Understanding the role of promoter priming prior to cell fate specification in zebrafish (*Danio rerio*)

A Thesis

Submitted to

Indian Institute of Science Education and Research Pune in partial fulfillment of the requirements of the BS-MS Dual Degree Programme

by

M Sarath

20181038

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

April, 2023

Under the guidance of

Supervisor: Prof. Dr. Sanjeev Galande,

From May 2022 to Mar 2023

CERTIFICATE

This is to certify that this dissertation entitled "**Understanding the role of promoter priming prior to cell fate specification in zebrafish**" towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune, represents study/work carried out by "**M Sarath** at the Indian Institute of Science Education and Research, Pune, under the supervision of **Prof. Dr. Sanjeev Galande**, Department of Biology during the academic year 2022-23.

Prof. Sanjeev Galande

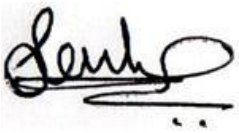
Committee:

Prof. Dr. Sanjeev Galande

Dr. Krishanpal Karmodiya

DECLARATION

I hereby declare that the matter embodied in the report entitled " **Understanding the role of promoter priming prior to cell fate specification in zebrafish**" is the result of the work carried out by me at the Department of Biology, Indian Institute of Science Education and Research, Pune, under the supervision of **Prof. Sanjeev Galande**, and the same has not been submitted elsewhere for any other degree.



M Sarath

10-04-2023

CONTENTS

ABBREVIATIONS

DNA	Deoxyribonucleic Acid
RNA	Ribonucleic Acid
hpf	Hours Post Fertilization
dpf	Days Post Fertilization
tpm	Transcripts Per Million
BSA	Bovine SerμM Albumin
gRNA	guide RNA
WT	Wild Type
PBS	Phosphate Buffered Saline
TBST	Tris Buffered Saline with 0.1% Tween 20 detergent
ChIP	Chromatin Immunoprecipitation
PFA	ParaFormaldehyde
qPCR	Quantitative Polymerase Chain Reaction
MLL	Mixed Lineage Leukemia
WDR5	WD Repeat Domain 5
AML	Acute Myeloid Leukemia
IGV	Integrative Genomve Viewer
KDM	Lysine demethylase

KMT	Lysine methyltransferase
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cas9	CRISPR-associated protein 9
dCas9	Deactivated Cas9
H3K4me1	Histone H3 lysine 4 monomethylation
H3K4me3	Histone H3 lysine 4 trimethylation
H3K27me3	Histone H3 lysine 27 trimethylation
ZGA	Zygotic Genome Activation
TALEN	Transcription activator-like effector nucleases
ZFN	Zinc Finger Nuclease
PHD	Plant Homeodomain

ABSTRACT

Understanding the regulation of gene expression in the context of development is important as it helps in comprehending the processes by which cells differentiate and develop into specific tissues and organs. Cis-regulatory regions play an important role in regulating gene expression by altering the efficiency of transcription machinery and the stability of transcripts. Promoter priming is a regulatory mechanism by which promoter regions of genes are pre-occupied with H3K4me3 modifications so that they are able to maintain proper temporal regulated expression of specific genes. The mechanism and the functional significance of this mark in regulating gene expression are still not known, and our study aims to understand this. The presence of specific histone modifications was tried to be validated using ChIP-qPCR. Their functional significance was explored in the context of cell fate specification. Drug based perturbation study was carried out to understand the significance of priming using MM-102 drug. The effect of the drug on various stages and its implications on gene expression, and thereby the development of the embryo, was explored using western blot and q-PCR. The study makes an attempt to understand the role of priming marks in regulating the precise expression of genes at respective time points of development in zebrafish.

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

I thank Prof. Sanjeev Galande for giving me the opportunity to work in the lab. I am thankful for his time and input during various experiments and during thesis writing. I would also like to thank my expert member, Dr. Krishanpal Karmodiya, for his time and suggestions during the project. I am thankful to my mentor, Greg Dsilva, for teaching me various techniques, providing scientific and technical inputs throughout the course of the project, and answering my multitude of questions. I would like to thank Ankita Sharma for teaching me ChIP-related techniques and for providing various suggestions in shaping the direction of the study. I cannot imagine a better group of people to work with. I would like to acknowledge the staff of the fish facility- Sameer and Aishwarya. I am thankful to SG and KK lab members for providing a friendly lab environment. I'll always be thankful to my parents and my friends for being there and for motivating me throughout.

CONTRIBUTIONS

Contributor name	Contributor role
M Sarath, Prof. Sanjeev Galande, Greg Dsilva, Ankita Sharma	Conceptualization Ideas
M Sarath, Greg Dsilva, Ankita Sharma	Methodology
M Sarath	Software
M Sarath	Validation
M Sarath	Formal analysis
M Sarath	Investigation
SG lab, Fish facility	Resources
—	Data Curation
M Sarath	Writing - original draft preparation
M Sarath, Prof. Sanjeev Galande, Greg Dsilva	Writing - review and editing
M Sarath	Visualization
Prof. Sanjeev Galande, Greg Dsilva	Supervision
—	Project administration
Prof. Sanjeev Galande	Funding acquisition

INTRODUCTION

Gene regulatory mechanisms during development

Animal development is orchestrated by spatiotemporal gene expression programmes that drive precise lineage commitment, proliferation, and migration events leading to large morphological change and functional specification in the organism (Dong & Liu, 2017). The specification of distinct cell types is established by precise control of gene expression, which is primarily achieved through the active utilization of cis-regulatory modules and by trans-acting transcription factors, as shown in Figure 1a. Cis-regulatory elements are broadly classified into promoters, enhancers, and insulators. A detailed understanding of each CRE will help us understand the regulatory programs that control important developmental processes like cell fate specification. Promoters are regions of DNA upstream of a gene that acts as docking sites for important proteins such as the RNA polymerase and transcription factors that initiate the transcription of that gene(Santos-Rosa et al., 2002). Promoters have been shown to direct spatiotemporal patterning in cells by interacting with distinct transcription factors and other complexes which help in transcriptional regulation. As shown in Figure 1B (Preissl et al., 2023), the functional state of cis-regulatory regions correlates with the modification state of the histones. Promoter regions are usually marked with Histone H3 lysine 4 trimethylation mark. An activated state of promoter is further marked by the presence of the Histone h3 lysine 27 acetylation mark, and a poised state is marked by the H3K27me3 mark.

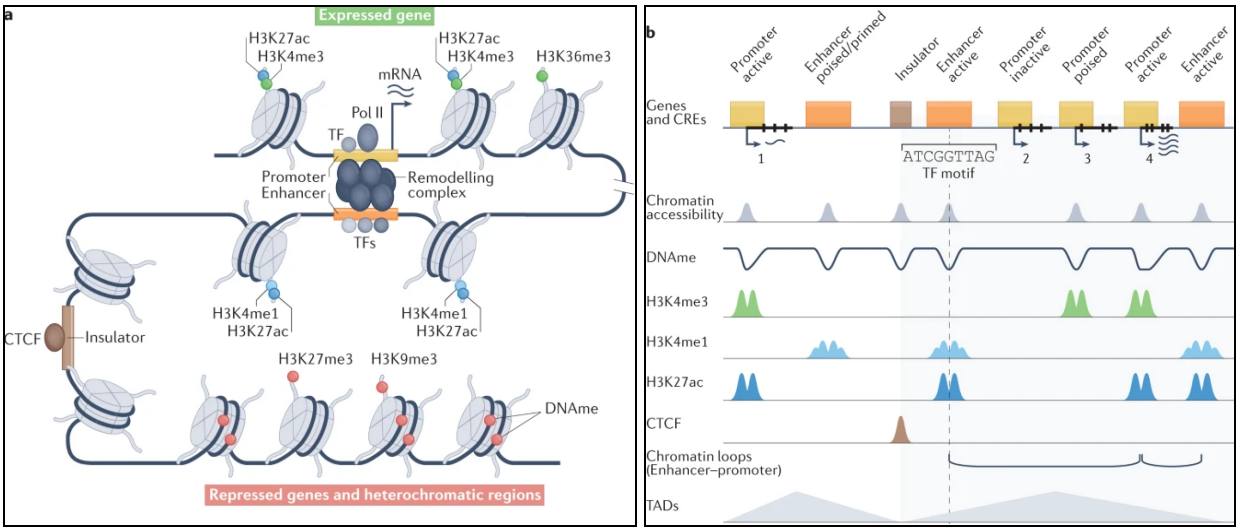


Figure 1: **Gene expression regulation model and functional state of cis regulatory regions based on modifications over them.** a) Representative model of gene expression regulation (Preissl et al., 2023). Distinct chromatin modifications can be used to identify the activity of cis-regulatory elements (CREs) and gene regions. Chromatin accessibility is high, while DNA methylation (DNAme) levels are low in the promoters of expressed genes. Additionally, active promoters exhibit high levels of histone H3 trimethylated at lysine 4 (H3K4me3) and acetylated lysine (H3K27ac) residue marks. b) Epigenetic features associated with different CREs as visualized in Integrated Genome (IGV). For example, active promoters (Genes 1 and 4) display a strong signal for H3K4me3 and H3K27ac, while active enhancers exhibit a strong signal for H3K4me1 and H3K27ac. Primed/poised promoters (Gene 3) are characterized by a strong H3K4me3 signal, while inactive promoters (Gene 2) lack H3K4me3.

Priming of Cis Regulatory Elements

Histone modifications play a critical role in regulating the DNA-histone interactions and thereby regulate the recruitment of specific proteins, which plays a role in the activation or repression of target genes. Priming is a regulatory mechanism that ensures timely and appropriate expression of genes at the correct level in a cell type-specific manner(Bonifer & Cockerill, 2017), as shown in Figure 2. Chromatin can be modified with certain transcription factors or histone modification marks that regulate the expression of the target genes. Chromatin priming has been shown to play a role in

lineage-specific cell type activation. GATA2, a transcription factor, acts as a priming mark for enhancers related to erythroid differentiation (Huang et al., 2016). Similarly, FOXA2 has been shown to be a requirement of enhancer priming by deposition of the H3K4me1 mark during pancreatic differentiation (Lee et al., 2019). Promoter priming refers to the regulatory mechanism where promoters are marked with H3K4me3 marks prior to the developmental state where the genes controlled by them are activated with further marks like H3K27ac. Lindeman et al (Lindeman et al., 2011) have earlier reported the presence of priming of the zebrafish genome by H3K4me3 prior to ZGA in genes related to homeostatic and developmental regulation. Promoters have been shown to be primed early prior to mid-blastula transition and are later activated temporally throughout the process of development. The data from Lindeman et al. suggest an instructive role that epigenetic modifications could play in regulating gene expression during development.

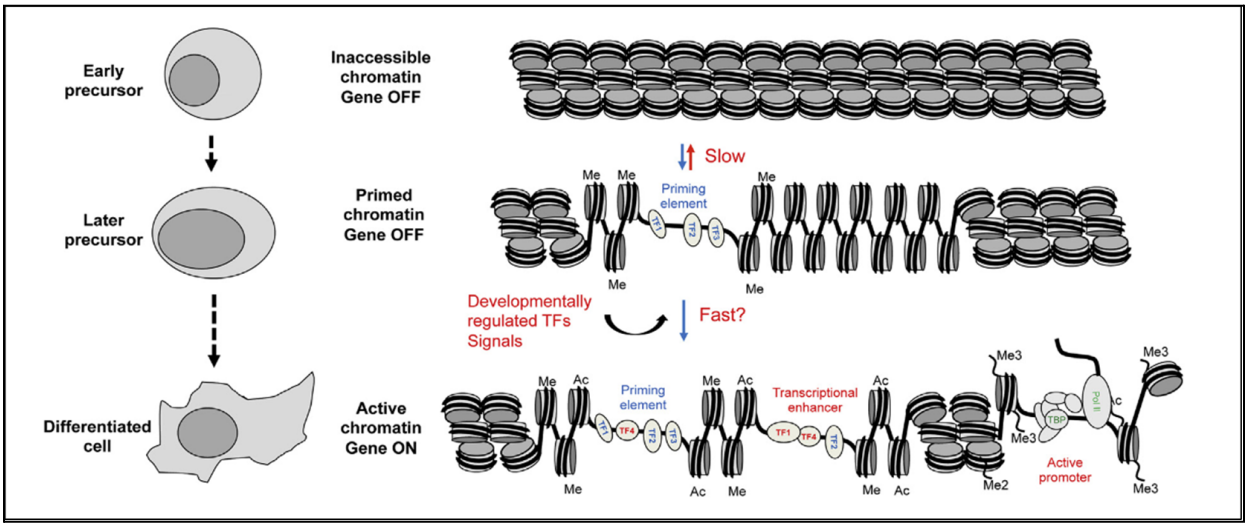


Figure 2: **Model showing how chromatin priming works**. In earlier stages of development, chromatin is compressed, and genes are in an off state; At a certain particular stage of development, they get primed by the addition of histone methylation marks, which slightly open up the chromatin, though the genes are off; At the stage of differentiation, other histone modification marks and activatory transcription factors bind to these primed regions and activate the genes near them rapidly (Bonifer et al. Experimental Hematology, 2017).

Priming in cell fate specification

Previous studies on cell fate specification have primarily focused on the role of morphogens as pre-patterning regulators, while the influence of temporal histone modifications remains an underexplored area. Further research in this direction could shed light on the mechanisms underlying cell fate determination. Histone modification has been shown to play a role in pre-patterning in establishing the anterior-posterior axis and, thereby, the regional identity of organisms (Halsall et al., 2021; K. Han et al., 2019). Enrichment of H3K27me3 has been shown in genes that are important for posterior identity, while H3K4me3 enrichment is important for anterior identity. Histone modifications also regulate the expression of key transcription factors which play crucial roles in fate specification (Akkers et al., 2009). Enrichment of specific histone modification marks has been shown to be acting as a pre-patterning condition for specific lineages. H3K4me2 is enriched in genes that are important for T cell lineage specification, while H3K27me3 patterns for other lineages during the differentiation of hematopoietic stem cells (Fanucchi et al., 2021; Tripathi & Lahesmaa, 2014). These modifications regulate the expression of critical transcription factors like TCF-1, which promotes the development of T cells. Similarly, neural differentiation into specific lineages is controlled by the histone modification marks of H3K27ac and H3k27me3. Neural fate-regulating transcription factors, NeuroD1 and Neurog2, are regulated by changes in histone modification patterns. (Park et al., 2022) Thus histone modifications help in regulating the expression of lineage-specific genes by regulating their accessibility and recruiting specific proteins, thereby regulating various cellular processes.

H3K4me3 modification as a priming mark for promoter regions

Methylation mark is added to Histone H3 by a family of enzymes called histone methyltransferases (HMTs). Depending on the location of the modification added, it can activate or repress the chromatin region. H3K4me3 is a well-studied histone modification mark associated with active promoter regions. It is typically found near the transcription start site (TSS) of actively transcribed genes. It can recruit and bind to specific proteins that can facilitate the recruitment of RNA Polymerase II and other

transcription factors to the region, thereby increasing the accessibility and active transcription in the region. There are several different types of H3K4me3 writers, known as Histone methyltransferases (HMTs), involving SET domain-containing proteins, MLL family members, and NSD family members, which transfer a methyl group from S-adenosyl methionine (SAM) to the lysine residue on histone H3.

H3K4me3 is reported to act as a prepattern mark by marking genes that can get activated in specific regions at specific time points in response to signaling pathways during development. By marking these genes with H3K4me3, the chromatin is primed for transcription, which enables rapid activation of these genes in response to developmental signals. Setd1b is an H3K4me3 writer which has been studied earlier in the context of priming. Its activity was reported to be providing preconditioning that helps in the rapid activation of immune response genes upon infection in zebrafish. MLL1 is another writer which has been reported to prime the immune response genes in mouse macrophages against lipopolysaccharide. In the context of development, MLL1 is majorly expressed in the blastula stage, and its importance has been studied in the context of mesoderm development. Thus the activity of these writers in the context of developmental processes like cell fate specification needs to be investigated. H3K4me3 modification serves as a docking site for various transcription factors and other regulators, which helps in altering the chromatin structure and allows the transcription machinery to proceed. Chromatin remodeling complexes like SWI/SNF interact with these marks and help in sliding the nucleosomes along the DNA, thereby exposing the methylated. regions for transcriptional activation. Reports have also suggested that histone acetyltransferases (HATs) and other coactivators leading to further activation of the chromatin region are also recruited by H3k4me3 modification. The addition of HATs to these primed regions neutralizes the positive charge on the lysine residues, increasing the chromatin accessibility for transcription machinery and increasing the expression of target genes. This implies the priming role that H3K4me3 modification could play in gene expression across development. It also helps in the establishment of regional identity in an organism. Histone modifications also play an important role in recruiting transcription factors which play a significant role in cell fate specification.

Specific modifications can promote or repress the progression of cells into specific lineages.

How the H3K4me3 marks act as a priming mark in development and what are the upstream signals and molecular pathways that regulate its deposition and maintenance in specific regions of the genome are some of the questions that need to be explored further. What are the differences and similarities in the roles of H3K4me3 in different cell types and tissues, and how do these differences contribute to the diversity of cell types and functions in multicellular organisms? How does H3K4me3 contribute to the maintenance of cellular identity, and what are the consequences of its dysregulation in developmental states?

