Norwegian University
of Life Sciences

Master's Thesis 2018 60 ECTS

Faculty of Environmental Sciences and Natural Resource Management (MINA)
Supervisor: Leif Christopher Lindeman

Chromatin Accessibility after Inhibition of Enhancer of Zeste Homolog 2 (Ezh2) in Zebrafish Embryos

Jarle Ballangby
Master of Science in Molecular Biology
Faculty of Chemistry, Biotechnology and Food Science

Acknowledgments

I want to thank the Centre of Excellence (CoE) Centre for Environmental Radioactivity (CERAD), Professor Peter Aleström, and Post-Doc Leif Lindeman who gave me the opportunity to write my master thesis about the exciting field of epigenetics. Epigenetics is an essential field in biology that I want to focus my career on. During my project, I became a part of the groups of Basic Sciences and Aquamedicine (BasAM) at Adamstuen and Faculty of Environmental Sciences and Natural Resource Management (MINA). Thank you for all the help I have received, Leif Lindeman, Erik Rasmussen, and Jorke Kamstra. You have always helped me no matter how hopeless the question was with everything from general laboratory questions to bioinformatics. It was great to have so many highly skilled people around me. Not least, I would like to thank all the people working at Adamstuen building 14 for creating such a great working and social environment. I want to thank Ana Tavara and Arturas Kavaliauskis for help with the collection of zebrafish embryos for my research as well as guidance for dechorionation of the embryos. Thanks to Peter Aleström to give me motivation and support while writing. Finally, I would like to thank friends and family, especially my girlfriend Andrea Hansen, who have always supported me during this master program.

Table of Contents

Summary

The thesis is a part of a Centre of Excellence Centre for Environmental Radioactivity (CoE CERAD) project, focusing on trans-generational studies of low-dose gamma-irradiation using zebrafish as a model organism. As a part of the epigenetic toolbox, Assay for Transposase-Accessible Chromatin (ATAC-seq) is established using a known epigenetic inhibitor (Ezh2i), PF-06726304 acetate, in a pilot study with the hypothesis that PF-06726304 acetate would induce a decline in H3K27me3 marks in chromatin, correlated with condensed chromatin and silent genes.

Subgoals were to establish methods for manual dechorionation and protocols for enzymatic dechorionation, establish methods for ATAC-seq library generation and clean-up, and establish a bioinformatic pipeline for preliminary ATAC-seq analysis. The ATAC-seq results from the treatment of PF-06726304 were validated with Western Blot and Chromatin Immunoprecipitation.

The thesis shows that low-dose exposure to the Ezh2 inhibitor, PF-06726304 acetate, indeed induces chromatin reorganization. Our results suggest that a GUI tool for the Analysis and Visualization of ATAC-seq data (GUAVA) can be used as a primary dataset analysis tool, but that more thorough post-analysis is necessary.

Sammendrag

Denne oppgaven er en del av et prosjekt under Senter for radioaktivitet, mennesker og miljø (SFF CERAD), med fokus på transgenerasjonsstudier av lavdose gamma-bestråling ved bruk av sebrafisk som modellorganisme. Som en del av den epigenetiske verktøykassen er metoden, «Assay for Transposase-Accessible Chromatin (ATAC-seq)» etablert ved bruk av en kjent epigenetisk inhibitor (Ezh2i), PF-06726304 acetat, i en pilotstudie med hypotesen om at PF-06726304 acetat ville indusere en nedgang i H3K27me3 i kromatin, som har sammenheng med pakket sammen kromatin og inaktive gener.

Delmål for oppgaven var å etablere en metode for manuell dechorionering, samt å etablere en metode for mer robust enzymatisk dechorionering, etablere metoder for rensning av ATAC-seq bibliotek, og etablere bioinformatisk nedstrømsanalyse for ATAC-seq analyse. ATAC-seq resultatene etter eksponering med PF-06726304 ble validert med Western Blot and Kromatin Immunopresipitering.

