-4608. doi:10.1038/sj.onc.1208678
109. Rane G, Kuan VLS, Wang S, et al. ZBTB48 is a pioneer factor regulating B-cell-specific CIITA expression. Published online July 27, 2023. doi:10.1101/2023.07.25.550481
110. Bie Y nan, Gu P, Chen Y ting, et al. TZAP plays an inhibitory role in the self-renewal of porcine mesenchymal stromal cells and is implicated the regulation of premature senescence via the p53 pathway. *J Transl Med*. 2019;17(1):72. doi:10.1186/s12967-019-1820-8
111. Howe K, Clark MD, Torroja CF, et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*. 2013;496(7446):498-503. doi:10.1038/nature12111
112. Choi TY, Choi TI, Lee YR, Choe SK, Kim CH. Zebrafish as an animal model for biomedical research. *Exp Mol Med*. 2021;53(3):310-317. doi:10.1038/s12276-021-00571-5
113. Adhish M, Manjubala I. Effectiveness of zebrafish models in understanding human diseases—A review of models. *Heliyon*. 2023;9(3):e14557. doi:10.1016/j.heliyon.2023.e14557
114. Veldman MB, Lin S. Zebrafish as a Developmental Model Organism for Pediatric Research. *Pediatr Res*. 2008;64(5):470-476. doi:10.1203/PDR.0b013e318186e609
115. Astell KR, Sieger D. Zebrafish In Vivo Models of Cancer and Metastasis. *Cold Spring Harb Perspect Med*. 2020;10(8). doi:10.1101/cshperspect.a037077
116. Hason M, Bartůněk P. Zebrafish Models of Cancer-New Insights on Modeling Human Cancer in a Non-Mammalian Vertebrate. *Genes (Basel)*. 2019;10(11). doi:10.3390/genes10110935
117. Major RJ, Poss KD. Zebrafish Heart Regeneration as a Model for Cardiac Tissue Repair. *Drug Discov Today Dis Models*. 2007;4(4):219-225. doi:10.1016/j.ddmod.2007.09.002
118. González-Rosa JM. Zebrafish Models of Cardiac Disease: From Fortuitous Mutants to Precision Medicine. *Circ Res*. 2022;130(12):1803-1826. doi:10.1161/CIRCRESAHA.122.320396
119. Ross Stewart KM, Walker SL, Baker AH, Riley PR, Brittan M. Hooked on heart regeneration: the zebrafish guide to recovery. *Cardiovasc Res*. 2022;118(7):1667-1679. doi:10.1093/cvr/cvab214
120. Tonelli F, Bek JW, Besio R, et al. Zebrafish: A Resourceful Vertebrate Model to Investigate Skeletal Disorders. *Front Endocrinol (Lausanne)*. 2020;11. doi:10.3389/fendo.2020.00489
121. Carneiro MC, de Castro IP, Ferreira MG. Telomeres in aging and disease: lessons from zebrafish. *Dis Model Mech*. 2016;9(7):737-748. doi:10.1242/dmm.025130
122. Henriques CM, Carneiro MC, Tenente IM, Jacinto A, Ferreira MG. Telomerase Is Required for Zebrafish Lifespan. *PLoS Genet*. 2013;9(1):e1003214. doi:10.1371/journal.pgen.1003214
123. Anchelin M, Murcia L, Alcaraz-Pérez F, García-Navarro EM, Cayuela ML. Behaviour of Telomere and Telomerase during Aging and Regeneration in Zebrafish. *PLoS One*. 2011;6(2):e16955. doi:10.1371/journal.pone.0016955
124. Anchelin M, Alcaraz-Pérez F, Martínez CM, Bernabé-García M, Mulero V, Cayuela ML. Premature aging in telomerase-deficient zebrafish. *Dis Model Mech*. Published online January 1, 2013. doi:10.1242/dmm.011635
125. Carneiro MC, Henriques CM, Nabais J, Ferreira T, Carvalho T, Ferreira MG. Short Telomeres in Key Tissues Initiate Local and Systemic Aging in Zebrafish. *PLoS Genet*. 2016;12(1):e1005798. doi:10.1371/journal.pgen.1005798

126. El Maï M, Bird M, Allouche A, et al. Gut-specific telomerase expression counteracts systemic aging in telomerase-deficient zebrafish. *Nat Aging*. 2023;3(5):567-584. doi:10.1038/s43587-023-00401-5