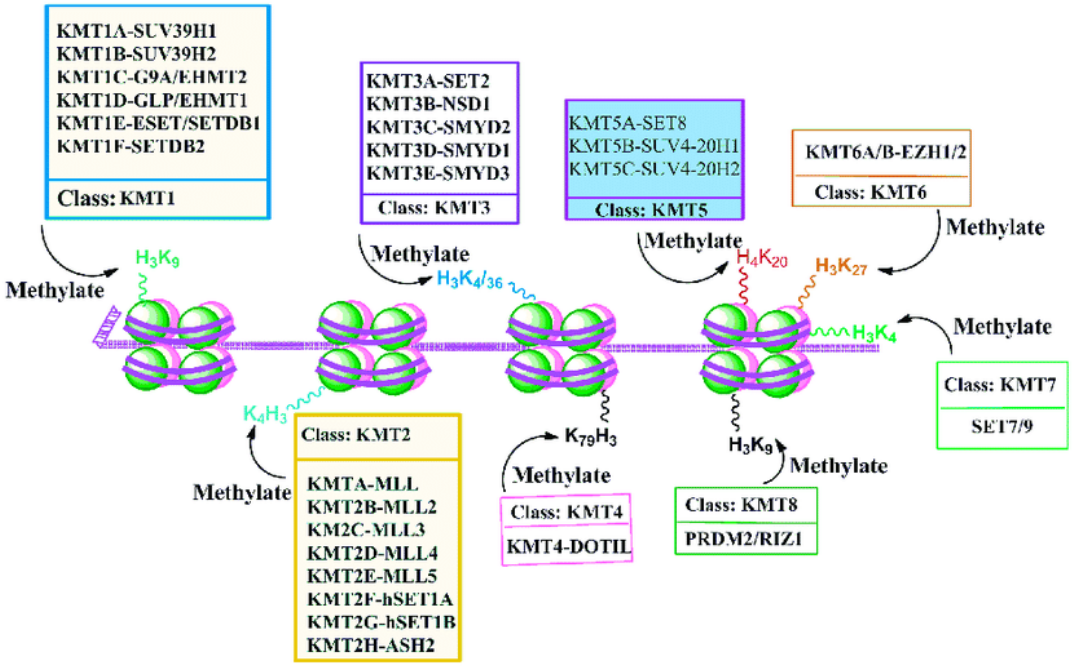
Zebrafish as an ideal model organism for these studies

Zebrafish is an ideal model organism for studying the histone dynamics happening inside the genome in a temporal manner. Priming behavior has been reported earlier in zebrafish. Various lineage tracing and other genetic manipulation experiments have provided insights into the genetic mechanisms happening in cell fate specification in zebrafish (Melby et al., 1996). The shield stage (6hpf) is the time point where the fate maps demarcating the position of precursors of various tissues are present, as shown in Figure 3 (Schier & Talbot, 2005). The embryonic shield structure formed at the blastoderm margin during this stage acts as an organizer during the gastrula stage. The rapid movement and migration of undifferentiated totipotent blastoderm cells, along with the increased nodal and Wnt signaling activity, results in a differentiated and fate-specified state during the shield stage (Economou et al., 2022; Hikasa & Sokol, 2013; Xie et al., 2011). Publicly accessible histone dynamics data is available in this shield stage and the stages earlier to it(dome stage-4.3hpf), making it an attractive model organism to study the priming behavior. Setd1a, Setd1b, Setd8, Setd9, Setd11, Mll1, Mll2 and Smyd2 are some of the H3K4me3 writers expressed in zebrafish. Of these, Setd1a, Setd1b, Mll1, Mll2, and Smyd2 have been reported to express from the very early stages of zebrafish development. This model system is particularly suitable for doing functional studies on these priming marks by perturbation techniques, as its

transparent embryos enable noninvasive observation of organ development and potential toxicity. The zebrafish's developmental time period also allows for the examination of recovery processes.

The goal of this research project is to investigate the role of pre-primed H3K4me3 promoter regions during the dome stage of development in cell fate specification. The study will involve the identification of pre-primed promoter regions that are associated with cell fate specification during the dome stage by analyzing ChIP-seq data. Once identified, the presence of these regions will be validated using experimental techniques. Finally, experiments will be conducted to determine the effects of manipulating the activity or expression of these regions on gene expression and cell fate. By gaining a better understanding of the molecular mechanisms underlying promoter priming in zebrafish, we hope to contribute to our broader understanding of the processes that regulate embryonic development and cell fate decisions in animals, which has significant implications in various fields like developmental biology, regenerative medicine, and tissue engineering.

a)



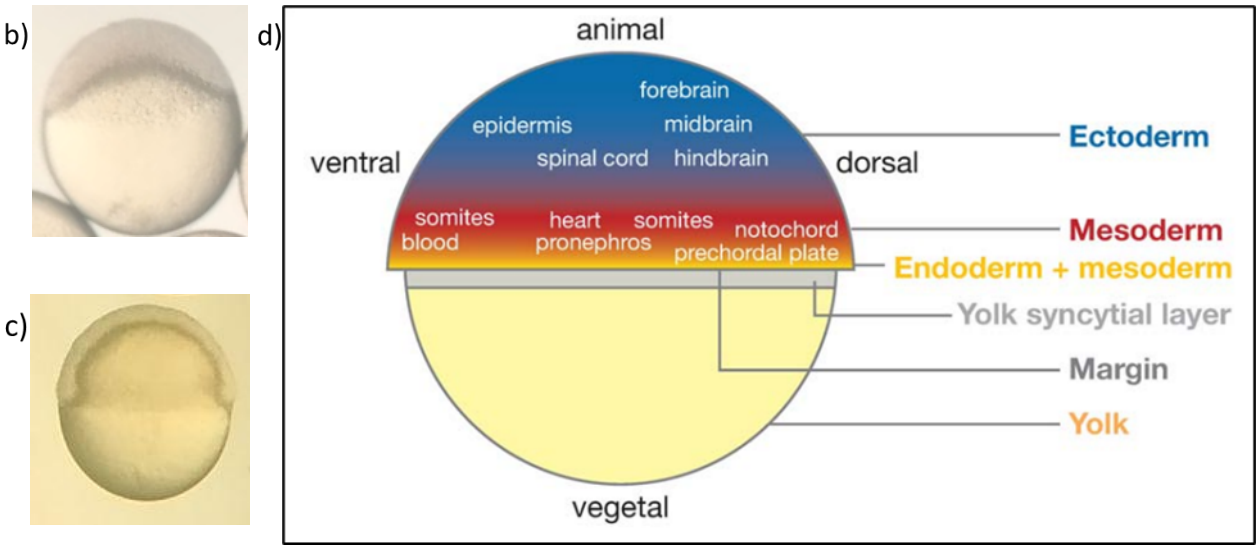


Figure 3: **Histone methyltransferases and cell fate specification in zebrafish embryos.** a) Histone methyltransferases and their target (Basavarajappa & Subbanna, 2016) The enzymes writing the H3K4me3 modification mark are subsets of KMT2 class. b) Dome stage of the zebrafish embryo c)shield stage of the zebrafish embryo, d) Cell fate specification map at shield stage of the embryo (Schier & Talbot, 2005).

AIM AND OBJECTIVES OF THE STUDY

The aim of the project is to understand the role of promoter priming during cell fate specification in zebrafish. the objectives of the study are as follows:

1. Finding pre-primed H3K4me3 regions at dome stage
2. Identifying primed promoter regions that have a role in cell fate specification
3. Validating pre-primed promoter regions
4. Finding the functional significance of them in-cell fate specification.

Materials and Methods

1. **ChIP seq analysis:**

H3K4me3 dome	GSE32483: GSM915189	ChIP
H3K4me3 shield	GSE32483	ChIP
H3K27ac dome	GSE32483: GSM915197	ChIP
H3K27ac shield	GSE32483	ChIP
H3K27me3 dome	GSE70847: GSM1820450	ChIP
H3K27me3 shield	GSE44269: GSM1081557	ChIP

Table 1: Datasets used in the study

Fastq files of the ChIP seq data of H3K4me3 and H3K27ac for dome and shield stage were downloaded from the Skarmeta et al. study from the Sequence read archive (https://www.ncbi.nlm.nih.gov/sra) database. Quality checks of the data were done using fast QC. Bowtie2 was used to align the reads to the danRer11 genome. Samtools was used to sort and convert it to the bam files. Pearson correlation plots, as shown in Figure4, were made using the HOMER tool (Boeva, 2016). It showed a strong relationship between the replicates. Picard and samtools were used to remove duplicate reads. Macs2 was used for peak calling with default

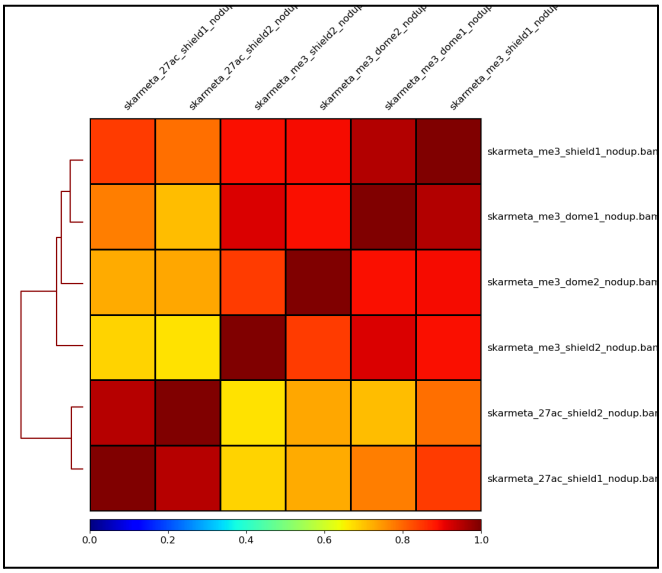


Figure 4: Pearson correlation plot : It shows the correlation between the replicates used. The replicates for respective histone marks can be seen highly correlating with each other.

parameters. Narrow peak calling was done for acetylation marks with a q-value cutoff of 0.05, while broad peak calling was done for methylation marks with a broad cutoff of 0.1. Bedtools intersect and subtract functions were used to find the overlapping peaks and primed peaks, respectively. Integrated Genome Viewer (IGV) was used to visualize the tracks.

2. Gene annotation and Motif Analysis

Homer (http://homer.ucsd.edu/homer/) was used for motif analysis and for peak annotation. Five hundred top-scoring motifs were retrieved using the analysis. Motif analysis was also carried out for regions including 1000 bp upstream and downstream of the peaks.

3. Gene Ontology analysis

Gene ontology analysis was done with the nearby genes to the peaks using the ShinyGO tool(http://bioinformatics.sdstate.edu/go/). Pathway enrichment plots were also made using the same. GeneOntology.org was also used for doing the gene ontology analysis(http://geneontology.org/).

4. Primer Design

NCBI PrimerBlast was used for designing primers. For ChIP qPCR control primer designing, regions visualized from the igv tracks without any peaks were chosen. Sequences were extracted from the genome using the getfasta command of bedtools. They were uploaded to the primer blast website, and primers that gave a product size in the range of 150-300 and melting temperature of around 60 ℃ were chosen. They were synthesized and tested using normal qPCR.

Table 2: Primer Sequences

Target	Primer Sequence	Method
Ccne2_FP	GCACTGGACACTGCGGACAA	qPCR

Ccne2_BP	GGGACTCTTCTATTGCACTCGCC	qPCR
Lamb1a_FP	CAGAGAGCCTGCTGGAACAA	qPCR
Lamb1a_BP	ATGTCTCCGACTCCACCGAA	qPCR
INO80B_FP	CAAGATCAAACTCGGCGGAC	qPCR
INO80B_RP	CCAGGTTACTGTCCTCATCCAG	qPCR
CHCHD2_FP	GCATTGCGAAGGTCAGTAGC	qPCR
CHCHD2_RP	GCTGGTGAGGAGTAACTTGGA	qPCR
SALL1A_FP	ACCACAAAAGGAAATCTCAAGGTGC	qPCR
SAL1A_RP	GACGCATATACTGGAGAGCACA	qPCR
Cntrl1_chr2: 30109567-30109675	TTTCCCATGTAGGTTTCACTCCC	ChIP-qPCR
Cntrl1_chr2: 30109567-30109676	TGGAGGTTGAGACAGAGCGG	ChIP-qPCR
Cntrl2_chr12: 19793024-19793215	GTCTCATTCTAACCGTAGCCGT	ChIP-qPCR
Cntrl2_chr12: 19793024-19793216	GAGGAACATCAGGGGTGAGC	ChIP-qPCR
INO80b_FP1	TATGCTTTCCCTGCAGTGGCT	ChIP-qPCR
INO80b_BP1	AGCTACGGGTCAGTAATGTGC	ChIP-qPCR
INO80b_FP2	AATATGCTTTCCCTGCAGTGGCT	ChIP-qPCR

INO80b_BP2	CGGTAAAGGAAGCTACGGGTC	ChIP-qPCR
Ccne2_FP1	GCTGTCAAACGAGGTAGCGA	ChIP-qPCR
Ccne2_BP1	GCATACTTAAGCTCTGCCGC	ChIP-qPCR
Ccne2_FP2	AGCTGTCAAACGAGGTAGCG	ChIP-qPCR
Ccne2_BP2	CCGCGTCGACCTTTGGTTTA	ChIP-qPCR
Lamb1a_FP1	AGGTACAGTTAGTGAAACACGGATA	ChIP-qPCR
Lamb1a_BP1	CTGCAGTAACATTCGGTTCCTT	ChIP-qPCR
Lamb1a_FP2	TAGGTACAGTTAGTGAAACACGGAT	ChIP-qPCR
Lamb1a_BP2	TGCAGTAACATTCGGTTCCTT	ChIP-qPCR
Sall1a_FP1	AAACTCGGCTGACCGACCTA	ChIP-qPCR
Sall1a_BP1	TTCCCGTGGCTTGACGTTTC	ChIP-qPCR
Sall1a_FP2	TCGAGACGCCGATAGTTGC	ChIP-qPCR
Sall1a_BP2	GAGCGTATCGCATATCCTGGT	ChIP-qPCR
ChChd2_FP1	ACTCAACACGCACTTCTCCAA	ChIP-qPCR
ChChd2_BP1	TGAGCTACGTGCTGTATTATCATGT	ChIP-qPCR
ChChd2_FP2	CAATACTCAACACGCACTTCTCC	ChIP-qPCR

ChChd2_BP2	TCGTGTGTCATTTACTCAACAAACA	ChIP-qPCR

5. Zebrafish maintenance and crosses

Zebrafish were maintained as described previously (Westerfield, 2000). The facility was maintained at 28℃. Crosses were set up the prior night by separating male and female zebrafish with a barrier. On the next day, the barriers were removed, and the sieve was kept at an incline to facilitate breeding. After some time, the fish water was filtered to collect the eggs laid. The embryos were incubated at 23-31 ℃. The strains used in this study are AB, TU, and ABTU.

6. ChIP qPCR

Approximately 500-600 embryos were collected and fixed using 1% formaldehyde in the deyolking buffer and stopped with 0.2 M glycine, followed by PBS washes. Cells were then lysed with a cell lysis buffer(10 mM Tris-Cl pH 8.0, 10 mM NaCl, 0.5% NP40 and PIC) followed by a 15 min spinning at 3000 rpm, 4℃, and a PBS wash. Nuclear lysis buffer(50 mM Tris-Cl pH 8.0, 10 mM EDTA, 1%SDS, 1% NP40, PIC) was added to the pelleted nuclei and was incubated on ice for 10 min. The DNA was fragmented using an S220 Focused-ultrasonicator (Covaris). This was done with a program consisting of seven cycles, each lasting 60 seconds, at 5% duty, 4 intensity, and 200 cycles/burst. After fragmentation, the SDS concentration was lowered to 0.1% using ChIP Dilution Buffer. The solubilized chromatin was then clarified through centrifugation for 15 minutes at 3000 rpm at 4°C and stored at -80℃. A small portion of the sample was set aside as input. The remaining solubilized chromatin was immunoprecipitated using antibodies against H3K27ac (Abcam, Cat#ab4729), H3K4me3 (Millipore, Cat# ab8895), and rabbit IgG. Antibody–chromatin complexes were further pulled down using protein A-Dynabeads. The pulled-down samples were washed three times using a low Salt Buffer (0.1%SDS, 1% TritonX100, 2 mM EDTA, 20 mM TrisCl-pH 8.0, 500 mM NaCl),

followed by one wash with TE (10 mM Tris-Cl pH 8.0, 1 mM EDTA). Finally, the complexes were eluted in 300µL of Elution Buffer (1%SDS, 0.1 M NaHCO3).

Following the reversal of cross-linking (performed overnight at 65°C) and proteinase K digestion, both DNA samples obtained from immunoprecipitation and the input were extracted using phenol-chloroform, ethanol-precipitated, and then treated with RNase.

ChIP qPCR primers were designed for the primed regions using NCBI PrimerBlast. The qPCR reaction was set up using these eluted samples. ViiA7 Software was used to set up the RT-qPCR. Comparative Ct with SyBR green reagents using standard properties was set up in the software. Wells were defined and assigned to the respective samples. 6 µL reaction was set up with 3 µL of TB green mix, 1 µL of NFW, 1µL of primer and 1 µL of template.

7. Drug treatments

20 Wild type dechorionated embryos were added to each of the wells containing specific concentrations of the drug. MM-102 drug ordered from Selleckchem was used for the study (Karatas et al., 2013). Drug Concentrations of 10 µm 50 µm and 100 µM were made by diluting them with E3. Treatment with these concentrations of the MM-102 drug was performed at the 2-4 cell stage. The embryos were incubated at 28℃ and harvested at dome and shield stages. DMSO was used as the vehicle control.

Stage of Development	Concentration of MM-102
Dome(4.3hpf)	100uM
Dome(4.3hpf)	50uM
Dome(4.3hpf)	10uM
Dome(4.3hpf)	0
Shield(6hpf)	50uM
Shield(6hpf)	10uM
Shield(6hpf)	0

Table 3: Concentration of various drugs and the developmental stages used in the drug treatment study.