Oppgaven viser at eksponering med lav dose til Ezh2-inhibitoren, PF-06726304 acetat, faktisk promoterer kromatinomorganisering. Resultatene tyder på at et GUAVA, som er et verktøy med grafisk brukergrensesnitt for analyse og visualisering av ATAC-seq-data (GUAVA) kan brukes som et foreløpig analyseverktøy, men at en grundigere analyse i etterkant er nødvendig.

Abbreviations

5mC	5-methylcytosine
6mA	N6-methyladenine
aspa	*aspartoacylase*
ATAC	Assay for Transposase-Accessible Chromatin
c/ebpα	*CCAAT/enhancer-binding protein α*
ChIP	Chromatin Immunoprecipitation
DMSO	Dimethyl Sulfoxide
EDTA	Ethylenediaminetetraacetic acid
EGTA	Triethylene Glycol Diamine Tetraacetic Acid
Hpf	Hours Post Fertilization
hnf4α	*hepatocyte nuclear factor 4, α*
hoxb1b	*homeobox B1b*
ilf2	*interleukin enhancer binding factor 2*
mthfd1b	*methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1b*
ncRNA	Non-coding RNA
PBS	Phosphate Buffered Saline
PMSF	Phenylmethanesulfonyl Fluoride
PTM	Post Translational Modification
RIPA	Radioimmunoprecipitation Assay
SDS	Sodium Dodecyl Sulfate
slc12a3	*solute carrier family 12 (sodium/chloride transporter), member 3*

TE	Tris-EDTA
TSS	Transcriptional Start Site
UTR	Untranslated Region

1 Introduction

1.1 Background

The zebrafish genome DNA consists of mainly non-coding RNA, with only 1-2 percent coding sequences regulated in a developmental as well as environmental induced manner (Genome Reference Consortium 2017). The Central Dogma of molecular biology (Crick 1958), states that sequence information in proteins cannot transfer back to the DNA sequence. The study of chromatin dynamics is within the field of epigenetics. Epigenetics explains phenotypes resulting from changes in chromatin (Li et al. 2018), without alterations in the DNA sequence (Berger et al. 2009). Changes of chromatin lead to changes in the transcriptome (Eccleston et al. 2013), causing effects on development (Wu et al. 2018), X chromosome inactivation (Payer et al. 2011), genomic imprinting (MacDonald 2012) and chromosome stability (Feil & Fraga 2012). How these epigenetic regulatory systems work is of great interest and could explain the differentiation of cells. Diseases and cancer follow improper regulation of epigenetics (Rodenhiser & Mann 2006). Individualized and personalized medicine that target disease-related epigenetic markers has a profound interest (Altucci & Rots 2016). Epigenetic inhibitors that inhibit specific enzyme complexes are examples that show great promise. For example, in humans abnormal expression of Polycomb Protein Complex 2 (PPC2) component, enhancer of zeste homolog 2, EZH2, which is mediating trimethylation of H3K27, is associated with malignancies (Busam et al. 2017; Karsli-Ceppioglu et al. 2017; Ntziachristos et al. 2012; Schaefer et al. 2015; Yoo & Hennighausen 2012). Inhibition of EZH2 is regarded as a promising treatment (Pourakbar et al. 2017).

1.2 Zebrafish as Model organism

Zebrafish (*Danio rerio*) is a freshwater teleost belonging to the Cyprinidae family (Gutierrez-Lovera et al. 2017). The organism was first presented as a model organism in the 1980s (Kimmel et al. 1981; Kimmel et al. 1988) and has been a popular vertebrate for studying epigenetic markers (Cavalieri & Spinelli 2017). The organism is widely used in human biomedicine with 70% shared genes (Aleström & Winther-Larsen 2016). The transparent embryos allow researchers to follow the development visually. The hundreds of embryos produced during one breeding promote a potentially statistical strong biological replicate (Kimmel et al. 1995). The genome is now in 11th version (Genome Reference Consortium 2017) and is well annotated.