127. Şerifoğlu N, Lopes-Bastos B, Ferreira MG. Lack of telomerase reduces cancer incidence and increases lifespan of zebrafish tp53M214K mutants. *Sci Rep*. 2024;14(1):5382. doi:10.1038/s41598-024-56153-8
128. Bednarek D, González-Rosa JM, Guzmán-Martínez G, et al. Telomerase Is Essential for Zebrafish Heart Regeneration. *Cell Rep*. 2015;12(10):1691-1703. doi:10.1016/j.celrep.2015.07.064
129. Imamura S, Uchiyama J, Koshimizu E, et al. A Non-Canonical Function of Zebrafish Telomerase Reverse Transcriptase Is Required for Developmental Hematopoiesis. *PLoS One*. 2008;3(10):e3364. doi:10.1371/journal.pone.0003364
130. Kishi S, Bayliss PE, Uchiyama J, et al. The Identification of Zebrafish Mutants Showing Alterations in Senescence-Associated Biomarkers. *PLoS Genet*. 2008;4(8):e1000152. doi:10.1371/journal.pgen.1000152
131. Ying Y, Hu X, Han P, et al. The non-telomeric evolutionary trajectory of TRF2 in zebrafish reveals its specific roles in neurodevelopment and aging. *Nucleic Acids Res*. 2022;50(4):2081-2095. doi:10.1093/nar/gkac065
132. Ma J, Tang D, Gao P, Liang S, Zhang R. Knockout of Shelterin subunit genes in zebrafish results in distinct outcomes. *Biochem Biophys Res Commun*. 2022;617:22-29. doi:10.1016/j.bbrc.2022.05.079
133. Xie Y, Yang D, He Q, Songyang Z. Zebrafish as a Model System to Study the Physiological Function of Telomeric Protein TPP1. *PLoS One*. 2011;6(2):e16440. doi:10.1371/journal.pone.0016440
134. Simon AJ, Lev A, Zhang Y, et al. Mutations in STN1 cause Coats plus syndrome and are associated with genomic and telomere defects. *Journal of Experimental Medicine*. 2016;213(8):1429-1440. doi:10.1084/jem.20151618
135. White RJ, Collins JE, Sealy IM, et al. A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife*. 2017;6. doi:10.7554/eLife.30860
136. Link V, Shevchenko A, Heisenberg CP. Proteomics of early zebrafish embryos. *BMC Dev Biol*. 2006;6(1):1. doi:10.1186/1471-213X-6-1
137. Chhetri G, Kalita P, Tripathi T. An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli. *MethodsX*. 2015;2:385-391. doi:10.1016/j.mex.2015.09.005
138. Senghaas N, Köster RW. Culturing and Transfecting Zebrafish PAC2 Fibroblast Cells. *Cold Spring Harb Protoc*. 2009;2009(6):pdb.prot5235. doi:10.1101/pdb.prot5235
139. He S, Salas-Vidal E, Rueb S, et al. Genetic and Transcriptome Characterization of Model Zebrafish Cell Lines. *Zebrafish*. 2006;3(4):441-453. doi:10.1089/zeb.2006.3.441
140. Horsfield J, Ramachandran A, Reuter K, et al. Cadherin-17 is required to maintain pronephric duct integrity during zebrafish development. *Mech Dev*. 2002;115(1-2):15-26. doi:10.1016/S0925-4773(02)00094-1
141. Wingert RA, Selleck R, Yu J, et al. The cdx Genes and Retinoic Acid Control the Positioning and Segmentation of the Zebrafish Pronephros. *PLoS Genet*. 2007;3(10):e189. doi:10.1371/journal.pgen.0030189
142. Zhou W, Boucher RC, Bollig F, Englert C, Hildebrandt F. Characterization of mesonephric development and regeneration using transgenic zebrafish. *American Journal of Physiology-Renal Physiology*. 2010;299(5):F1040-F1047. doi:10.1152/ajprenal.00394.2010
143. Carvalho L, Heisenberg CP. The yolk syncytial layer in early zebrafish development. *Trends Cell Biol*. 2010;20(10):586-592. doi:10.1016/j.tcb.2010.06.009
144. Laue K, Rajshekar S, Courtney AJ, Lewis ZA, Goll MG. The maternal to zygotic transition regulates genome-wide heterochromatin establishment in the zebrafish embryo. *Nat Commun*. 2019;10(1):1551. doi:10.1038/s41467-019-09582-3
145. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Developmental Dynamics*. 1995;203(3):253-310. doi:10.1002/aja.1002030302
146. Christou-Savina S, Beales PL, Osborn DPS. Evaluation of Zebrafish Kidney Function Using a Fluorescent Clearance Assay. *Journal of Visualized Experiments*. 2015;(96). doi:10.3791/52540
147. Kotb AM, Müller T, Xie J, Anand-Apte B, Endlich K, Endlich N. Simultaneous assessment of glomerular filtration and barrier function in live zebrafish. *American Journal of Physiology-Renal Physiology*. 2014;307(12):F1427-F1434. doi:10.1152/ajprenal.00029.2014
148. Mrakovčić M, Haley LE. Inbreeding depression in the Zebra fish *Brachydanio rerio* (Hamilton Buchanan). *J Fish Biol*. 1979;15(3):323-327. doi:10.1111/j.1095-8649.1979.tb03612.x
149. Thisse B, Heyer V, Lux A, et al. Spatial and Temporal Expression of the Zebrafish Genome by Large-Scale In Situ Hybridization Screening. In: ; 2004:505-519. doi:10.1016/S0091-679X(04)77027-2

150. Tábara LC, Burr SP, Frison M, et al. MTFP1 controls mitochondrial fusion to regulate inner membrane quality control and maintain mtDNA levels. *Cell*. 2024;187(14):3619-3637. e27. doi:10.1016/j.cell.2024.05.017
151. Sur A, Wang Y, Capar P, Margolin G, Prochaska MK, Farrell JA. Single-cell analysis of shared signatures and transcriptional diversity during zebrafish development. *Dev Cell*. 2023;58(24):3028-3047. e12. doi:10.1016/j.devcel.2023.11.001
152. Farrell JA, Wang Y, Riesenfeld SJ, Shekhar K, Regev A, Schier AF. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science (1979)*. 2018;360(6392). doi:10.1126/science.aar3131
153. Jiang M, Xiao Y, E W, et al. Characterization of the Zebrafish Cell Landscape at Single-Cell Resolution. *Front Cell Dev Biol*. 2021;9. doi:10.3389/fcell.2021.743421
154. Sposato AL, Llewellyn DR, Weber JM, et al. Germ cells do not progress through spermatogenesis in the infertile zebrafish testis. Published online September 5, 2023. doi:10.1101/2023.09.05.556432
155. Liu Y, Kossack ME, McFaul ME, et al. Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary. *Elife*. 2022;11. doi:10.7554/eLife.76014
156. Krøvel AV, Olsen LC. Expression of a vas::EGFP transgene in primordial germ cells of the zebrafish. *Mech Dev*. 2002;116(1-2):141-150. doi:10.1016/S0925-4773(02)00154-5
157. Hartung O, Forbes MM, Marlow FL. Zebrafish vasa is required for germ-cell differentiation and maintenance. *Mol Reprod Dev*. 2014;81(10):946-961. doi:10.1002/mrd.22414
158. Raz E. Primordial germ-cell development: the zebrafish perspective. *Nat Rev Genet*. 2003;4(9):690-700. doi:10.1038/nrg1154
159. Houwing S, Kamminga LM, Berezikov E, et al. A Role for Piwi and piRNAs in Germ Cell Maintenance and Transposon Silencing in Zebrafish. *Cell*. 2007;129(1):69-82. doi:10.1016/j.cell.2007.03.026
160. Ye M, Chen Y. Zebrafish as an emerging model to study gonad development. *Comput Struct Biotechnol J*. 2020;18:2373-2380. doi:10.1016/j.csbj.2020.08.025
161. Chen W, Ge W. Gonad differentiation and puberty onset in the zebrafish: Evidence for the dependence of puberty onset on body growth but not age in females. *Mol Reprod Dev*. 2013;80(5):384-392. doi:10.1002/mrd.22172
162. Dai X, Shu Y, Lou Q, et al. Tdrd12 Is Essential for Germ Cell Development and Maintenance in Zebrafish. *Int J Mol Sci*. 2017;18(6). doi:10.3390/ijms18061127
163. Pathak R, Bhangu SK, Martin GJO, Separovic F, Ashokkumar M. Ultrasound-induced protein restructuring and ordered aggregation to form amyloid crystals. *European Biophysics Journal*. 2022;51(4-5):335-352. doi:10.1007/s00249-022-01601-4