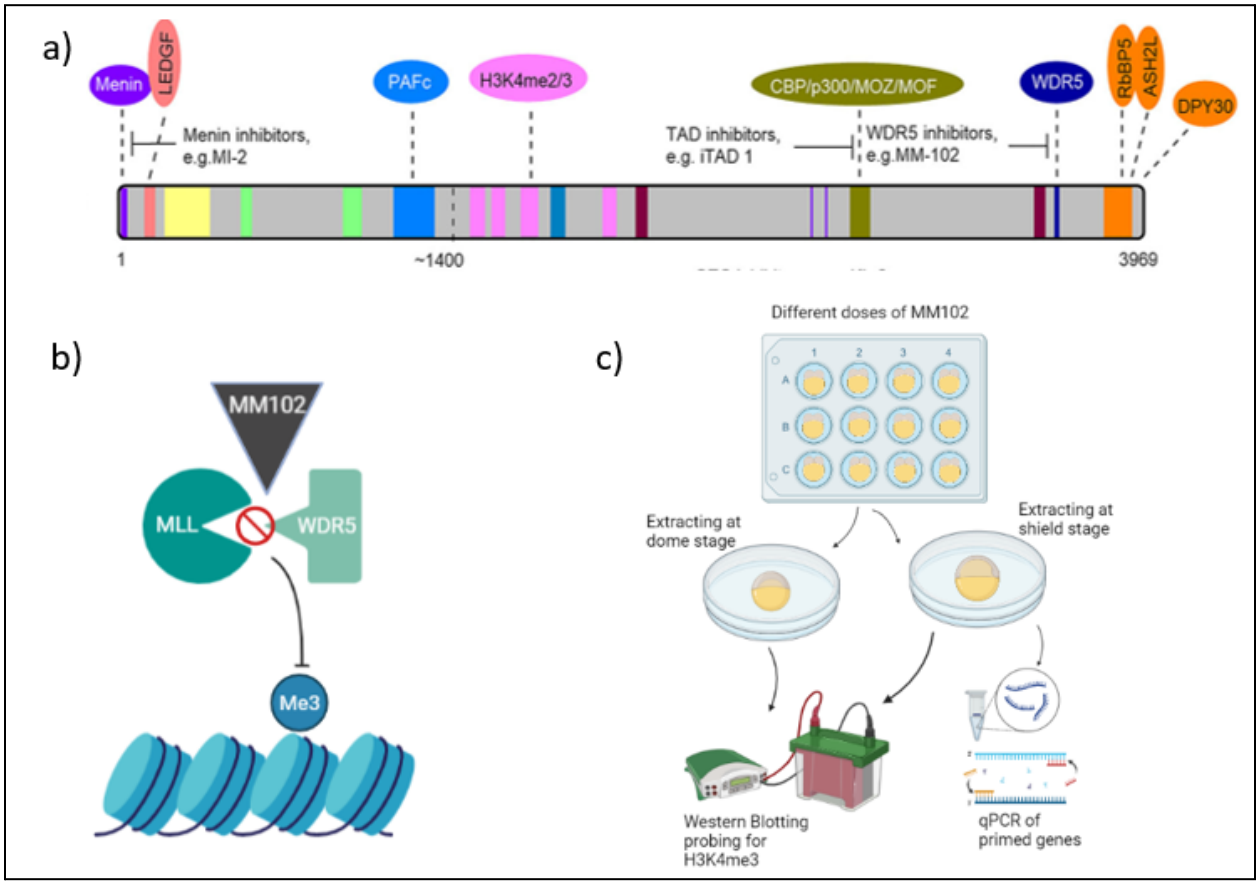


Figure 5: **MM-102 drug treatment study targeting H3K4me3 writer MLL1.** a) Domains of MLL1 protein where various adapter proteins like WDR5 bind for its action (Li & Song, 2021). b) Schematics showing the action of MM-102 drug on H3K4me3 writer c) Schematic representation of the drug treatment study. Embryos treated with MM-102 drug from the 2 cell stage were collected at the dome and shield stage and were processed for western blotting and RT-qPCR.

8. RNA isolation

Dechorionated embryos at dome and shield stages were collected and suspended in TRIzol reagent. Chloroform (1/5th volume) was added, mixed, incubated and centrifuged at 12000 rpm. The aqueous layer was collected, washed again with chloroform. Isopropanol was added and kept overnight to precipitate the RNA. It was centrifuged and the resultant pellet was washed with ethanol, dried, and dissolved in

Nuclease-free water. RNA concentration was determined by Nanodrop and diluted as needed.

9. cDNA preparation

All RNA samples were extracted as mentioned above. cDNA was prepared from 600 ng of RNA using the iScript RT kit (Bio-rad). The cDNA was diluted with NFW in a 1:1 ratio. qPCR was set up using either TB Green II. The quantification was done using the Comparative Ct method (Equation 1), and the values were plotted using GraphPad Prism software. Fold change= 2 (-C T) (Equation 1) where CT= (CTgene - CTinternal control)treated – (CTgene - CTinternal control) control

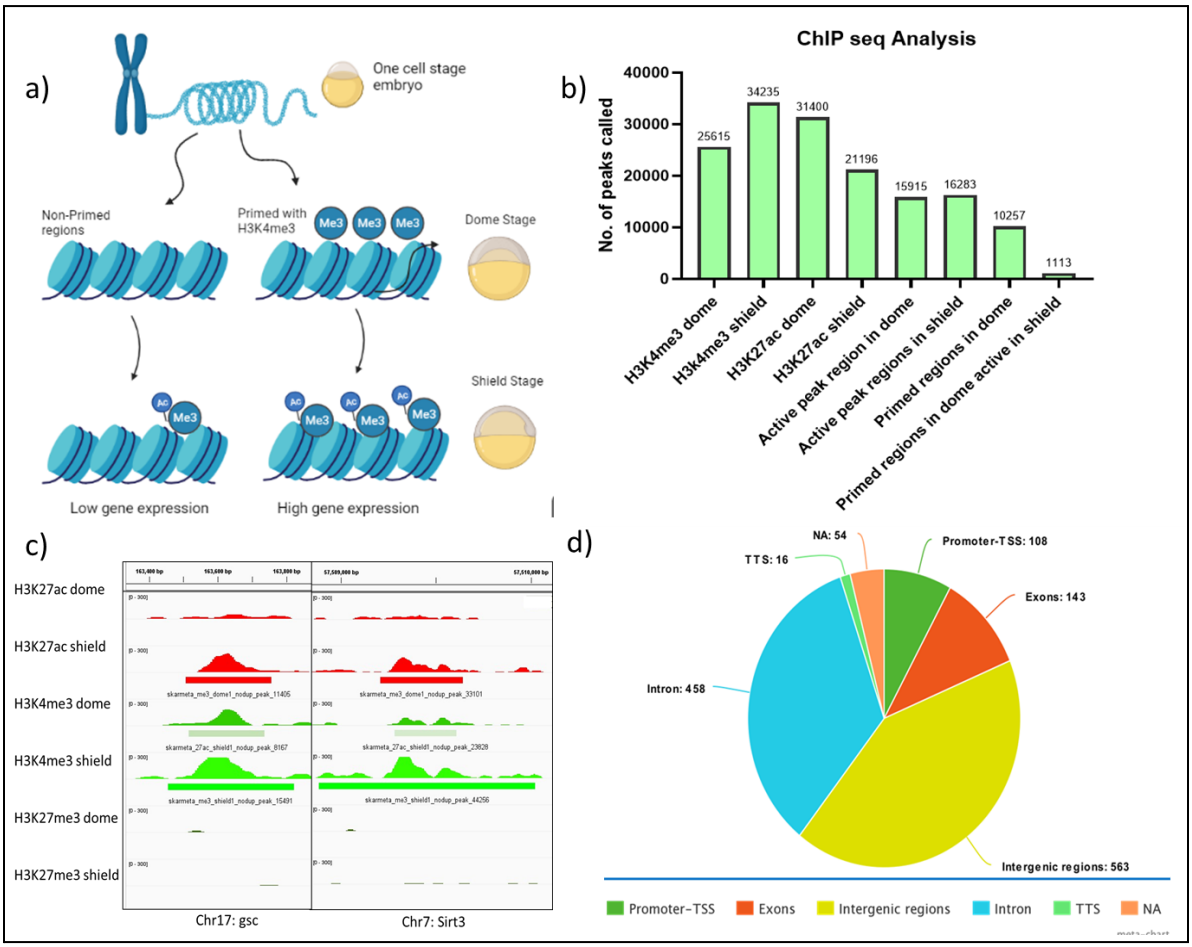
10. Protein Extraction and Western Blotting

Dechorionated embryos were collected at various stages and washed in deyolking buffer (55 mM NaCl, 1.8 mM KCl, 1.25 NaHCO3) to remove the yolk. 6X Laemmli buffer (375 mM Tris-HCl pH 6.8, 10% w/v SDS, 4.8% Glycerol, 10% -mercaptoethanol, 25 0.02% Bromophenol blue) was added to the pellet after deyolking. The sample was boiled at 98°C for 10 minutes and then loaded onto a 15% SDS-PAGE gel. The proteins were then transferred onto a PVDF membrane (Millipore) using wet transfer at 0.6 A for 1.5 hours at 4°C. The membrane was then blocked with 5% Bovine Serum Albumin (BSA) for 1 hour at room temperature. The blot was then incubated with the Millipore H3K4me3 (ab8895) primary antibody for 12-16 hours at 4°C. Gamma tubulin was used as a control. 4 TBST (0.05 M Tris pH 7.4, 0.15 M NaCl, 0.01% Tween 20) washes of 10 minutes each given. The blot was then incubated in HRP conjugated antibody against appropriate species, followed by 4 TBST washes. The membrane was then developed using Clarity ECL Western substrate (Biorad) and imaged using LasQuant 4000. The primary antibodies used in this study are anti-H3K4me3 (ab8895) , H3K27ac (Abcam #ab4729), and anti-Tubulin (Sigma).

Results and Discussion

1. ChIP seq analysis of histone modification data to find primed regions in the dome stage getting activated in the shield stage

1113 regions were found in the genome, which were primed in the dome and got active in the shield stage, as shown in Figure6(b). Examples of peak tracks of two regions, gsc and sirt3, as visualized in igv, are shown in Figure 6(c). These peak regions were annotated, and only 9.7% of the peak regions were lying in the promoter regions. The majority of them were in the intergenic regions, which includes enhancers. The H3K4me3 mark is generally reported to be associated with promoters, but there are studies showing that H3K4me2/3 is present in great amounts in the enhancer regions as well. Thus, though it was expected that the majority of them might lie within the promoter region, enhancer regions and other intergenic regions took the major share of them.



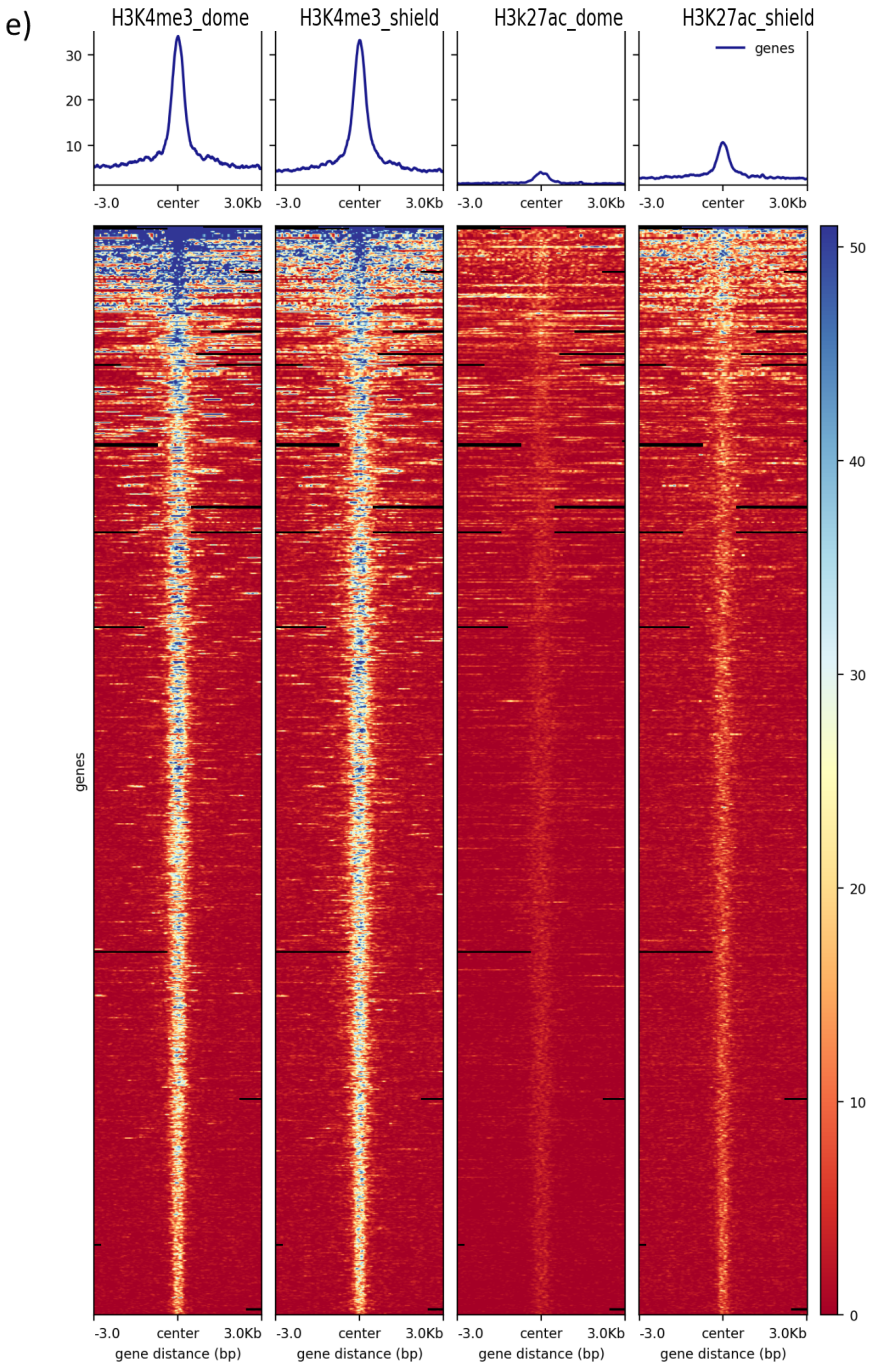


Figure 6: **Representation of ChIP seq analysis data to find dome primed shield active promoter regions.** a) Schematic representing the selection of primed regions b) A bar plot showing the number of peaks obtained in ChIPseq analysis. ChIP-seq peaks of dome-primed shield active regions visualized in IGV. Example of two Dome primed Shield active regions having peaks in dome H3K4me3, shield H3K4me3, and in shield

H3K27ac, but not in dome H3K27ac. No specific enrichment with respect to H3K27me3 mark is also observed c) The annotation of the 1113 peak regions obtained from ChIP seq analysis d) Annotation of dome-primed shield active regions. e) Heatmap representing the 1113 regions which were primed in dome with H3K4me3 and getting activated at shield stage with the addition of H3K27ac. The enrichment of H3K4me3 at dome stage is shown in the left panel. The enrichment of H3K4me3 at shield stage, H3K27ac at dome and H3K27ac at shield stage over the same locations is shown in the subsequent panels. H3K27ac levels are very low over these regions at dome stage and hence the priming pattern can be seen.

We also wanted to check whether it was just the priming behavior or whether there was poising happening in the enriched regions. H3K27me3 peak regions of dome and shield regions were visualized in igv and checked whether the dome primed shield active regions were having any poised behavior. No significant enrichment of H3K27me3 was found in those regions, as shown in Figure 6(c). Thus the poised behavior was not present.

2. Gene Ontology analysis of primed regions

Promoters are usually present upstream of a gene. Hence the nearest neighboring method was used to find the genes using the HOMER annotation tool. Though many of the genes obtained were not well studied/related to cell fate specification, there were some well-studied genes as well in the list, which includes gsc, fgf4, notch3, foxc1a, foxg1a, wnt2bb, sirt3, lhx5, pcdh18b, bmp4, pbx2. The genes list includes those shown to play a role in cell fate specification and in important signaling mechanisms and is summarized in Table 4.

Gene	Functions associated
foxc1a	cardiac structure formation, tract development, and eye formation
notch3	cell differentiation during development.
fgf4	cell growth, morphogenesis, tissue repair, tumor growth, and invasion
lhx5	development of the forebrain, neuronal differentiation, and migration.

gsc	formation and patterning of embryo
bmp4	Hematopoietic differentiation, neurogenesis, vascular development
wnt2bb	liver development and swim bladder development
pcdh18b	Motor axon growth and arborization.
foxg1a	neural cell/progenitor proliferation and brain development
sirt3	neuroprotection, cardiovascular diseases, cancer and aging
pbx2	segmentation of hindbrain and pharyngeal arches.

Table 4: Significant genes found near the dome primed shield active regions and their functions.

Gene ontology analysis was carried out with the gene list and the pathways enriched are shown in Figure 7. Geneontology.org was also used for gene ontology analysis and the analysis enriched for pathways for pattern specification, brain, eye, and tube development. Mesodermal cell fate specification is also one of the pathways enriched in the analysis implying the role of primed promoter regions in mesodermal fate specification genes. Precursors of different mesodermal cell types like notochord, prechordal plates, trunk somites, heart, blood, and pronephros are arranged in the dorso ventral axis in the shield stage. The various signaling pathways and the phenotypes associated with the perturbation of many genes associated with mesodermal fate specification have been well documented in zebrafish. Hence it would be interesting to look into the effect of the role of priming in the expression of these genes. Embryonic eye morphogenesis is another enriched pathway from the analysis. The optic primordia develop only by 12 hpf in the zebrafish. Thus, the early expression of embryonic eye morphogenetic and developmental genes could have some role in the fate specification of unspecified cells to that of optic precursors. Similarly, the hindbrain and other organ developmental genes getting activated at this earlier stage could be acting as the earliest fate-specified precursors for the respective organ development. As a whole, the genes involved in the pathways for differentiation and development of various organs can be observed from the data.