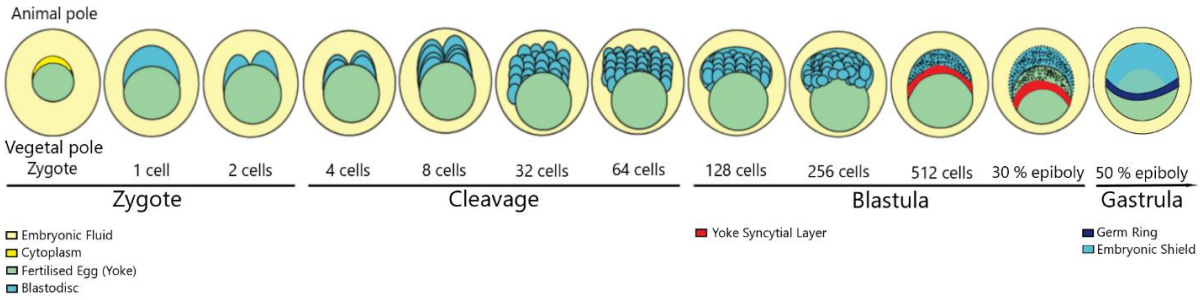

Figure 1: Zygote period (0 – 0.75 hpf): Fertilization and transition from 1 cell to 2 cells. The Animal and vegetal pole separate the blastodisc from the cytoplasm. Cleavage period (0.75 – 2.25 hpf): Transition from 2 cells to 64 cells. Rapid cleavage every 15 minutes. Blastula period (2.25 – 5.25 hpf): Transition from 128 cells to 50% epiboly. Rapid cleavage up to the activation of the zygotic genome (3.5 hpf) and the establishment of histone PTM signatures. Initial Gastrula (5.25 hpf): Formation of germ ring.

The initial development of a zebrafish up to early gastrula period can be divided into zygote-, cleavage-, and blastula period (Figure 1). In the zygote period (0 – 0.75 hpf), the animal and vegetal pole segregate the blastodisc (singular terminology for dividing cells) from the cytoplasm, and the transition from 1 cell to 2 cells takes place. The cells continue to segregate to the cleavage period (0.75 – 2.25 hpf) where cells cleave synchronized every 15 minutes. The cells continue with the rapid cleavage until midterm (3.5 hpf) blastula period (2.25 – 5.25) where the zygotic genome is activated (Kimmel et al. 1995) and histone PTM signatures established (Lindeman 2011, Vastenhough 2010). Single-cell RNA- seq reveal that the cells are mainly similar up to gastrula period (6 hpf) and are differentiated to all germ layers at 8 hpf (Figure 2) (Wagner et al. 2018).