164. Tian R, Feng J, Huang G, et al. Ultrasound driven conformational and physicochemical changes of soy protein hydrolysates. *Ultrason Sonochem*. 2020;68:105202. doi:10.1016/j.ultsonch.2020.105202
165. Zhu X, Kumar R, Mandal M, et al. Cell cycle-dependent modulation of telomerase activity in tumor cells. *Proc Natl Acad Sci U S A*. 1996;93(12):6091-6095. doi:10.1073/pnas.93.12.6091
166. Prince VE, Price AL, Ho RK. Hox gene expression reveals regionalization along the anteroposterior axis of the zebrafish notochord. *Dev Genes Evol*. 1998;208(9):517-522. doi:10.1007/s004270050210
167. Stemple DL. Structure and function of the notochord: an essential organ for chordate development. *Development*. 2005;132(11):2503-2512. doi:10.1242/dev.01812
168. Talbot WS, Trevarrow B, Halpern ME, et al. A homeobox gene essential for zebrafish notochord development. *Nature*. 1995;378(6553):150-157. doi:10.1038/378150a0
169. Szeto DP, Griffin KJP, Kimelman D. hrT is required for cardiovascular development in zebrafish. *Development*. 2002;129(21):5093-5101. doi:10.1242/dev.129.21.5093
170. Stemple DL, Solnica-Krezel L, Zwartkruis F, et al. Mutations affecting development of the notochord in zebrafish. *Development*. 1996;123(1):117-128. doi:10.1242/dev.123.1.117
171. Odenthal J, Haffter P, Vogelsang E, et al. Mutations affecting the formation of the notochord in the zebrafish, *Danio rerio*. *Development*. 1996;123(1):103-115. doi:10.1242/dev.123.1.103
172. Wopat S, Bagwell J, Sumigray KD, et al. Spine Patterning Is Guided by Segmentation of the Notochord Sheath. *Cell Rep*. 2018;22(8):2026-2038. doi:10.1016/j.celrep.2018.01.084
173. Bagwell J, Norman J, Ellis K, et al. Notochord vacuoles absorb compressive bone growth during zebrafish spine formation. *Elife*. 2020;9. doi:10.7554/eLife.51221
174. Schoels M, Zhuang M, Fahrner A, et al. Single-cell mRNA profiling reveals changes in solute carrier expression and suggests a metabolic switch during zebrafish pronephros development. *American Journal of Physiology-Renal Physiology*. 2021;320(5):F826-F837. doi:10.1152/ajprenal.00610.2020
175. Wingert RA, Davidson AJ. The zebrafish pronephros: A model to study nephron segmentation. *Kidney Int*. 2008;73(10):1120-1127. doi:10.1038/ki.2008.37

176. Wingert RA, Davidson AJ. Zebrafish nephrogenesis involves dynamic spatiotemporal expression changes in renal progenitors and essential signals from retinoic acid and irx3b. *Developmental Dynamics*. 2011;240(8):2011-2027. doi:10.1002/dvdy.22691
177. Li Y, Cheng CN, Verdun VA, Wingert RA. Zebrafish nephrogenesis is regulated by interactions between retinoic acid, mecom, and Notch signaling. *Dev Biol*. 2014;386(1):111-122. doi:10.1016/j.ydbio.2013.11.021
178. Nguyen TK, Petrikas M, Chambers BE, Wingert RA. Principles of Zebrafish Nephron Segment Development. *J Dev Biol*. 2023;11(1). doi:10.3390/jdb11010014
179. Naylor RW, Qubisi SS, Davidson AJ. Zebrafish Pronephros Development. In: ; 2017:27-53. doi:10.1007/978-3-319-51436-9_2
180. Naylor RW, Przepiorski A, Ren Q, Yu J, Davidson AJ. HNF1 β is essential for nephron segmentation during nephrogenesis. *J Am Soc Nephrol*. 2013;24(1):77-87. doi:10.1681/ASN.2012070756
181. McCampbell KK, Wingert RA. New tides: using zebrafish to study renal regeneration. *Translational Research*. 2014;163(2):109-122. doi:10.1016/j.trsl.2013.10.003