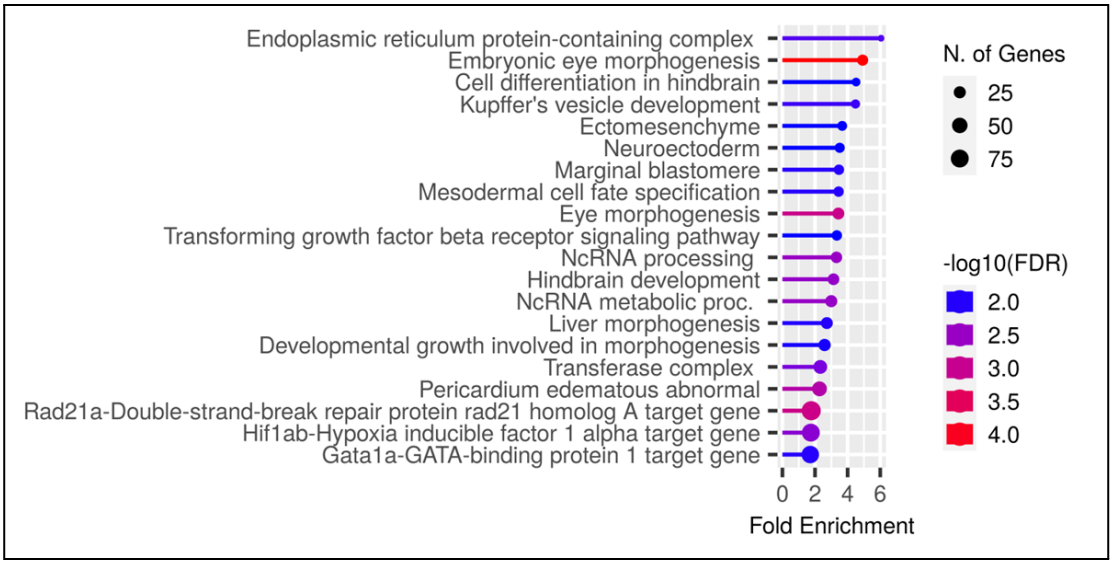


Figure 7: **GO analysis of dome primed shield active genes.** Significant pathways enriched by the genes found in dome primed shield active regions were found using gene ontology analysis. Enriched pathways include Mesodermal fate specification and multiple organ development pathways.

3. Correlation with stage-wise expression data of the genes

Transcriptome data of the genes present nearby the dome primed shield active peak regions were examined. What we expected was a baseline/null expression in the dome stage and an increase in expression from the shield stage when the activation mark is added. But most of the genes didn't have this pattern. This could be because histone marks need not be completely correlated with the gene expression all the time. There were some genes that exhibited this pattern. These genes include fgf4, notch3, foxc1a, lhx5, pcdh18b, bmp4, and pbx2. Their temporal expression data is shown in Figure8. Organogenesis needs further time from the shield stage. Thus the significance of the expression of these organ-specific genes at an earlier time point needs to be further investigated.

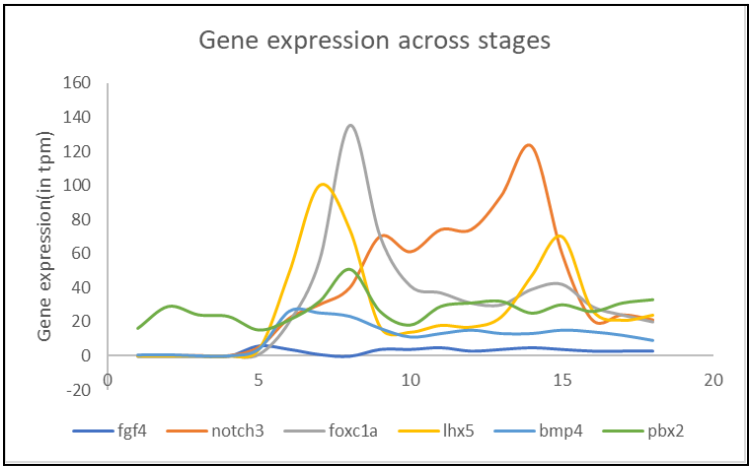


Figure 8: **RNA-seq data of significant enriched genes in the analysis across developmental time periods.** Expression data of dome-primed shield active genes expressing baseline expression in the dome and showing higher activation in the shield.

4. Motif analysis of dome primed shield active peak regions

a)

Rank	Motif	Name	P-value	log P-pvalue
1		MYB119(MYB)/colamp-MYB119-DAP-Seq(GSE60143)/Homer	1e-46	-1.069e+02
2		MYB118(MYB)/colamp-MYB118-DAP-Seq(GSE60143)/Homer	1e-41	-9.647e+01
3		Ronin(THAP)/ES-Thap11-ChIP-Seq(GSE51522)/Homer	1e-34	-7.984e+01
4		bZIP50(bZIP)/colamp-bZIP50-DAP-Seq(GSE60143)/Homer	1e-34	-7.968e+01
5		GFY-Staf(?,Zf)/Promoter/Homer	1e-32	-7.429e+01
6		HY5(bZIP)/colamp-HY5-DAP-Seq(GSE60143)/Homer	1e-32	-7.429e+01
7		ELF1(ETS)/Jurkat-ELF1-ChIP-Seq(SRA014231)/Homer	1e-31	-7.197e+01
8		AT2G15740(C2H2)/col-AT2G15740-DAP-Seq(GSE60143)/Homer	1e-30	-7.101e+01
9		CAMTA1(CAMTA)/col-CAMTA1-DAP-Seq(GSE60143)/Homer	1e-30	-6.995e+01

b)

Rank	Motif	Name	P-value	log P-pvalue
1		MYB119(MYB)/colamp-MYB119-DAP-Seq(GSE60143)/Homer	1e-42	-9.718e+01
2		MYB118(MYB)/colamp-MYB118-DAP-Seq(GSE60143)/Homer	1e-40	-9.364e+01
3		bZIP50(bZIP)/colamp-bZIP50-DAP-Seq(GSE60143)/Homer	1e-35	-8.148e+01
4		Ronin(THAP)/ES-Thap11-ChIP-Seq(GSE51522)/Homer	1e-35	-8.083e+01
5		ELF1(ETS)/Jurkat-ELF1-ChIP-Seq(SRA014231)/Homer	1e-33	-7.739e+01
6		GFY-Staf(?,Zf)/Promoter/Homer	1e-32	-7.441e+01
7		AT2G15740(C2H2)/col-AT2G15740-DAP-Seq(GSE60143)/Homer	1e-31	-7.145e+01
8		HY5(bZIP)/colamp-HY5-DAP-Seq(GSE60143)/Homer	1e-30	-6.964e+01
9		AT3G10030(Trihelix)/colamp-AT3G10030-DAP-Seq(GSE60143)/Homer	1e-29	-6.856e+01
10		ATAF1(NAC)/col-ATAF1-DAP-Seq(GSE60143)/Homer	1e-29	-6.835e+01
11		MYNN(Zf)/HEK293-MYNN.eGFP-ChIP-Seq(Encode)/Homer	1e-29	-6.757e+01
12		CAMTA1(CAMTA)/col-CAMTA1-DAP-Seq(GSE60143)/Homer	1e-28	-6.494e+01

Figure 9: Motif analysis of dome primed shield active regions. a) Significant motifs found using HOMER motif analysis on dome primed shield active regions includes, MYB, Ronin, GFY, etc. b) Motifs enriched when 1000 bp upstream and downstream sequences were taken for motif analysis gave the same top enriched motifs of MYB and Ronin.

Motif analysis was carried out using the HOMER find motifs tool, and the results are shown in Figure 9. MYB119, MYB118, Ronin, bZIP50, GFY, ELF1, CAMTA1, and NPAF1 are the transcription factors whose binding sites were found to be significantly enriched in the dome-primed shield active peak regions. In the motif analysis results, we observe that 11 of the Fox family proteins are enriched, implying a role for them. Amongst these, FoxA2(1e-7), FoxA3(1e-4), and FoxH1(1e-6) have their RNA seq expression data higher in shield compared to the dome. Previous studies have reported priming activity related to the Fox family of transcription factors. FoxA2 has been previously shown to have a role in enhancer priming during pancreatic differentiation(Lee et al., 2019). It helps the genomic regions to acquire proper

chromatin accessibility during pancreatic differentiation. Similarly, some of the transcription factors in the ETS family of TFs (Gabpa(1e-10, Etv4(1e-20), Elf1(1e-31), Elf3(1e-3), E2F4(1e-3)) are also enriched in the motif analysis and are shown to have the expected expression pattern. Sox2(1e-3), Sox3(1e-5), Gata3(1e-8), and NfyA(1e-10) are also enriched factors with expected expression patterns.

Motif analysis was done with the 108 promoters which we got in the analysis of dome primed and active in shield peaks, and the results are shown in Figure 10. The important motifs found in their regions were found out, and their significance of enrichment is also shown in the figures. Elk4, NRF, Elk1, Ronin, Fli1, GFY, ETV4, ATAF1, and MYB118 are the most enriched motif regions in the promoter peak regions. We see that the significance of the motif has significantly gone down compared to the previous analysis. This is because of the low number of sequences that we are using for the motif analysis. To check other transcription factor binding regions around the peak regions, 1000 base pairs upstream and downstream of these promoter peak regions were also retrieved, and their motif analysis was done. Its result is also shown below, with NRF, Elk4, Elk1, Ronin, Fli1, and GFY still being the top enriched motif regions. All of these genes encode proteins that are involved in the regulation of gene expression and play important roles in various cellular processes, such as cell growth, differentiation, apoptosis, and response to extracellular signals.

a)

Rank	Motif	Name	P-value	log P-pvalue
1		Elk4(ETS)/Hela-Elk4-ChIP-Seq(GSE31477)/Homer	1e-11	-2.584e+01
2		NRF(NRF)/Promoter/Homer	1e-11	-2.560e+01
3		Elk1(ETS)/Hela-Elk1-ChIP-Seq(GSE31477)/Homer	1e-11	-2.553e+01
4		NRF1(NRF)/MCF7-NRF1-ChIP-Seq(Unpublished)/Homer	1e-10	-2.435e+01
5		Ronin(THAP)/ES-Thap11-ChIP-Seq(GSE51522)/Homer	1e-8	-1.852e+01
6		Fli1(ETS)/CD8-FLI-ChIP-Seq(GSE20898)/Homer	1e-7	-1.835e+01
7		GFY-Staf(?,Zf)/Promoter/Homer	1e-7	-1.720e+01
8		ETV4(ETS)/HepG2-ETV4-ChIP-Seq(ENCODE)/Homer	1e-7	-1.637e+01
9		ATAF1(NAC)/col-ATAF1-DAP-Seq(GSE60143)/Homer	1e-6	-1.548e+01
10		AT2G15740(C2H2)/col-AT2G15740-DAP-Seq(GSE60143)/Homer	1e-5	-1.318e+01
11		SeqBias: CG bias	1e-5	-1.304e+01

b)

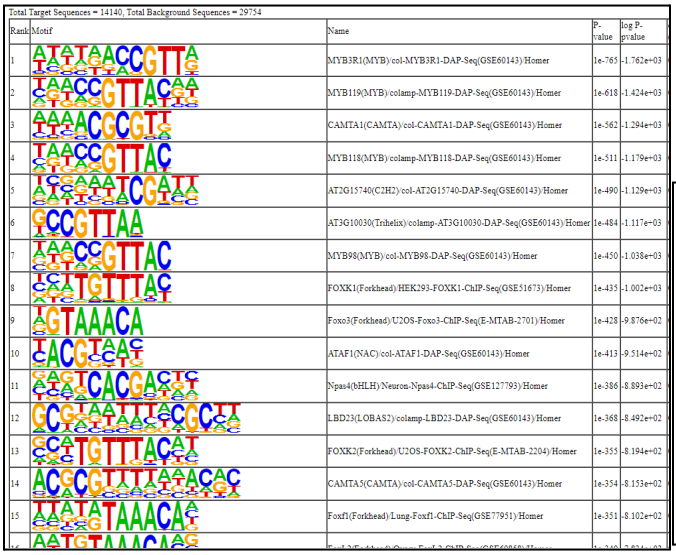
Rank	Motif	Name	P-value	log P-pvalue
1		NRF(NRF)/Promoter/Homer	1e-11	-2.602e+01
2		Elk4(ETS)/Hela-Elk4-ChIP-Seq(GSE31477)/Homer	1e-11	-2.553e+01
3		NRF1(NRF)/MCF7-NRF1-ChIP-Seq(Unpublished)/Homer	1e-11	-2.533e+01
4		Elk1(ETS)/Hela-Elk1-ChIP-Seq(GSE31477)/Homer	1e-10	-2.471e+01
5		Ronin(THAP)/ES-Thap11-ChIP-Seq(GSE51522)/Homer	1e-7	-1.804e+01
6		Fli1(ETS)/CD8-FLI-ChIP-Seq(GSE20898)/Homer	1e-7	-1.760e+01
7		GFY-Staf(?,Zf)/Promoter/Homer	1e-7	-1.651e+01
8		ETV4(ETS)/HepG2-ETV4-ChIP-Seq(ENCODE)/Homer	1e-6	-1.530e+01
9		ATAF1(NAC)/col-ATAF1-DAP-Seq(GSE60143)/Homer	1e-6	-1.481e+01
10		SeqBias: CG bias	1e-5	-1.346e+01
11		AT2G15740(C2H2)/col-AT2G15740-DAP-Seq(GSE60143)/Homer	1e-5	-1.291e+01

Figure 10: **Motif analysis of dome primed shield active promoter regions.** a) Motif analysis was carried out with the 108 promoter peak regions obtained from the dome-primed shield active regions. b) 1000 base pairs upstream and downstream of the peak regions were collected, and the motif analysis result is shown in the second figure.

5. Motif analysis and gene ontology analysis of active regions in shield

Since the primed regions identified from the analysis was significantly less, we tried to identify all the important motifs and related gene ontology functions in the shield stage. The important motifs found in the active regions of the shield development stage were found by overlapping H3K4me3 and H3K27ac regions. Motif analysis was done with the peaks, and important motifs with a high significance of enrichment were found in them and as shown in Figure 11. The genes nearby them were annotated using the nearest neighboring method, and those genes were used to do the gene ontology analysis.

a)



b)

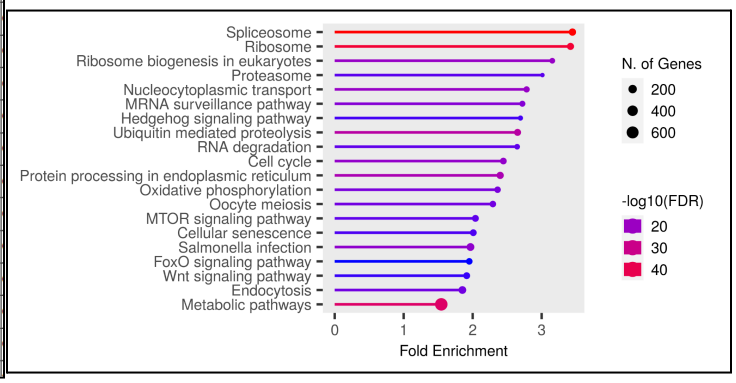


Figure 11: **Analysis of active regions in the shield stage.** a) Motif regions found from active regions in shields. b) Gene ontology analysis results of genes next to these active regions.

The p values for enrichment for motifs in active peak regions in the shield stage are highly significant. The enriched motifs continue to be the same as those of primed regions. MYB119, MYB118, CAMTA1, myb3r1, FoxK1, FoxO3, and FoxK2 are some of the top enriched motifs. Signaling mechanisms have been shown to control cell fate and embryonic patterning(Perrimon et al., 2012). Genes for such signaling pathways like

wnt, hedgehog, and FoxO can be seen enriched in the gene ontology analysis, as shown in Figure 11. The Wnt signaling pathway plays an important role in the specification of mesodermal fate (Hikasa & Sokol, 2013). Previous research utilizing methods such as in situ hybridization have confirmed the presence of these signaling pathways during the shield stage, thus the GO terms and thereby the enriched genes are in aligned with their temporal expression profile in zebrafish.

6. Selection of genes for validation and checking expression

Among the 108 regions found to be dome-primed and active in the shield, genes that were related to cell fate specification and having an expression pattern relatable to priming were investigated.

Promoter primed genes showing the desired expression		**Promoter primed genes related to cell fate specification**		
➢ Lamb1a	➢ Sco2	➢ Casp9	➢ Pdk4	➢ Ccne2
➢ Ccne2	➢ Zdhhc16a	➢ Spcs2	➢ Dcaf6	➢ Ino80
➢ Mnd1	➢ Chchd2	➢ Pex19	➢ Sall1a	➢ Chchd2
➢ Ino80b	➢ Psmg2	➢ Crlf3	➢ Lamb1a	
➢ Pfas	➢ Hip1	➢ Sco2	➢ Chid1	
➢ Sall1a	➢ Taf8	➢ Sirt3	➢ Scamp3	

Table 5: Primed genes showing cell fate specification related function, Primed genes showing priming related expression.

By intersecting both these tables, the following genes were taken for further studies.
Lamb1a, Ccne2, Ino80b, ChChd2 and Sall1a.