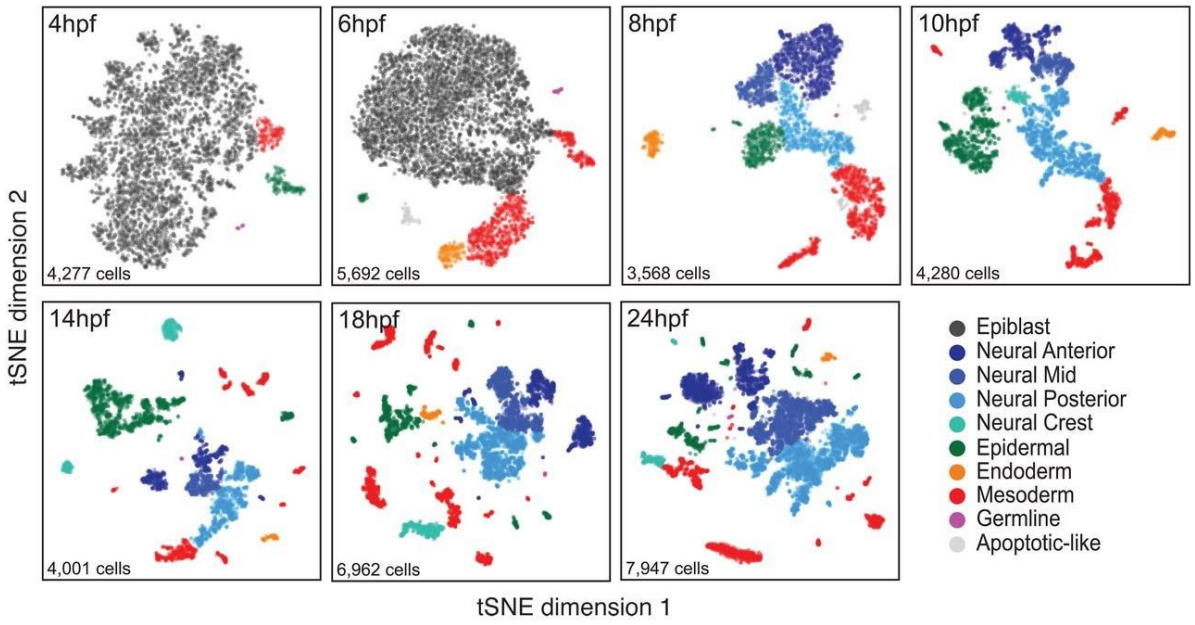

Figure 2: Single-cell RNA sequencing of zebrafish embryos at different developmental stages correlates with cellular differentiation. At 4 hpf and 6 hpf, the embryo mainly consists of one cell type, while from 8 hpf and onwards the cells are differentiated to neural, epidermal, endodermal, and mesoderm cells. The germ cell is visible at all investigated stages. Adapted from Wagner et al. 2018 with permission.

1.3 Epigenetics

The term epigenetics generally means “upon” or “above” the genes that hold the information that makes up an organism. The term was first introduced in 1942 by Conrad Waddington to describe events that could not be explained by known genetic principles. The term included all molecular pathways regulating the expression of genes into specific phenotypes (Waddington 1942). The definition has since evolved, and are now more commonly known as the changes in phenotypes, without alterations in the genome (Berger et al. 2009). Epigenetics revolves around the surroundings of the DNA with the epigenetic code, affecting the level of transcription through activating- or repressive marks.

1.3.1 DNA methylation

DNA methylation refers to the covalent modification of the DNA molecule with methyl groups. The DNA methylation occurs mostly on cytosine to form 5-methylcytosine (5mC), but also on adenine (A) to form N^6-methyladenine (6mA) (Meyer & Jaffrey 2016; Wu et al. 2016). DNA methylation is found in specific genomic sites in vertebrates and is for example linked to transcription regulation in development- and activity-dependent scenarios (Meyer & Jaffrey 2016), cancer (Ehrlich 2002) and aging (Horvath 2013).

1.3.2 Histone post-translational modifications

Two of each of histone H2A, H2B, H3 and H4 form an octamer protein complex of which DNA is wrapped around at a length of 147 base pairs (bp) forming a nucleosome (Figure 3, left), the repeating unit of chromatin (Kornberg 1974). The histones tails are susceptible to covalent modification of the amino acids by histone modifying enzymes. The modifications can be methylation, acetylation, phosphorylation, as well as more than 80 other variants (Tsompana & Buck 2014). The modifications represent an epigenetic "histone code" (Jenuwein & Allis 2001) that, in combination with the other epigenetic mechanisms, determines the structural outcome of chromatin as specific unique docking sites that promotes recruitment of different protein complexes (Berger 2007; Turner 2007). Well described are modifications of lysines (K) on the tail of histone (H) 3, with acetylation (ac) of H3K9 and trimethylation (me3) of H3K4 associated with activated chromatin, and H3K9me3 and H3K27me3 associated with repressed chromatin (Wiles & Selker 2017).