182. Song HD, Sun XJ, Deng M, et al. Hematopoietic gene expression profile in zebrafish kidney marrow. *Proceedings of the National Academy of Sciences*. 2004;101(46):16240-16245. doi:10.1073/pnas.0407241101
183. Chen AT, Zon LI. Zebrafish blood stem cells. *J Cell Biochem*. 2009;108(1):35-42. doi:10.1002/jcb.22251
184. Wu L, Zhang C, Zhang J. HMBOX1 negatively regulates NK cell functions by suppressing the NKG2D/DAP10 signaling pathway. *Cell Mol Immunol*. 2011;8(5):433-440. doi:10.1038/cmi.2011.20
185. Robertson AL, Avagyan S, Gansner JM, Zon LI. Understanding the regulation of vertebrate hematopoiesis and blood disorders – big lessons from a small fish. *FEBS Lett*. 2016;590(22):4016-4033. doi:10.1002/1873-3468.12415
186. Davidson AJ, Zon LI. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene*. 2004;23(43):7233-7246. doi:10.1038/sj.onc.1207943
187. Siderakis M, Tarsounas M. Telomere regulation and function during meiosis. *Chromosome Res*. 2007;15(5):667-679. doi:10.1007/s10577-007-1149-7
188. Uchida D, Yamashita M, Kitano T, Iguchi T. Oocyte apoptosis during the transition from ovary-like tissue to testes during sex differentiation of juvenile zebrafish. *Journal of Experimental Biology*. 2002;205(6):711-718. doi:10.1242/jeb.205.6.711
189. Tondera D, Czauderna F, Paulick K, Schwarzer R, Kaufmann J, Santel A. The mitochondrial protein MTP18 contributes to mitochondrial fission in mammalian cells. *J Cell Sci*. 2005;118(14):3049-3059. doi:10.1242/jcs.02415
190. Popgeorgiev N, Bonneau B, Ferri KF, Prudent J, Thibaut J, Gillet G. The Apoptotic Regulator Nr3 Controls Cytoskeletal Dynamics via the Regulation of Ca²⁺ Trafficking in the Zebrafish Blastula. *Dev Cell*. 2011;20(5):663-676. doi:10.1016/j.devcel.2011.03.016
191. Donnarumma E, Kohlhaas M, Vimont E, et al. Mitochondrial Fission Process 1 controls inner membrane integrity and protects against heart failure. *Nat Commun*. 2022;13(1):6634. doi:10.1038/s41467-022-34316-3
192. Patitucci C, Hernández-Camacho JD, Vimont E, et al. Mtfp1 ablation enhances mitochondrial respiration and protects against hepatic steatosis. *Nat Commun*. 2023;14(1):8474. doi:10.1038/s41467-023-44143-9
193. Xiao T, Sun J, Xing Z, Xie F, Yang L, Ding W. MTFP1 overexpression promotes the growth of oral squamous cell carcinoma by inducing ROS production. *Cell Biol Int*. 2020;44(3):821-829. doi:10.1002/cbin.11278
194. van der Blik AM, Shen Q, Kawajiri S. Mechanisms of mitochondrial fission and fusion. *Cold Spring Harb Perspect Biol*. 2013;5(6). doi:10.1101/cshperspect.a011072
195. Nolfi-Donagan D, Braganza A, Shiva S. Mitochondrial electron transport chain: Oxidative phosphorylation, oxidant production, and methods of measurement. *Redox Biol*. 2020;37:101674. doi:10.1016/j.redox.2020.101674
196. Liu YJ, McIntyre RL, Janssens GE, Houtkooper RH. Mitochondrial fission and fusion: A dynamic role in aging and potential target for age-related disease. *Mech Ageing Dev*. 2020;186:111212. doi:10.1016/j.mad.2020.111212
197. Handy DE, Loscalzo J. Redox regulation of mitochondrial function. *Antioxid Redox Signal*. 2012;16(11):1323-1367. doi:10.1089/ars.2011.4123
198. Wang C, Youle RJ. The role of mitochondria in apoptosis. *Annu Rev Genet*. 2009;43:95-118. doi:10.1146/annurev-genet-102108-134850
199. Westermann B. Mitochondrial fusion and fission in cell life and death. *Nat Rev Mol Cell Biol*. 2010;11(12):872-884. doi:10.1038/nrm3013

200. Renaudin X. Reactive oxygen species and DNA damage response in cancer. In: ; 2021:139-161. doi:10.1016/bs.ircmb.2021.04.001