Lamb1a is a gene that encodes for the Lamb1a (Laminin subunit alpha-1) protein which is a component of the extracellular matrix of many tissues (Sun et al., 2008). This includes muscle tissues, the nervous system, and skin. ECM plays an essential role in cell fate specification. Lamb1a expression is required for the formation and maintenance of neural stem cell niches. Defects in neural development have been reported in the absence of lamb1a expression. Lamb1a has been shown to be involved in the migration of neural crest cells in zebrafish (Henderson & Copp, 1997). Similarly, lamb1a has been

shown to play an important role in muscle development(Thorsteinsdóttir et al., 2011). Myotubes, which are precursors of muscle fibers, require its expression. In zebrafish, its expression has been shown in somites, which are precursors for muscle cells. It has been shown in mice that Lamb1a deficiency leads to defects in muscle development with reduced muscle mass. Its expression is regulated by various signaling pathways, including the notch and wnt pathways, which regulate many aspects of cell fate specification

Ccne2 gene encodes for the Cyclin E2 protein, which plays a critical role in cell cycle regulation. The Cyclin E2 protein binds to and activates the cyclin-dependant kinase 2 (CDK-2), leading to the initiation of DNA synthesis and cell division (Fagundes & Teixeira, 2021). Ccne2 gene expression is upregulated around the mid-blastula stage in zebrafish embryos, where major changes in gene expression and cellular organization happen. Its deficiency at this time period has been shown to affect the proper timing of the cell cycle during this transition, and its deficiency has defects in the early development of embryos. It has also been reported that Ccne2 regulates cell cycle speed which in turn influences their differentiation potential. Mesodermal progenitors (MEP) with higher levels of Ccne2 expression have shorter cell cycle and are biased toward muscle cell formation, while the one with lower levels and longer cell cycle is biased toward blood vessel formation (Lu et al., 2018). Ccne2 has also been shown to be required for hematopoietic stem cell fate specification (Caldon & Musgrove, 2010; Umemoto et al., 2018). Its deficiency in mice has been shown to have defects in HSC specification and function. Overall, these studies suggest that the Ccne2 gene plays a crucial role in the specification of MEPs, including HSCs, by regulating the cell cycle speed and thereby influencing fate decisions

Ino80b (Ino80 complex subunit B) encodes for a protein of the Ino80 chromatin-remodeling complex, which plays a critical role in regulating the structure and accessibility of chromatin, thereby regulating gene expression (Wang et al., 2014). Han et al. has earlier reported that Ino80b is required for the specification of cardiac progenitor cells (CPCs) during zebrafish development (P. Han et al., 2011). They showed that Ino80b, whose expression is regulated by the notch signaling pathway, modulates the chromatin accessibility at key cardiac transcription factors such as tbx5 and nkx2.5, thereby regulating their expression. Ino80b deficiency led to defects in CPC

specification and showed cardiac hypoplasia. It has also been shown to play a role in the differentiation of oligodendrocytes. It modulates accessibility at oligodendrocyte-specific genes such as myelin basic protein (Mbp) and thereby regulates myelin formation. Also, H2A.Z are deposited at promoter regions by the Ino80 complex (Papamichos-Chronakis et al., 2011). H2A.Z deposition at promoter regions has been shown to be a form of priming for activating genes. Thus the role of INO80 in the context of the activation of lineage-specific genes by this method needs to be investigated.

Sall1a (Sal-like protein 1a) gene is expressed and has been shown to play a role in the specification of a variety of tissues during zebrafish embryo development, including the heart, eyes, brain, and kidneys (Diotel et al., 2015). Sall1a regulates the expression of pax2a and lim1 transcription factors, which helps in the specification and development of pronephros, the embryonic kidney in zebrafish (Basta et al., 2017; Schebesta et al., 2006). It also regulates the expression of pax6a and six3b genes which plays an important role in eye development. The key transcription factors for CPC specification, tbx5 and nkx2.5, as mentioned above, are also regulated by Sall1a (Harvey & Logan, 2006). The deficiency of all these genes mentioned above results in developmental abnormalities in zebrafish.

Chchd2 (Coiled-coil-helix-coiled-coil-helix domain containing 2) encodes for a mitochondrial targeting protein shown to play a role in brain, eye, heart, and somite development (Petel Légaré et al., 2023). It regulates the expression of key neural transcription factors such sa neuorog1 and ascl1a, which regulates neural differentiation (Zhu et al., 2016). Chchd2 has been reported to play an important role in the specification of the anterior lateral plate mesoderm (ALPM) in zebrafish. Chchd2 deficient embryos fail to form functional hearts, as ALPM specification is affected. It is also reported to be expressed in presomitic mesoderm and thereby regulates genes involved in somitogenesis.

The ChIP seq data visualized in IGV around these regions are shown in Figure 12(b). The RNA-seq expression of these genes across developmental stages of zebrafish is shown in Figure 12(a). These genes show a suddenly increased expression around the shield stage.

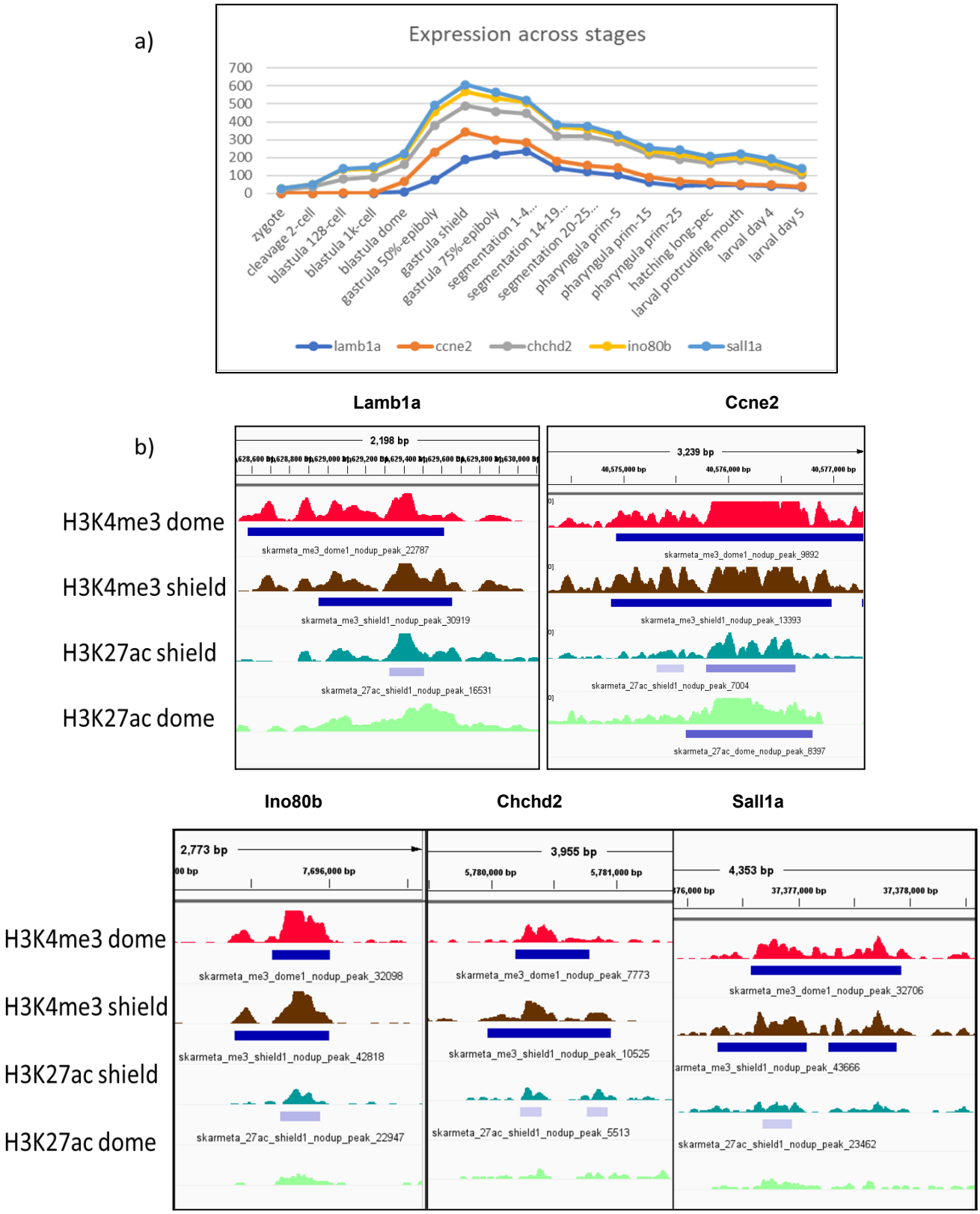


Figure 12: **Selection of genes based on the expression profiles and ChIP enrichment.** a) Gene expression profile of selected genes at early developmental time

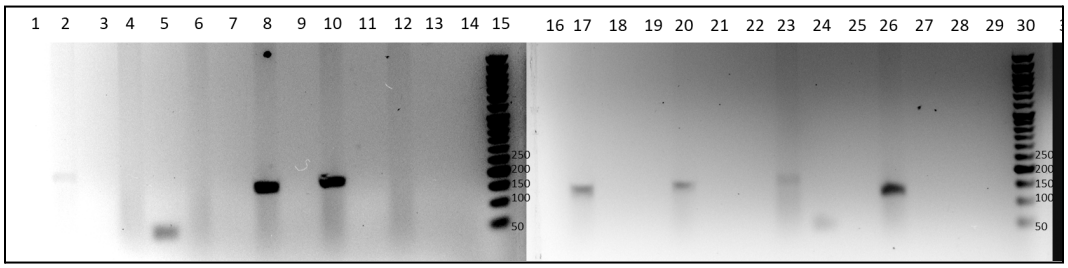
period. These genes exhibit an increased expression from the dome stage until the shield stage. b) IGV tracks of selected genes showing priming activity. Peak calling happened at H3K4me3 at dome and shield stages, while peak calling for H3K27ac happened only in the shield stage for these regions.

7. Validation of ChIP data

The presence of H3K4me3 and H3K27ac marks over these regions was analyzed using ChIP-qPCR. Specific primers were designed to span the peak regions so as to pull down efficiently. The specificity and efficiency of primers were tested using a q-PCR with zebrafish genomic DNA. Specific regions were amplified using the primers, and samples were run on a 2 % gel after the qPCR, as shown in Figure13(a).

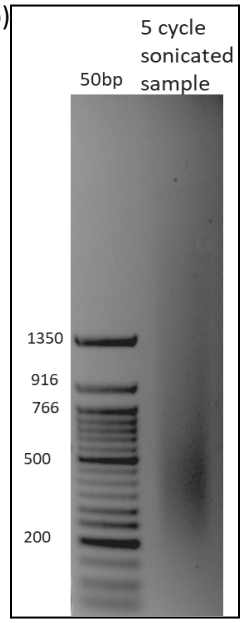
q-PCR was performed using ChIP lysates for H3K4me3 and H3K27ac at the dome and shield stage to validate the priming marks. The efficiency of chromatin shearing was checked after sonication by reverse cross-linking and purifying a small amount of the sonicated sample. This was run on a 2% gel and is shown in Figure13(b). IP was carried out with this sonicated sample using H3K4me3, H3K27ac, IgG and Input at dome and shield stages. qPCR analysis was carried out, and the samples were run on a 2% gel, as shown in Figure 13(c,d). Input shield samples didn't show amplification because of some experimental error. Hence it needs to be repeated. The results of ChIP qPCR analysis of H3K4me3 and H3K27ac pull-down at the dome stage are shown in Figure 13(e). Enrichment of H3K4me3 at dome and shield is clear from the gel images. Enrichment of H3K27ac needs to be further compared with shield stage and validated.

a)

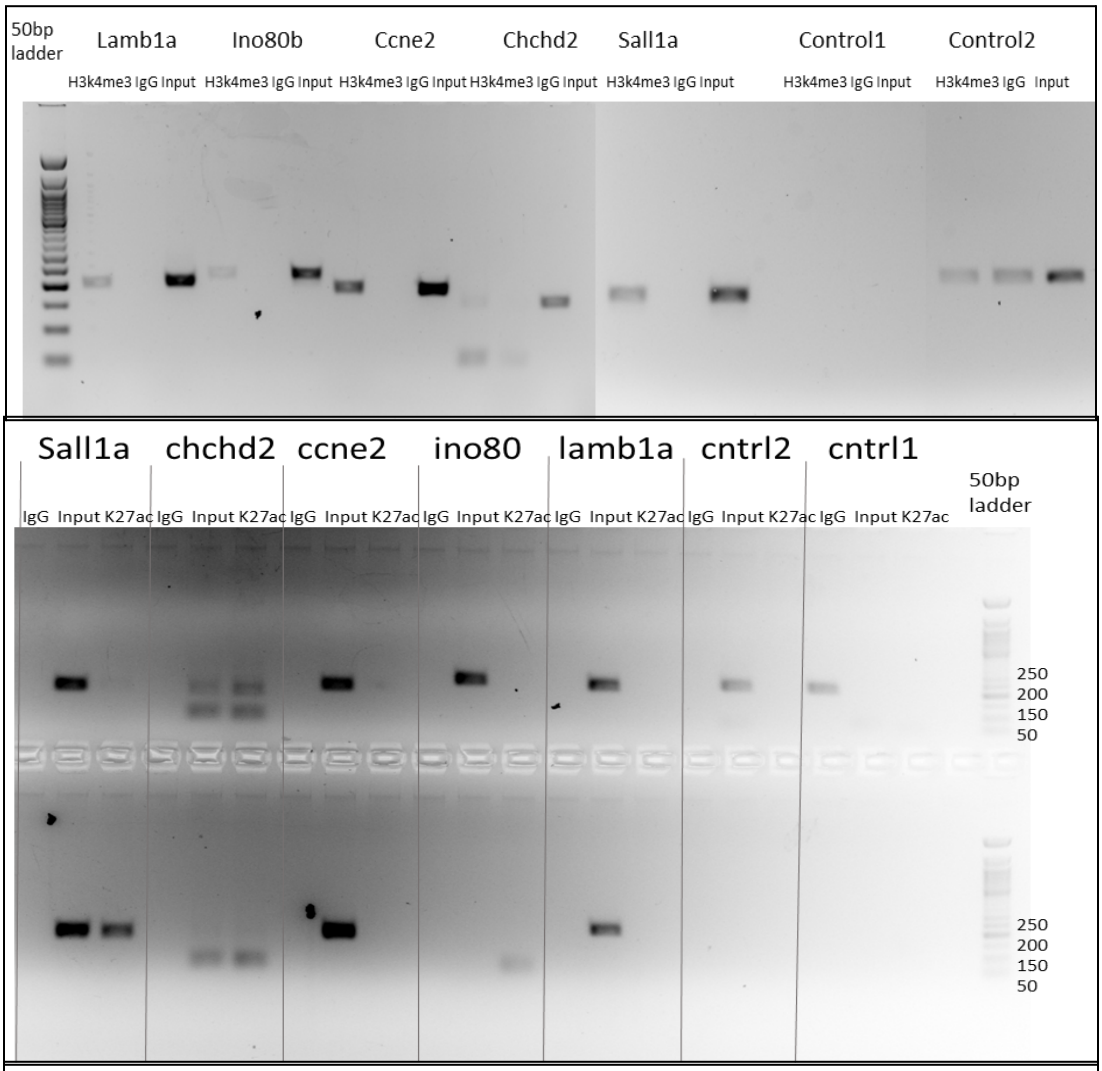


1-Empty
2- Ino80b
3-Ino80b NTC
4-Ino80b
5-Ino80b NTC
6-CCNE2
7-CCNE2_NTC
8-CCNE2
9-CCNE2_NTC
10-Lamb1a
11-Lamb1a_NTC
12-Lamb1a
13-Lamb1a_NTC
14-Empty
15-50bp Ladder
16-Empty
17-Sall1a
18-Sall1a_NTC
19-Empty
20-Sall1a
21-Sall1a_NTC
22-Empty
23-ChChd2
24-ChChd2_NTC
25Empty
26-ChChd2
27-ChChd2_NTC
28-Empty
29-Empty
30-50bp ladder

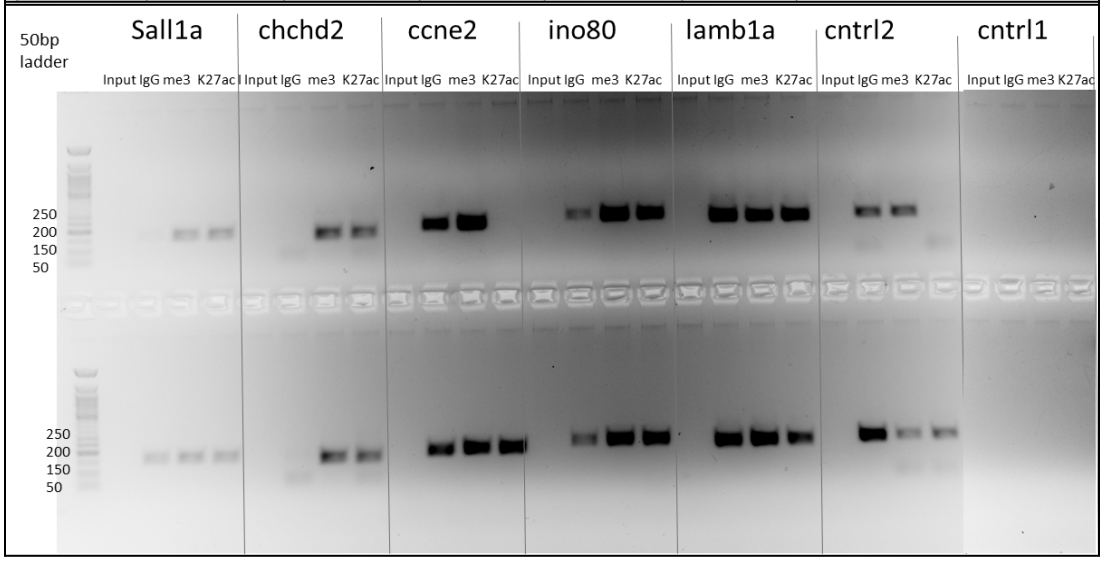
b)



c)


50bp ladder
Lamb1a
Ino80b
Ccne2
Chchd2
Sall1a
Control1
Control2
H3k4me3 IgG Input
Sall1a
chchd2
ccne2
ino80
lamb1a
cntrl2
cntrl1
IgG Input K27ac
50bp ladder
250
200
150
50

d)


50bp ladder
Sall1a
chchd2
ccne2
ino80
lamb1a
cntrl2
cntrl1
Input IgG me3 K27ac
250
200
150
50

e)