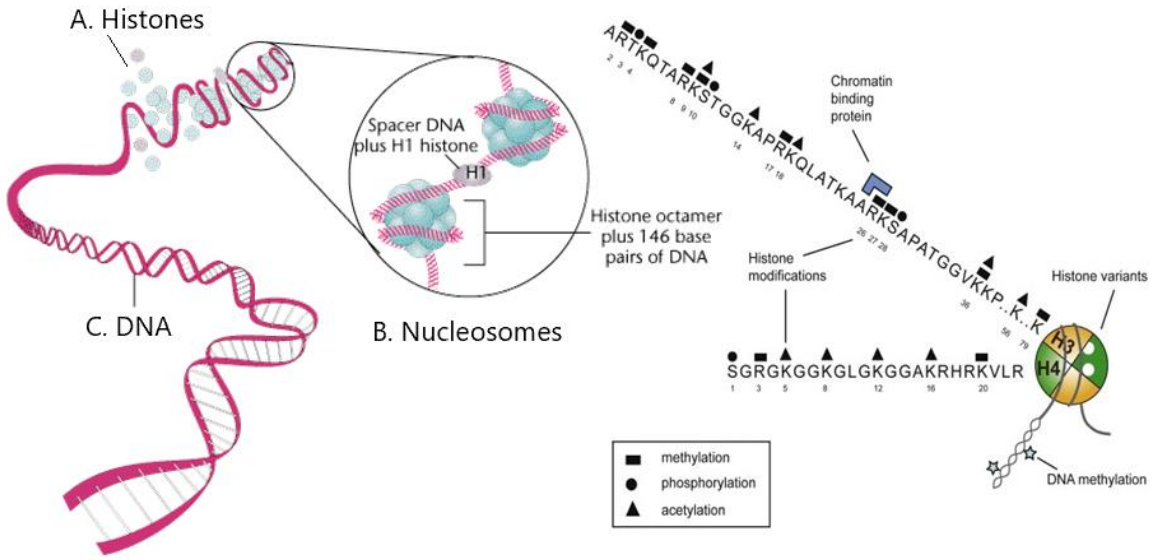

Figure 3: Left: Conceptual drawing, showing the histones (**A**) with DNA wrapped around creating nucleosomes (**B**), leading to the unraveling of DNA (**C**). Right: Drawing of a nucleosome with two amino acid tails from histone H3 and histone H4. Three of the possible post-translational modifications shows as either a square (methylation), circle (phosphorylation and triangle (acetylation) (Scharf & Imhof 2011).

1.3.3 Non-coding RNA

Non-coding RNA (ncRNA) are RNA molecules not translated into proteins, e.g., transfer RNAs, ribosomal RNAs, small RNAs, and long ncRNAs. Small RNA, e.g., micro RNA (miRNA), are essential for vertebrate embryogenesis (Giraldez et al. 2005) and known to regulate of gene expression post-transcriptionally by the degradation of mRNA (Ambros 2004; Bartel 2004; Zhu et al. 2018). Both small

and long noncoding RNA influence transgenerational inheritance as they act as recruiters for various proteins, e.g., chromatin remodeling factors to achieve repression by DNA methylation and repressive histone PTMs (Castel & Martienssen 2013; Khalil et al. 2009; Tsai et al. 2010; Yu & Li 2015). Long ncRNA are also found upregulated in breast-, esophageal-, lung-, and gastric cancer (Yu & Li 2015).

1.3.1 Epitranscriptomics

Epitranscriptomics refer to the epigenetic modifications of RNA. There are currently 112 known modifications, mostly methyl groups (The RNA Institute 2018). RNA methylation regulates the processing, export, and stability of RNA (Peer et al. 2017). One of the most abundant is 6 mA modification of mRNA (Peer et al. 2017). 6mA is found in a variety of organisms in mRNA, including viruses (Moss et al. 1977), yeast (Bodi et al. 2010), plants (Nichols 1979), humans, and mice (Dominissini et al. 2012). The modification affects transcription as well as translation (Liu & Pan 2015).

1.3.2 Chromatin and topography

The epigenetic mechanisms together form dynamic chromatin including both open and more compact structures. There are two subdivisions of chromatin structures, primary- and higher-order structures. Primary order includes the euchromatin and the heterochromatin. Euchromatin can be exemplified with the ”beads on the string” concept (Figure 4) (Alberts et al. 2015), the “string” represents the DNA, and the “beads” represent the nucleosome cores. Euchromatin is more open and loosely packed than heterochromatin allowing access of, e.g., transcription factors (Li & Reinberg 2011). Euchromatin correlates with high levels of acetylation of the histone tails, high levels of methylation on H3K4, H3K36, and H3K79, as well as low levels of repressive marks, e.g., methylation on H3K9 (Alvarenga et al. 2016) and H3K27 (Wiles & Selker 2017).