201. Gonzalez-Hunt CP, Wadhwa M, Sanders LH. DNA damage by oxidative stress: Measurement strategies for two genomes. *Curr Opin Toxicol.* 2018;7:87-94. doi:10.1016/j.cotox.2017.11.001
202. Singh A, Kukreti R, Saso L, Kukreti S. Oxidative Stress: Role and Response of Short Guanine Tracts at Genomic Locations. *Int J Mol Sci.* 2019;20(17). doi:10.3390/ijms20174258
203. Morita M, Prudent J, Basu K, et al. mTOR Controls Mitochondrial Dynamics and Cell Survival via MTFP1. *Mol Cell.* 2017;67(6):922-935.e5. doi:10.1016/j.molcel.2017.08.013
204. Papadopoli D, Boulay K, Kazak L, et al. mTOR as a central regulator of lifespan and aging. *F1000Res.* 2019;8. doi:10.12688/f1000research.17196.1
205. Walsh ME, Shi Y, Van Remmen H. The effects of dietary restriction on oxidative stress in rodents. *Free Radic Biol Med.* 2014;66:88-99. doi:10.1016/j.freeradbiomed.2013.05.037
206. López-Torres M, Gredilla R, Sanz A, Barja G. Influence of aging and long-term caloric restriction on oxygen radical generation and oxidative DNA damage in rat liver mitochondria. *Free Radic Biol Med.* 2002;32(9):882-889. doi:10.1016/S0891-5849(02)00773-6
207. Zheng Q, Huang J, Wang G. Mitochondria, Telomeres and Telomerase Subunits. *Front Cell Dev Biol.* 2019;7. doi:10.3389/fcell.2019.00274
208. Sahin E, Colla S, Liesa M, et al. Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature.* 2011;470(7334):359-365. doi:10.1038/nature09787
209. Singhapol C, Pal D, Czapiewski R, Porika M, Nelson G, Saretzki GC. Mitochondrial Telomerase Protects Cancer Cells from Nuclear DNA Damage and Apoptosis. *PLoS One.* 2013;8(1):e52989. doi:10.1371/journal.pone.0052989
210. Haendeler J, Hoffmann J, Diehl JF, et al. Antioxidants Inhibit Nuclear Export of Telomerase Reverse Transcriptase and Delay Replicative Senescence of Endothelial Cells. *Circ Res.* 2004;94(6):768-775. doi:10.1161/01.RES.0000121104.05977.F3
211. Scholz J, Besir H, Strasser C, Suppmann S. A new method to customize protein expression vectors for fast, efficient and background free parallel cloning. *BMC Biotechnol.* 2013;13(1):12. doi:10.1186/1472-6750-13-12
212. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc.* 2013;8(11):2281-2308. doi:10.1038/nprot.2013.143
213. Aleström P, D'Angelo L, Midtlyng PJ, et al. Zebrafish: Housing and husbandry recommendations. *Lab Anim.* 2020;54(3):213-224. doi:10.1177/0023677219869037
214. Hwang WY, Fu Y, Reyon D, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol.* 2013;31(3):227-229. doi:10.1038/nbt.2501
215. Green MR, Sambrook J. Precipitation of RNA with Ethanol. *Cold Spring Harb Protoc.* 2020;2020(3):pdb.prot101717. doi:10.1101/pdb.prot101717
216. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.* 2008;26(12):1367-1372. doi:10.1038/nbt.1511
217. Fradera-Sola A, Nischwitz E, Bayer ME, Luck K, Butter F. RNA-dependent interactome allows network-based assignment of RNA-binding protein function. *Nucleic Acids Res.* 2023;51(10):5162-5176. doi:10.1093/nar/gkad245
218. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21. doi:10.1093/bioinformatics/bts635
219. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923-930. doi:10.1093/bioinformatics/btt656
220. Ramírez F, Ryan DP, Grüning B, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 2016;44(W1):W160-W165. doi:10.1093/nar/gkw257
221. Hahne F, Ivanek R. Visualizing Genomic Data Using Gviz and Bioconductor. In: ; 2016:335-351. doi:10.1007/978-1-4939-3578-9_16

Resume

(Blank)

(Blank)

(Blank)