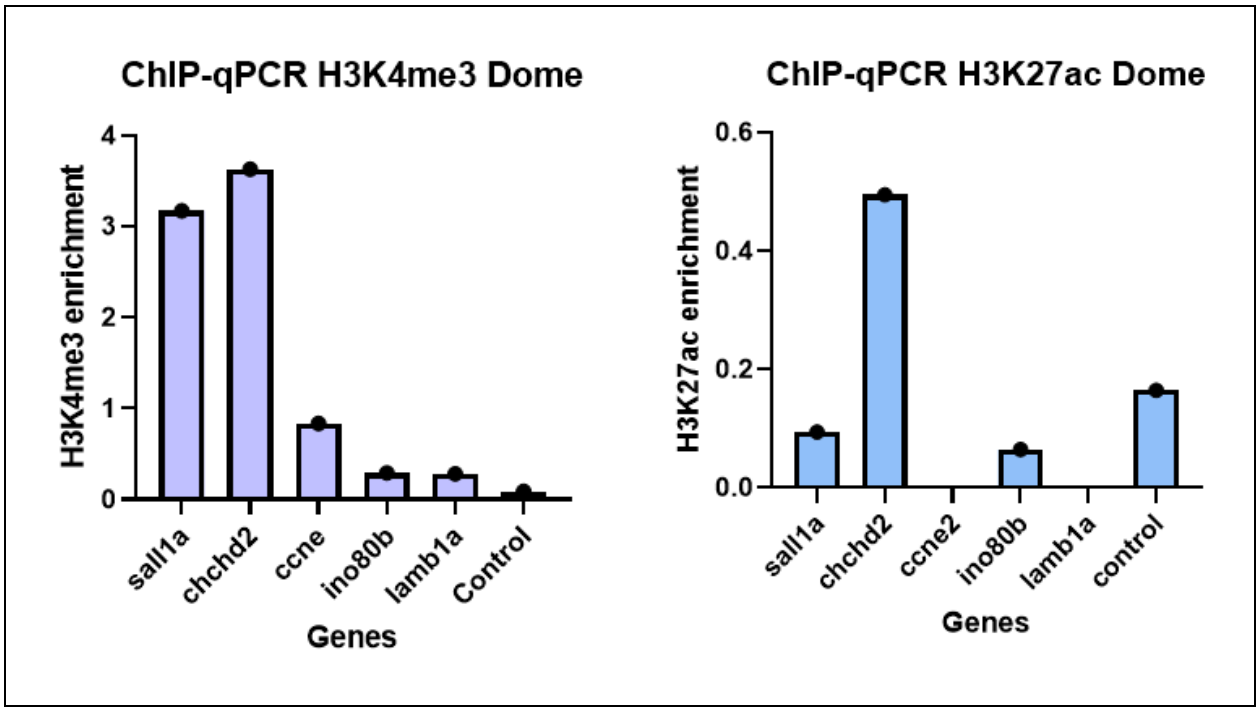


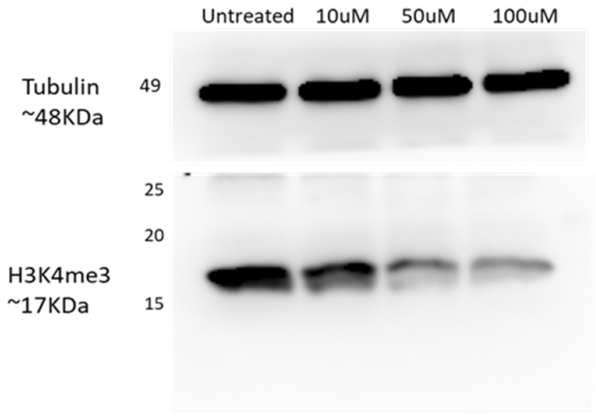
Figure 13: **Validation of selected genes using ChIP-qPCR.** a)ChIP qPCR primers were validated by setting up an RT- qPCR reaction. They were run on a 2% gel to validate that the amplification of the desired size was happening. b) Gel electrophoresis image of sonicated chromatin run on a 2% gel c) Gel electrophoresis image of ChIP qPCR amplified sample, bands can be seen in the H3K4me3 pull down for the selected gene regions, while IgG pull down doesn't have these gene regions enriched in them. d) Gel electrophoresis image of ChIP qPCR amplified samples at shield stage, bands can be seen in the H3K4me3, IgG pull down. But input didn't give bands and amplification in the RT-qPCR, hence the qPCR needs to be repeated. e) Fold enrichment of H3K4me3 and H3K27ac (n=1) in the selected gene regions at dome stage. The enrichment of these modification marks at this specific sequence location can be seen. H3K4me3 mark is enriched at the selected gene regions in the dome stage. H3K27ac is also enriched, but in a lower level at the dome stage in these genomics regions. The control region, as expected, doesn't have any enrichment for both the marks.

8. Drug treatment: MM-102 Dosage study

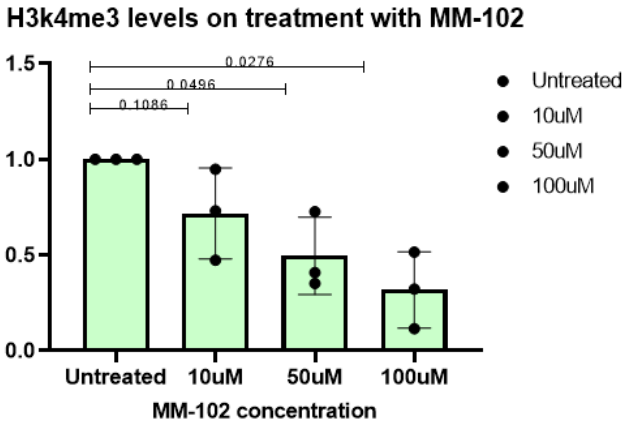
Although MM-102 has been used earlier in zebrafish, there is currently no available data on the quantification of the H3K4me3 mark upon its treatment. Previous studies have reported varying effects of the drug on zebrafish embryos at different concentrations. Chan et al. found that MM-102 treatment with concentrations of 1500µm 750µm and 300µM led to the death of all embryos, while 150µM resulted in some surviving embryos. Meanwhile, Westphal et al. utilized MM-102 concentrations of 30µm 10µm 3µm and 1µM to evaluate the expression of neuronal markers in zebrafish, and no mortality was reported in their study. To investigate the effects of the drug without causing mortality, we selected concentrations of 10µm 50µm and 100uM.

The effect of MM-102 drug on H3K4me3 levels at the dome stage was investigated by treating with different doses of the drug and analyzing its impact. Western blot analysis revealed a gradient of H3K4me3 levels inversely proportional to the drug concentration, confirming the drug's ability to successfully inhibit kmt2a expression and reduce H3K4me3 levels in a dosage-dependent manner. However, no significant deformity and mortality rate were obtained upon drug treatment at this concentration, as shown in Figure 14(c,d). Images of the embryos upon MM-102 treatment and DMSO treatment are shown in Figure 14(e).

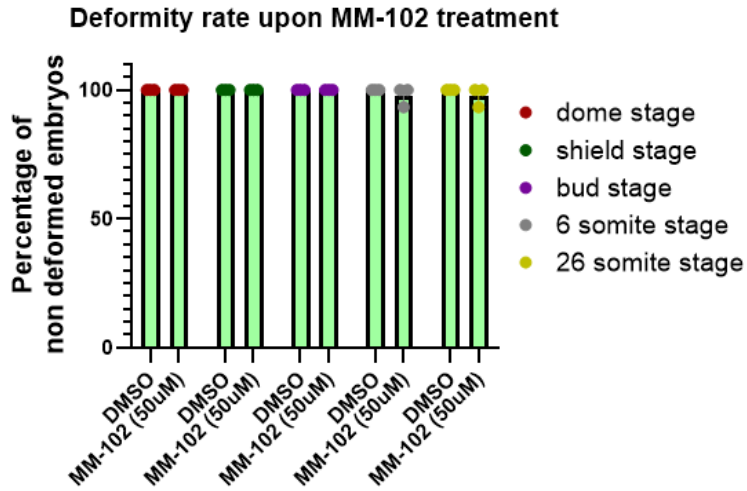
a)


Untreated
10uM
50uM
100uM
Tubulin
~48KDa
49
25
20
H3K4me3
~17KDa
15

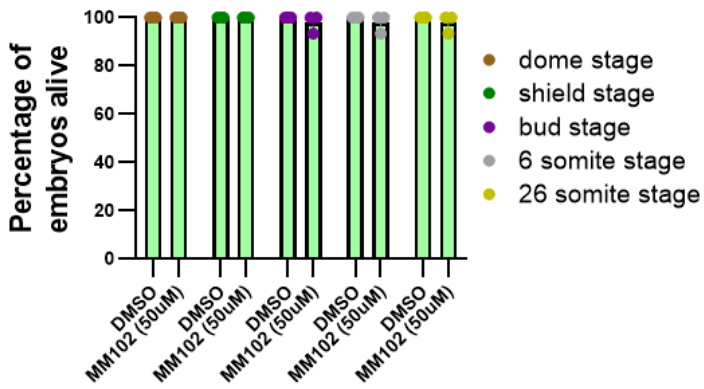
b)


H3k4me3 levels on treatment with MM-102
0.0276
0.0496
0.1086
1.5
1.0
0.5
0.0
Untreated
10uM
50uM
100uM
MM-102 concentration

c)


Deformity rate upon MM-102 treatment
Percentage of non deformed embryos
100
50
0
DMSO
MM-102 (50uM)
dome stage
shield stage
bud stage
6 somite stage
26 somite stage

d)


Mortality rate upon MM-102 treatment
Percentage of embryos alive
100
80
60
40
20
0
DMSO
MM102 (50uM)
dome stage
shield stage
bud stage
6 somite stage
26 somite stage

e)

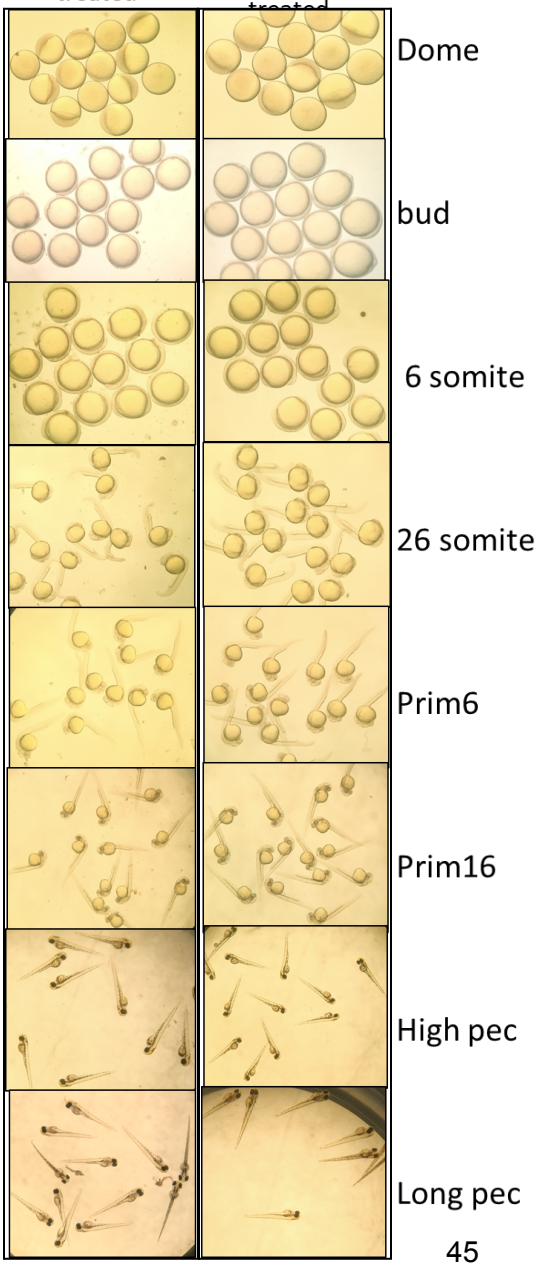

DMSO treated
50uM MM-102
Dome
bud
6 somite
26 somite
Prim6
Prim16
High pec
Long pec

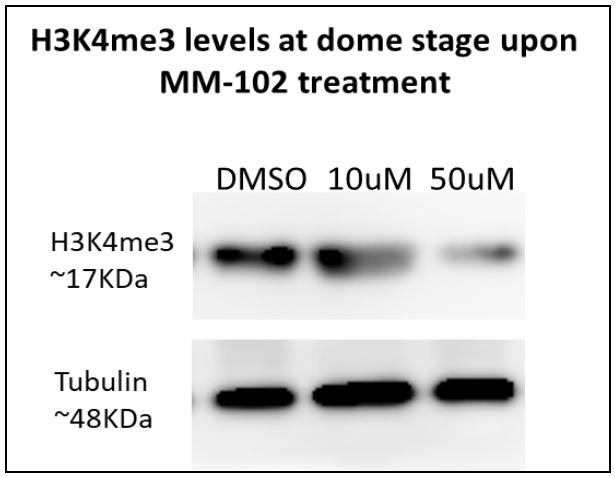
Figure 14: **Effect of MM-102 drug treatment on developing zebrafish embryos.** a) Western blot image of different concentrations of MM-102 drug-treated zebrafish embryos. b) Quantification of the H3K4me3 levels in different samples (n=3) c) Deformity rate of embryos upon MM-102 treatment d) Mortality rate of embryos upon MM-102 treatment. e) Image of embryos at different time points of development upon MM-102 treatment compared to DMSO control.

9. H3K4me3 levels of MM-102 treated samples at dome and shield stages

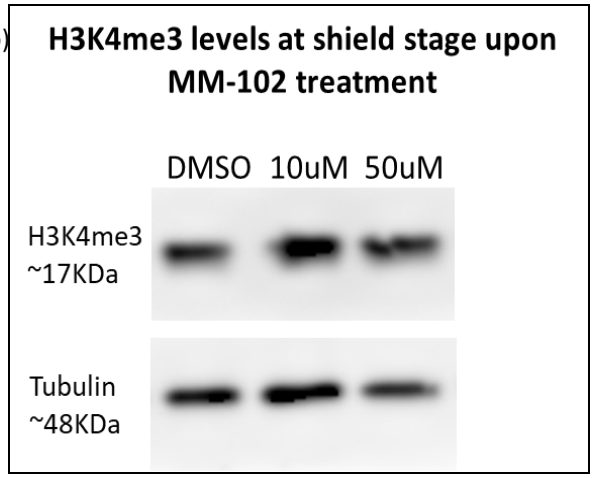
Embryos were treated with MM-102 drug till the dome and shield stage to observe the effects of MM-102. Table X summarizes the doses of MM-102 and the developmental stages used in the drug treatment study. As shown in Figure15(a,c), western blot data showed that MM-102 was able to reduce the expression of H3K4me3 levels at the dome stage. But surprisingly, western blot data of MM-102 indicated that it didn't have any impact at the shield stage. MM-102 was not able to reduce the expression of H3K4me3 levels at the shield stage, as shown in Figure15(b,d). These results indicated that MM-102 was able to successfully remove only the H3K4me3 levels till the dome stage, which is ideally required for studying the priming behavior.

As reported earlier, no phenotypic changes were also associated with MM-102 treatment. This could be because the H3K4me3 protein expression level is catching up by the time it reaches the shield stage. This could be either because MM-102 is getting consumed at an early stage or a different writer could be coming into action after the dome stage, thereby adding the H3K4me3 marks. As shown in Figure16(e), the writer Kmt2a, inhibited by MM-102, is mainly acting only till the dome stage (Papatheodorou et al., 2018). Thus the H3K4me3 levels are getting reduced in the dome stage upon drug treatment. But Kmt2a has a lower expression in the stages from dome to shield, thereby MM-102 need not have any role in regulating the expression of H3K4me3 marks. A different writer could be playing a role in adding the marks at this stage. Other H3K4me3 writers expressing at this stage need to be looked into. But for our study, this came to be an advantage as we are able to target a specific writer who mainly acts in the stage where priming marks are added, thereby giving us a temporal advantage in studying the priming effects.