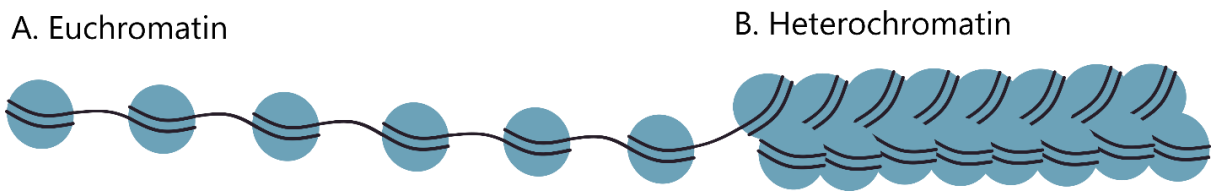

Figure 4: Schematics of the primary order of chromatin. The DNA (black) wraps around the nucleosome cores (blue). **A.** Euchromatin. A more open and loosely packaged form of chromatin. **B.** Heterochromatin. A More compact and tightly packed form of chromatin.

Heterochromatin is a compact form of chromatin and is exemplified with the 30 nm fiber and higher orders of packaging, making the DNA not accessible for the transcriptional machinery due to the tight packaging of the nucleosomes — the compact packaging results in "silent" chromatin. Heterochromatin is associated with low levels of the activating histone PTMs while higher levels of repressive marks. Heterochromatin is categorized into constitutive and facultative heterochromatin. The former, constitutive, is condensed chromatin found in the centromeric and telomeric parts of the chromosome, mainly composed of satellite repeats and transposon repeats (Saksouk et al. 2015). Facultative heterochromatin is more dynamic and switches between eu- and heterochromatin depending on the requirement of the cell (Saksouk et al. 2015; Trojer & Reinberg 2007).

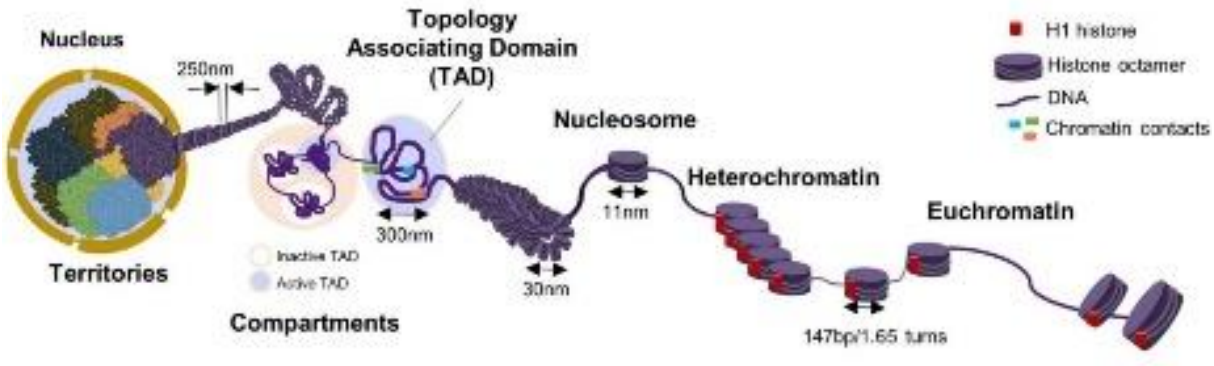

Figure 5: Levels of chromatin compaction from left, higher order structures, to right, primary order structures (heterochromatin and euchromatin) (adapted from Chang et al. 2018).

Topology Associating Domains (TADs, figure 5) are higher order structures, first discovered in 2012, defined as domains enriched of interacting chromatin (Dixon et al. 2012; Pombo & Dillon 2015). The interacting chromatin forms loops, linking enhancers and promoters (Rao et al. 2014). TADs are found to be stable over different cell types as well as different mammalian species suggesting that it is a natural property of the chromatin (Dixon et al. 2012). Evidence suggests that genes within a TAD responds to similar transcriptional stimuli and are more likely to be co-regulated (Le Dily et al. 2014). Aberrant interactions in TADs may be followed by aberrant gene regulation (Pombo & Dillon 2015). The formation of TADs are linked to the protein transcriptional regulator CCCTC-binding factor (CTCF) and cohesin (Rao et al. 2014). The mechanism is not fully understood, but one promising model is the loop extrusion model. In the loop extrusion model, the primary order chromatin slides through loop extruding factors, such as cohesin, to form larger loops. A process that continues until cohesin encounters TAD boundary proteins such as CTCF (Fudenberg et al. 2016). Note that this thesis focus on the primary order of chromatin described in the previous paragraph.