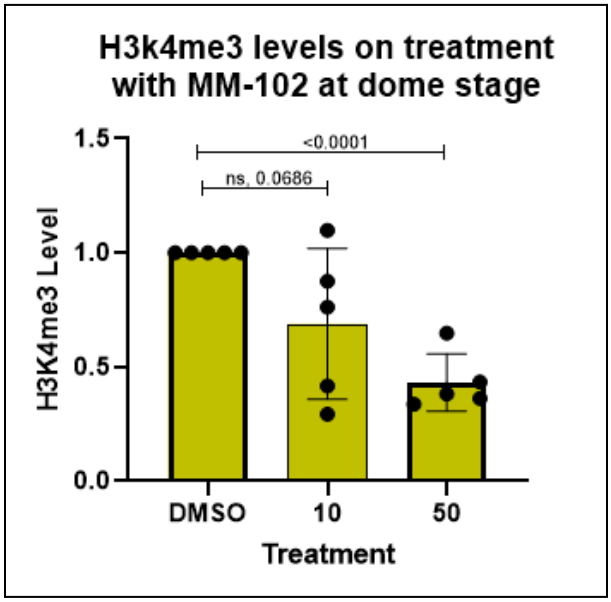
a)


H3K4me3 levels at dome stage upon MM-102 treatment
DMSO 10uM 50uM
H3K4me3 ~17KDa
Tubulin ~48KDa

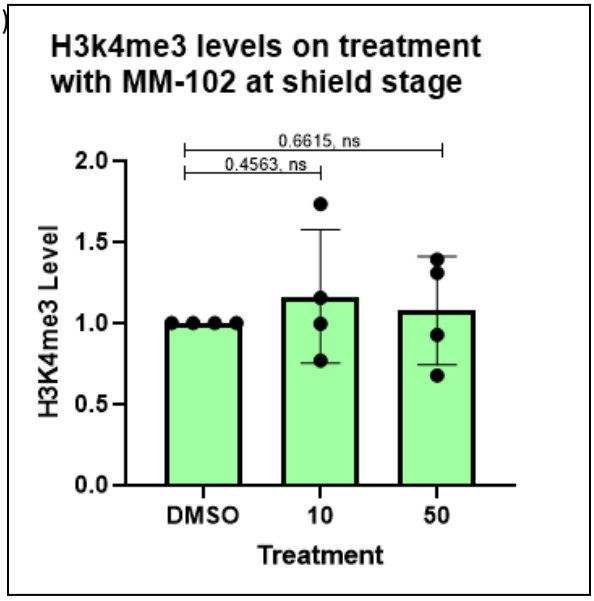
b)


H3K4me3 levels at shield stage upon MM-102 treatment
DMSO 10uM 50uM
H3K4me3 ~17KDa
Tubulin ~48KDa

c)


H3k4me3 levels on treatment with MM-102 at dome stage
<0.0001
ns, 0.0686
H3K4me3 Level
1.5
1.0
0.5
0.0
DMSO
10
50
Treatment

d)


H3k4me3 levels on treatment with MM-102 at shield stage
0.6615, ns
0.4563, ns
H3K4me3 Level
2.0
1.5
1.0
0.5
0.0
DMSO
10
50
Treatment

e)

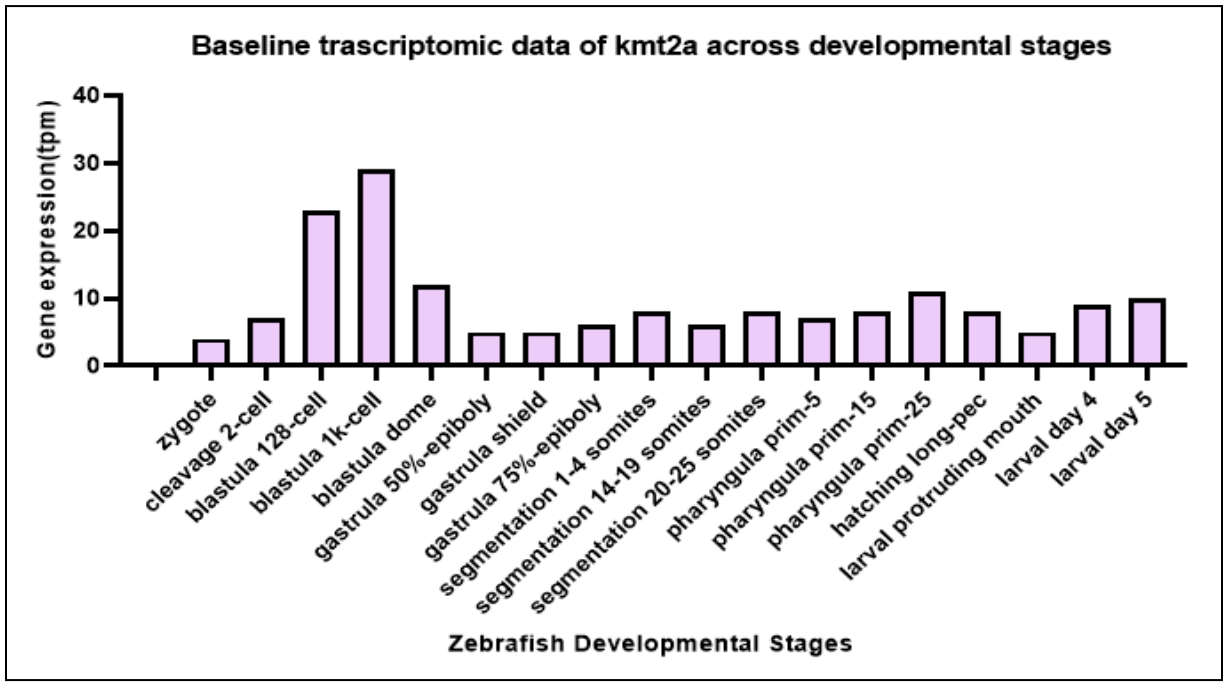

Baseline trascriptomic data of kmt2a across developmental stages
Gene expression(tpm)
40
30
20
10
0
zygote
cleavage 2-cell
blastula 128-cell
blastula 1k-cell
blastula dome
gastrula 50%-epiboly
gastrula shield
gastrula 75%-epiboly
segmentation 1-4 somites
segmentation 14-19 somites
segmentation 20-25 somites
pharyngula prim-5
pharyngula prim-15
pharyngula prim-25
hatching long-pec
larval protruding mouth
larval day 4
larval day 5
Zebrafish Developmental Stages

Figure 15: **Analysis of the effect of MM-102 treatment at dome and shield stage and analyzing the expression of its target KMT2a across developmental timepoints**. a,b) Western blot image data probing H3K4me3 levels upon MM-102 treatment. c,d) Quantification of the western bands(n=4) obtained at different concentrations of MM-102. There is a significant reduction in the H3K4me3 levels upon treatment by a concentration of 50μM MM-102 with a p-value<0.00001 in the dome stage. But the depletion was not that significant with 10μM treatment as the p-value is 0.0686. But at the shield stage, 10μM and 50μM MM-102 treatment don't have any significant impact on the H3K4me3 levels as the p-values are 0.4563 and 0.6615, respectively. e)Expression of MLL1 writer (Kmt2a) across developmental stages

10. RT-qPCR with MM-102 treated embryos

Since the drug treatment is able to successfully remove the H3k4me3 levels at the dome stage, qPCR was carried out to check whether this perturbation has an impact on the expression of the selected genes. qPCR was carried out at the shield stage as this is the stage where the primed genes are expressed actively. As shown in Figure16, qPCR of drug-treated samples at the shield stage showed that there is a reduction in the expression of the genes compared to DMSO-treated samples. This is significant as H3K4me3 levels were not reduced at the shield stage upon MM-102 treatment. A reduction of H3K4me3 levels in the dome stage was able to create an impact on the expression of these genes at the shield stage.

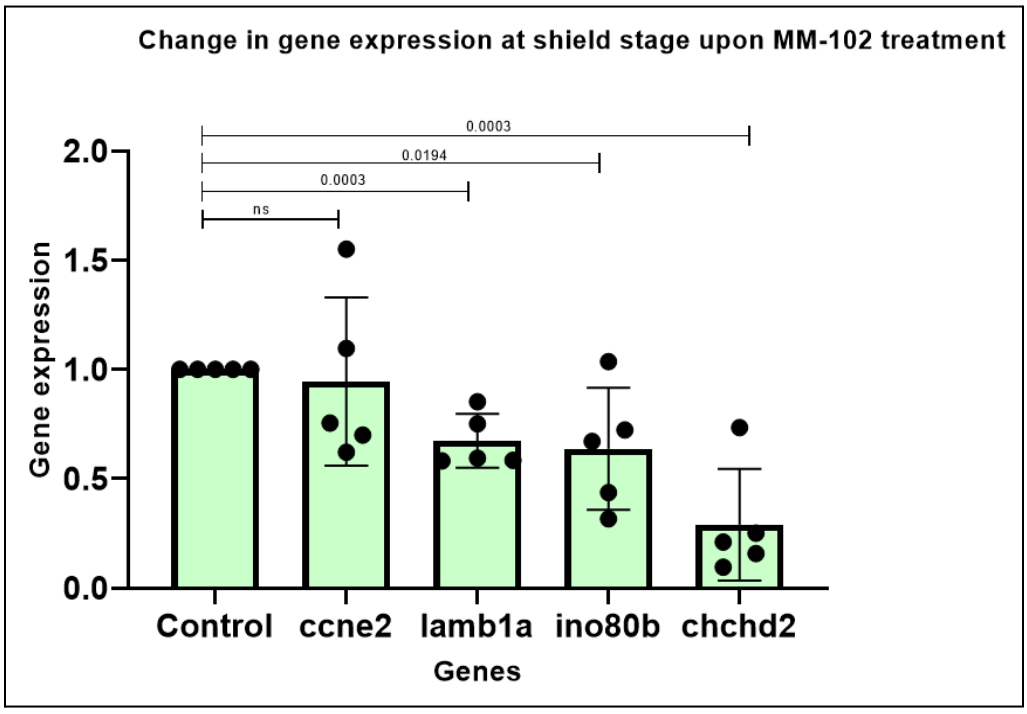


Figure 16: **qPCR analysis of MM-102 treated embryos at shield stage**. Images of RT-qPCR data (n=5) of embryos treated with 50 μM of MM-102 vs. DMSO control at till shield stage. Ccne2 didn't have a significant reduction in its expression, as shown in the Figureure. Lamb1a, Ino80b, and Chchd2 genes had a significant reduction in their expression upon MM-102 treatment with a p-value of 0.0003, 0.,0194, and 0.0003 respectively.

This points toward the activatory role that H3K4me3 could be playing in regulating their expression. It was earlier reported that depletion of H3K4me3 in embryonic stem cells resulted in the downregulation of pluripotency genes and the upregulation of lineage-specific genes, indicating a loss of stemness and a drive toward differentiation (Grandy et al., 2016). Also, studies in the context of cancer have reported that the removal of H3K4me3 reduced cell proliferation and tumor growth as it reduced the transcription of genes critical for tumor cell survival and proliferation. The removal of the H3K4me3 prepatterning mark can make it more difficult for the transcriptional machinery to access the DNA, which in turn can lead to a decrease in the level of transcription of that gene. It could also affect the ability of other marks to be added or removed from the histones in the vicinity of the gene, leading to further changes in gene expression.

The impact of MM-102 treatment on Ccne2 was found to be minimal, despite being classified as a primed gene. Upon revisiting the histone peak data in IGV, it was found that Ccne2 exhibited a dispersed acetylation pattern, with an acetylation peak detected in close proximity to the region that was considered primed. The presence of acetylation mark (H3K27ac) in the earlier stage itself could have helped it in overcoming the gene expression challenge induced by H3K4me3 depletion. This signifies the role of priming in gene expression. The observation that only primed genes were affected by the drug-treated genes signifies the mechanism of priming in regulating gene expression at the right threshold at the right time point. How the presence of H3K4me3 marks in the dome stage facilitates the activation, and increased expression of the primed genes needs to be further explored.

11. Effect of MM-102 treatment on H3K27ac levels

Since H3K27ac acts as an activatory mark for genes, we monitored the levels of H3K27ac upon MM-102 treatment at the shield stage where the expression of genes was found to be reduced. It was observed that the levels of H3K27ac reduces upon MM-102 treatment at shield stage as quantified from western blot of drug treated samples as shown in Figure17. As observed in H3K4me3 levels upon MM-102 treatment at dome stage, there is no significant reduction observed in H3K27ac at shield stage as well at his concentration. H3K4me3 levels are not getting reduced at the shield stage while H3K27ac levels are getting reduced at shield stage upon drug treatment. Thus presence of H3K4me3 mark at earlier stage could be helping in the addition of H3K27ac marks. Reduction in H3K4me3 at the dome stage could be preventing the addition of activatory acetylation mark over them, which could be reducing the gene expression at shield stage upon MM-102 treatment.

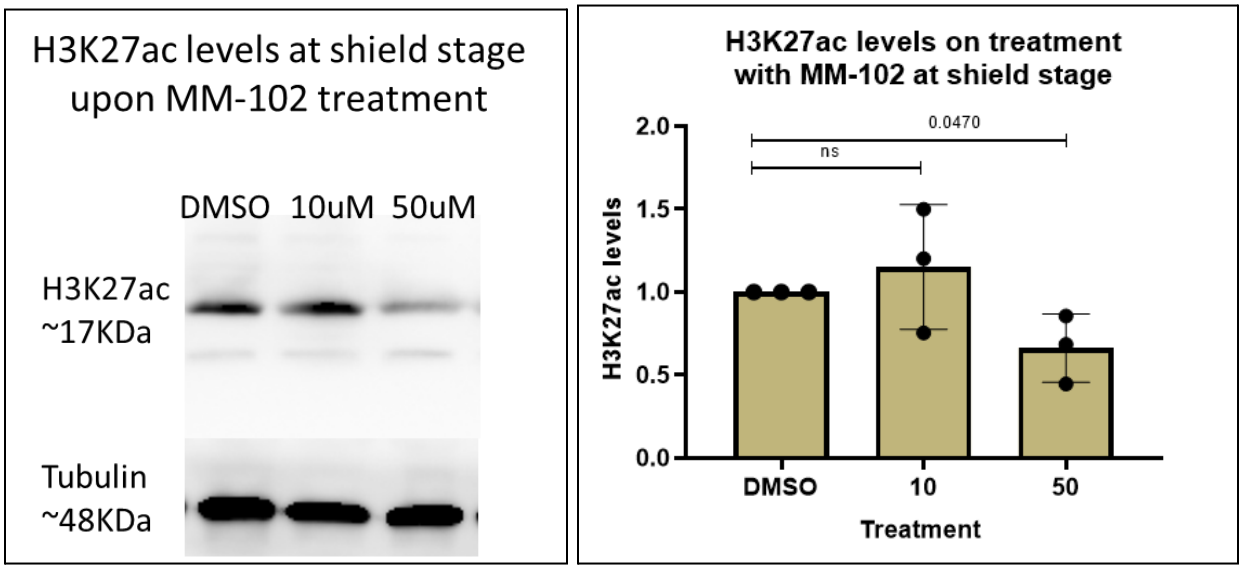


Figure 17: **Effect of MM-102 treatment on H3K27ac levels**. Western blot image of H3k27ac levels upon MM-102 treatment at shield stage (n=3). Quantification shows a significant reduction in the levels of H3K27ac with increased concentration (50 uM) of MM-102 with a p-value of 0.0470.No significant reduction was seen at 10μM MM-102 concentration.

12. Targeted perturbation technique using dCas9

Specific functional domains, like demethylases as in our case, can be cloned to dcas9 and can be targeted to specific locations using gRNAs. There are multiple lysine demethylases reported, which are termed as KDMs, as shown in Figure 17(a). KDM5 and KDM2 are two demethylases known to target H3K4me2/3, specifically in zebrafish. These are further classified into KDM5A, KDM5B, KDM5C, KDM5D, and KDM2A, KDM2B respectively. Oberkofler et al. earlier tried targeted perturbation of H3K4me3 at a specific locus by expressing a histone H3K4-specific demethylase at a specific locus in Arabidopsis to evaluate the role of H3K4 methylation in heat shock memory (Oberkofler & Bäurle, 2022). They showed that the effector construct (jmjd-dcas9) binds to the targeted *APX2* locus, where it decreases H3K4 methylation and, thereby, HS-induced transcriptional memory.

A similar strategy can be employed in the zebrafish system to study the effect of priming. Kdm5a, as shown in Figure 17(b), is a specific demethylase that has been reported to be reducing H3k4me3 levels in zebrafish (Kang et al., 2017; Romani et al., 2019). Kdm5a consists of 8 domains that have specific functions. The PHD3 domain of Kdm5a in homo sapiens has been shown to bind to H3K4me3. Though the same domain, PHD3, is present in zebrafish also, its function needs to be further validated. Thus either specific domains of the Kdm5a, like the PHD3 domain or a combination of domains or the whole of the KDM5a, can be cloned along with dcas9 and target the priming activity to check the efficiency of removing the trimethylation mark at specific primed sites using gRNAs. The temporal and spatial regulation offered by this strategy makes it an ideal one to study priming behavior.

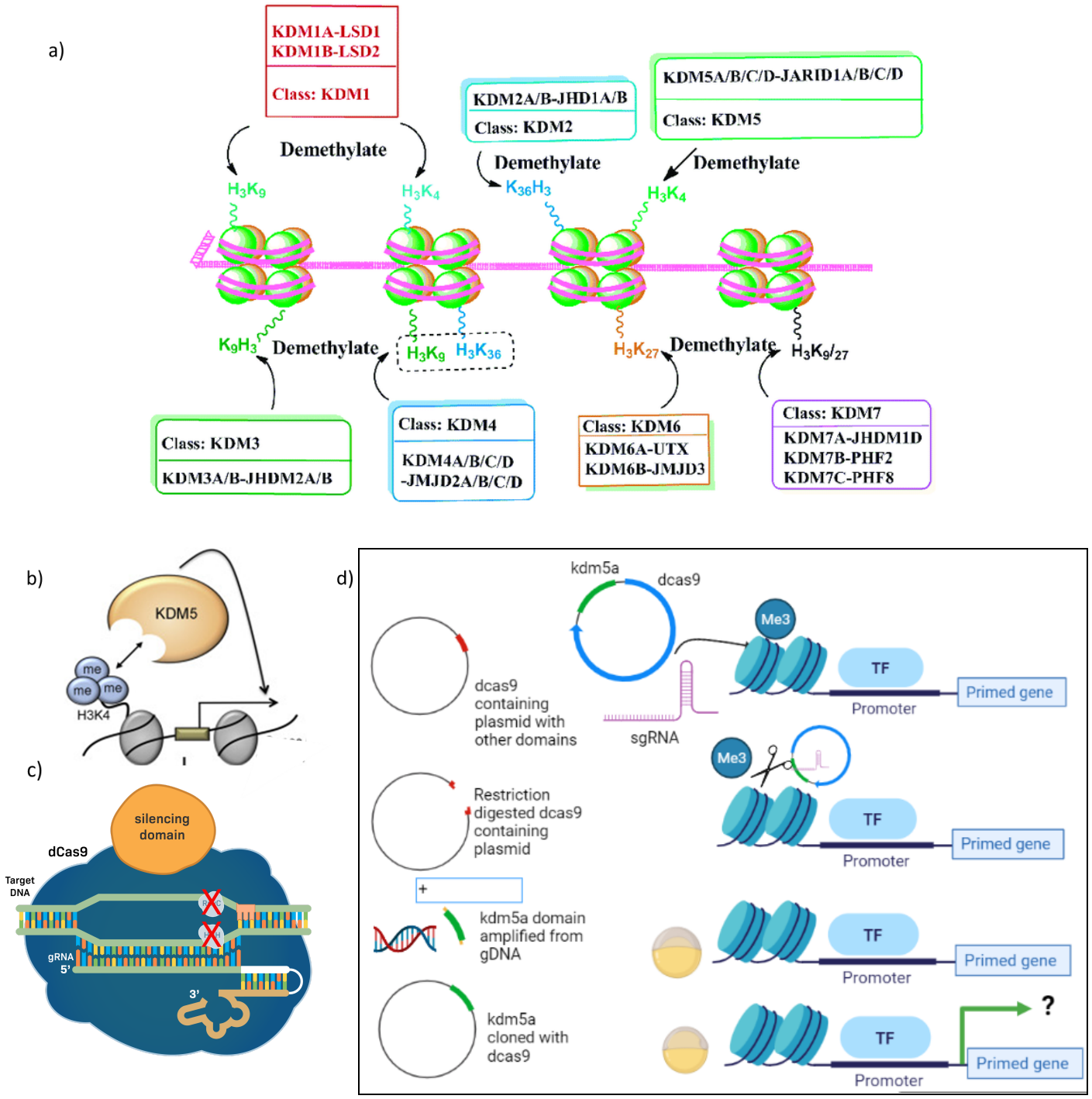


Figure 18: **targeted perturbation of H3K4me3 using dCas9-KDM5a construct.** a) Lysine demethylases (KDMs) and their target location of action (Basavarajappa & Subbanna, 2016) b) KDM5a domain functioning c) CRISPR-dcas9 targeted modification system (Image by LabXchange) d) Flowchart representing the planned workflow of dcas9 strategy.

13. Summary and Future directions

We have attempted to fish out regions in the zebrafish genome which are primed by H3K4me3 in the dome stage and are getting activated in the shield stage, thus finding the role of priming behavior in the cell fate specification process. By analyzing the ChIPseq data, such regions were found out, annotated and their significance was found. The expression of these genes were correlated with the RNAseq data at the respective stages available for these genes. We find the important pathways and the processes regulated by the set of genes using gene ontology analysis. One of the enriched processes in the gene ontology study ended up being the specification of the mesodermal cells. Precursors of different mesodermal cell types like notochord, prechordal plates, trunk somites, heart, blood, and pronephros are arranged in the dorso ventral axis in the shield stage. The various signalling pathways and the phenotypes associated with perturbation of many genes associated with mesodermal fate specification has been well documented in zebrafish. Hence it would be interesting to look into the effect of the role of priming in the expression of these genes. Organogenesis needs further time from the shield stage. Thus the organ specific genes getting activated in the shield stage could be playing a role in fate specification. The functional importance of the temporal enrichment of H3K4me3 marks in earlier stages in cell fate specification needs to be characterized further.

Perturbation using MM-102 treatment has been successful in showing the importance of priming. By reducing the expression of primed genes, it was able to convey the importance of priming in maintaining the correct expression of genes in a temporal manner. But some of the limitations of the study include, MM-102 being a global removal of H3K4me3 removes H3K4me3 from all genes that are active in the dome stage, and not just the primed genes. Affecting the other genes could have some impact on the primed genes under study. This could be noted as a limitation of our drug treatment study. Still, the RT-qPCR data from Ccne2, which is a not properly primed gene, implies that this limitation didn’t affect much on the genes which already have the acetylation mark. Thus a proposed model could be that the removal of H3K4me3 mark

perturbs the expression of primed genes by delaying the recruitment of acetylation marks. Another limitation that needs to be addressed is that the treatment is given from 2 cells to either dome/shield stage; thus, the MLL1 writer is targeted from an early stage itself till the stage where we collect them. This limitation in the temporal regulation should be considered while inferring the results of the study.

This study strengthens the concept of priming and implies the need to understand this regulatory mechanism in more detail. Some of the very interesting questions include how these regions are selected and how these priming marks are added. Has earlier reported that a broader H3K4me3 mark is related to more transcriptional activity and cell identity (Liu et al., 2016). The length of histone modification enrichment in relation to these priming functions is another area that can be explored. It was earlier reported that Setd1a is recruited to promoters of immune response genes upon bacterial infection, which catalyzes the addition of histone marks at these locations, and these primed marks act as a memory for the rapid activation of these genes upon subsequent exposure (Mehta & Jeffrey, 2015). Exploring the same concept of selection of regions to be primed in the context of development will be an interesting direction to be pursued. Overall, this pre-conditioning provided by the early addition of H3K4me3 marks helps in the efficient activation of these genes at a stage when it is required.

REFERENCES

- Akkers, R. C., van Heeringen, S. J., Jacobi, U. G., Janssen-Megens, E. M., Françoijs, K.-J., Stunnenberg, H. G., & Veenstra, G. J. C. (2009). A Hierarchy of H3K4me3 and H3K27me3 Acquisition in Spatial Gene Regulation in Xenopus Embryos. *Developmental Cell*, *17*(3), 425–434. https://doi.org/10.1016/j.devcel.2009.08.005
- Basavarajappa, B., & Subbanna, S. (2016). Epigenetic Mechanisms in Developmental Alcohol-Induced Neurobehavioral Deficits. *Brain Sciences*, *6*(2), 12. https://doi.org/10.3390/brainsci6020012
- Basta, J. M., Robbins, L., Denner, D. R., Kolar, G. R., & Rauchman, M. (2017). Sall1-NuRD interaction regulates multipotent nephron progenitors and is required for loop of Henle formation. *Development*, dev.148692. https://doi.org/10.1242/dev.148692
- Boeva, V. (2016). Analysis of Genomic Sequence Motifs for Deciphering Transcription Factor Binding and Transcriptional Regulation in Eukaryotic Cells. *Frontiers in Genetics*, *7*, 24. https://doi.org/10.3389/fgene.2016.00024
- Bonifer, C., & Cockerill, P. N. (2017). Chromatin priming of genes in development: Concepts, mechanisms and consequences. *Experimental Hematology*, *49*, 1–8. https://doi.org/10.1016/j.exphem.2017.01.003
- Caldon, C., & Musgrove, E. A. (2010). Distinct and redundant functions of cyclin E1 and cyclin E2 in development and cancer. *Cell Division*, *5*(1), 2. https://doi.org/10.1186/1747-1028-5-2
- Diotel, N., Rodriguez Viales, R., Armant, O., März, M., Ferg, M., Rastegar, S., & Strähle, U. (2015). Comprehensive expression map of transcription regulators in the adult zebrafish telencephalon reveals distinct neurogenic niches. *The Journal of Comparative Neurology*, *523*(8), 1202–1221. https://doi.org/10.1002/cne.23733
- Dong, P., & Liu, Z. (2017). Shaping development by stochasticity and dynamics in gene regulation. *Open Biology*, *7*(5), 170030. https://doi.org/10.1098/rsob.170030
- Economou, A. D., Guglielmi, L., East, P., & Hill, C. S. (2022). Nodal signaling establishes a competency window for stochastic cell fate switching. *Developmental Cell*, *57*(23), 2604-2622.e5. https://doi.org/10.1016/j.devcel.2022.11.008
- Fagundes, R., & Teixeira, L. K. (2021). Cyclin E/CDK2: DNA Replication, Replication Stress and Genomic Instability. *Frontiers in Cell and Developmental Biology*, *9*, 774845.

https://doi.org/10.3389/fcell.2021.774845

- Fanucchi, S., Domínguez-Andrés, J., Joosten, L. A. B., Netea, M. G., & Mhlanga, M. M. (2021). The Intersection of Epigenetics and Metabolism in Trained Immunity. *Immunity*, *54*(1), 32–43. https://doi.org/10.1016/j.immuni.2020.10.011
- Grandy, R. A., Whitfield, T. W., Wu, H., Fitzgerald, M. P., VanOudenhove, J. J., Zaidi, S. K., Montecino, M. A., Lian, J. B., van Wijnen, A. J., Stein, J. L., & Stein, G. S. (2016). Genome-Wide Studies Reveal that H3K4me3 Modification in Bivalent Genes Is Dynamically Regulated during the Pluripotent Cell Cycle and Stabilized upon Differentiation. *Molecular and Cellular Biology*, *36*(4), 615–627. https://doi.org/10.1128/MCB.00877-15
- Halsall, J. A., Andrews, S., Krueger, F., Rutledge, C. E., Ficz, G., Reik, W., & Turner, B. M. (2021). Histone modifications form a cell-type-specific chromosomal bar code that persists through the cell cycle. *Scientific Reports*, *11*(1), 3009. https://doi.org/10.1038/s41598-021-82539-z
- Han, K., Ren, R., Cao, J., Zhao, S., & Yu, M. (2019). Genome-Wide Identification of Histone Modifications Involved in Placental Development in Pigs. *Frontiers in Genetics*, *10*, 277. https://doi.org/10.3389/fgene.2019.00277
- Han, P., Hang, C. T., Yang, J., & Chang, C.-P. (2011). Chromatin Remodeling in Cardiovascular Development and Physiology. *Circulation Research*, *108*(3), 378–396. https://doi.org/10.1161/CIRCRESAHA.110.224287
- Harvey, S. A., & Logan, M. P. O. (2006). Sall4 acts downstream of tbx5 and is required for pectoral fin outgrowth. *Development (Cambridge, England)*, *133*(6), 1165–1173. https://doi.org/10.1242/dev.02259
- Henderson, D. J., & Copp, A. J. (1997). Role of the extracellular matrix in neural crest cell migration. *Journal of Anatomy*, *191 (Pt 4)*(Pt 4), 507–515. https://doi.org/10.1046/j.1469-7580.1997.19140507.x
- Hikasa, H., & Sokol, S. Y. (2013). Wnt Signaling in Vertebrate Axis Specification. *Cold Spring Harbor Perspectives in Biology*, *5*(1), a007955–a007955. https://doi.org/10.1101/cshperspect.a007955
- Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T. V., Zon, L. I., Yuan, G.-C., Orkin, S. H., & Xu, J. (2016). Dynamic Control of Enhancer

Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. *Developmental Cell*, *36*(1), 9–23. https://doi.org/10.1016/j.devcel.2015.12.014

- Kang, M., Mehrazarin, S., Park, N.-H., & Wang, C.-Y. (2017). Epigenetic gene regulation by histone demethylases: Emerging role in oncogenesis and inflammation. *Oral Diseases*, *23*(6), 709–720. https://doi.org/10.1111/odi.12569
- Karatas, H., Townsend, E. C., Cao, F., Chen, Y., Bernard, D., Liu, L., Lei, M., Dou, Y., & Wang, S. (2013). High-affinity, small-molecule peptidomimetic inhibitors of MLL1/WDR5 protein-protein interaction. *Journal of the American Chemical Society*, *135*(2), 669–682. https://doi.org/10.1021/ja306028q
- Lee, K., Cho, H., Rickert, R. W., Li, Q. V., Pulecio, J., Leslie, C. S., & Huangfu, D. (2019). FOXA2 Is Required for Enhancer Priming during Pancreatic Differentiation. *Cell Reports*, *28*(2), 382-393.e7. https://doi.org/10.1016/j.celrep.2019.06.034
- Li, X., & Song, Y. (2021). Structure, function and inhibition of critical protein–protein interactions involving mixed lineage leukemia 1 and its fusion oncoproteins. *Journal of Hematology & Oncology*, *14*(1), 56. https://doi.org/10.1186/s13045-021-01057-7
- Lindeman, L. C., Andersen, I. S., Reiner, A. H., Li, N., Aanes, H., Østrup, O., Winata, C., Mathavan, S., Müller, F., Aleström, P., & Collas, P. (2011). Prepatterning of Developmental Gene Expression by Modified Histones before Zygotic Genome Activation. *Developmental Cell*, *21*(6), 993–1004. https://doi.org/10.1016/j.devcel.2011.10.008
- Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., Zhang, Y., Gao, Y., & Gao, S. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature*, *537*(7621), 558–562. https://doi.org/10.1038/nature19362
- Lu, Y.-C., Sanada, C., Xavier-Ferrucio, J., Wang, L., Zhang, P.-X., Grimes, H. L., Venkatasubramanian, M., Chetal, K., Aronow, B., Salomonis, N., & Krause, D. S. (2018). The Molecular Signature of Megakaryocyte-Erythroid Progenitors Reveals a Role for the Cell Cycle in Fate Specification. *Cell Reports*, *25*(8), 2083-2093.e4. https://doi.org/10.1016/j.celrep.2018.10.084
- Mehta, S., & Jeffrey, K. L. (2015). Beyond receptors and signaling: Epigenetic factors in the regulation of innate immunity. *Immunology & Cell Biology*, *93*(3), 233–244.

https://doi.org/10.1038/icb.2014.101

- Melby, A. E., Warga, R. M., & Kimmel, C. B. (1996). Specification of cell fates at the dorsal margin of the zebrafish gastrula. *Development*, *122*(7), 2225–2237. https://doi.org/10.1242/dev.122.7.2225
- Oberkofler, V., & Bäurle, I. (2022). Inducible epigenome editing probes for the role of histone H3K4 methylation in Arabidopsis heat stress memory. *Plant Physiology*, *189*(2), 703–714. https://doi.org/10.1093/plphys/kiac113
- Papamichos-Chronakis, M., Watanabe, S., Rando, O. J., & Peterson, C. L. (2011). Global Regulation of H2A.Z Localization by the INO80 Chromatin-Remodeling Enzyme Is Essential for Genome Integrity. *Cell*, *144*(2), 200–213. https://doi.org/10.1016/j.cell.2010.12.021
- Papatheodorou, I., Fonseca, N. A., Keays, M., Tang, Y. A., Barrera, E., Bazant, W., Burke, M., Füllgrabe, A., Fuentes, A. M.-P., George, N., Huerta, L., Koskinen, S., Mohammed, S., Geniza, M., Preece, J., Jaiswal, P., Jarnuczak, A. F., Huber, W., Stegle, O., … Petryszak, R. (2018). Expression Atlas: Gene and protein expression across multiple studies and organisms. *Nucleic Acids Research*, *46*(D1), D246–D251. https://doi.org/10.1093/nar/gkx1158
- Park, J., Lee, K., Kim, K., & Yi, S.-J. (2022). The role of histone modifications: From neurodevelopment to neurodiseases. *Signal Transduction and Targeted Therapy*, *7*(1), 217. https://doi.org/10.1038/s41392-022-01078-9
- Perrimon, N., Pitsouli, C., & Shilo, B.-Z. (2012). Signaling Mechanisms Controlling Cell Fate and Embryonic Patterning. *Cold Spring Harbor Perspectives in Biology*, *4*(8), a005975. https://doi.org/10.1101/cshperspect.a005975
- Petel Légaré, V., Rampal, C. J., Gurberg, T. J. N., Aaltonen, M. J., Janer, A., Zinman, L., Shoubridge, E. A., & Armstrong, G. A. B. (2023). Loss of mitochondrial Chchd10 or Chchd2 in zebrafish leads to an ALS‐like phenotype and Complex I deficiency independent of the mitochondrial integrated stress response. *Developmental Neurobiology*, *83*(1–2), 54–69. https://doi.org/10.1002/dneu.22909
- Preissl, S., Gaulton, K. J., & Ren, B. (2023). Characterizing cis-regulatory elements using single-cell epigenomics. *Nature Reviews Genetics*, *24*(1), 21–43. https://doi.org/10.1038/s41576-022-00509-1

- Romani, M., Daga, A., Forlani, A., Pistillo, M. P., & Banelli, B. (2019). Targeting of Histone Demethylases KDM5A and KDM6B Inhibits the Proliferation of Temozolomide-Resistant Glioblastoma Cells. *Cancers*, *11*(6), 878. https://doi.org/10.3390/cancers11060878
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., Mellor, J., & Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature*, *419*(6905), 407–411. https://doi.org/10.1038/nature01080
- Schebesta, M., Lien, C.-L., Engel, F. B., & Keating, M. T. (2006). Transcriptional profiling of caudal fin regeneration in zebrafish. *TheScientificWorldJournal*, *6 Suppl 1*, 38–54. https://doi.org/10.1100/tsw.2006.326
- Schier, A. F., & Talbot, W. S. (2005). Molecular genetics of axis formation in zebrafish. *Annual Review of Genetics*, *39*, 561–613. https://doi.org/10.1146/annurev.genet.37.110801.143752
- Sun, Y.-M., Cooper, M., Finch, S., Lin, H.-H., Chen, Z.-F., Williams, B. P., & Buckley, N. J. (2008). Rest-mediated regulation of extracellular matrix is crucial for neural development. *PloS One*, *3*(11), e3656. https://doi.org/10.1371/journal.pone.0003656
- Thorsteinsdóttir, S., Deries, M., Cachaço, A. S., & Bajanca, F. (2011). The extracellular matrix dimension of skeletal muscle development. *Developmental Biology*, *354*(2), 191–207. https://doi.org/10.1016/j.ydbio.2011.03.015
- Tripathi, S. K., & Lahesmaa, R. (2014). Transcriptional and epigenetic regulation of T-helper lineage specification. *Immunological Reviews*, *261*(1), 62–83. https://doi.org/10.1111/imr.12204
- Umemoto, T., Hashimoto, M., Matsumura, T., Nakamura-Ishizu, A., & Suda, T. (2018). Ca2+–mitochondria axis drives cell division in hematopoietic stem cells. *Journal of Experimental Medicine*, *215*(8), 2097–2113. https://doi.org/10.1084/jem.20180421
- Wang, L., Du, Y., Ward, J. M., Shimbo, T., Lackford, B., Zheng, X., Miao, Y., Zhou, B., Han, L., Fargo, D. C., Jothi, R., Williams, C. J., Wade, P. A., & Hu, G. (2014). INO80 Facilitates Pluripotency Gene Activation in Embryonic Stem Cell Self-Renewal, Reprogramming, and Blastocyst Development. *Cell Stem Cell*, *14*(5), 575–591. https://doi.org/10.1016/j.stem.2014.02.013

- Xie, X., Liu, J.-X., Hu, B., & Xiao, W. (2011). Zebrafish foxo3b Negatively Regulates Canonical Wnt Signaling to Affect Early Embryogenesis. *PLoS ONE*, *6*(9), e24469. https://doi.org/10.1371/journal.pone.0024469
- Zhu, L., Gomez-Duran, A., Saretzki, G., Jin, S., Tilgner, K., Melguizo-Sanchis, D., Anyfantis, G., Al-Aama, J., Vallier, L., Chinnery, P., Lako, M., & Armstrong, L. (2016). The mitochondrial protein CHCHD2 primes the differentiation potential of human induced pluripotent stem cells to neuroectodermal lineages. *Journal of Cell Biology*, *215*(2), 187–202. https://doi.org/10.1083/jcb.201601061