1.3.3 Transposition

DNA is subject to recombination of genomic regions. The recombination can happen in the form of two classes, conservative site-specific recombination or transposition. Conservative site-specific recombination is known as either deletion, insertion, or inversion of DNA, while transposition is the recombination of genetic elements called transposons. Transposons exist as three classes, virus-like retrotransposons, poly-A retrotransposons, and DNA transposons (James D. Watson 2008). The latter is the relevant transposon class for this thesis.

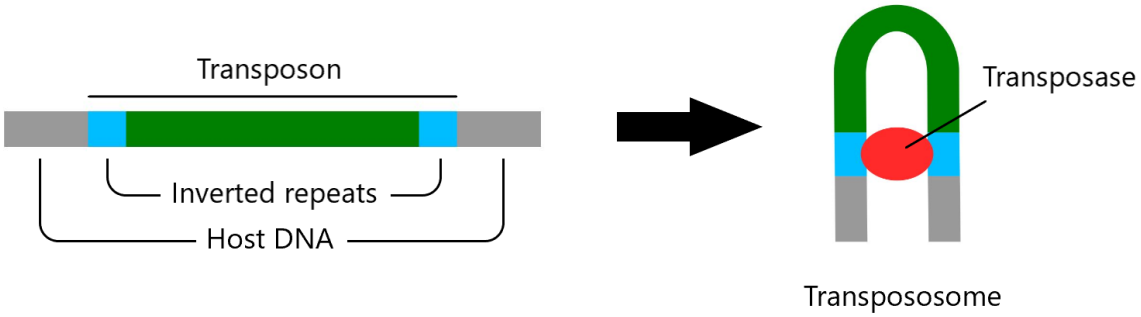

Figure 6: The transposase form a hairpin complex generated from DNA transposons with terminal inverted repeats (IS50L and IS50R) inserted within a host's DNA. The green middle section includes recombination sites, as well as three genes coding for antibiotic resistance (James D. Watson 2008).

Transposases catalyze the transposition that happens by one of two mechanisms, by cut-and-paste-mechanism, moving the transposon, or a replicative mechanism, duplicating the transposon (James D. Watson 2008). The function of the *in vivo* Tn5 transposase is to locate the Tn5 transposon (its coding region), copy it and move the copy. The general mechanism of the Tn5 transposition is the cut-and-paste mechanism. The transposase binds to the inverted terminal repeats of the end of the transposon, resulting in the formation of a hairpin protein-DNA complex called a transpososome (Figure 6) (James D. Watson 2008).

The transposase attached to the transpososome will catalyze the cleavage of DNA at the junction between the host DNA and the inverted repeats of the transposon resulting in excised 3'-OH ends. These 3'-OH ends will carry out nucleophilic attack of the opposite strand through transesterification where the 3'-OH ends attack a phosphodiester bond directly across on the opposite strand, resulting in cleavage and the formation of a DNA-hairpin intermediate. The hairpin ends of the intermediate are hydrolyzed by the transposase resulting in a completely excised transposon (Figure 7) (James D. Watson 2008).

Figure 7: Transposon forms a DNA-hairpin intermediate. The hairpin is hydrolyzed to form an excised transposon.

The 3'-OH ends of the excised transposon are now ready to be joined with the target DNA strand through another transesterification. The transpososome ensures that both 3'-OH ends of a transposon attacks the two DNA strand of the same target site together resulting in a gap on each strand initial to the 5' end of the transposon. The gap is filled with nucleotides by DNA-repair polymerase (Figure 8). The *in vivo* Tn5 system is functionally inactive (James D. Watson 2008). This thesis will focus on an enhanced hyperactive variant of Tn5 transposase developed (Goryshin & Reznikoff 1998).

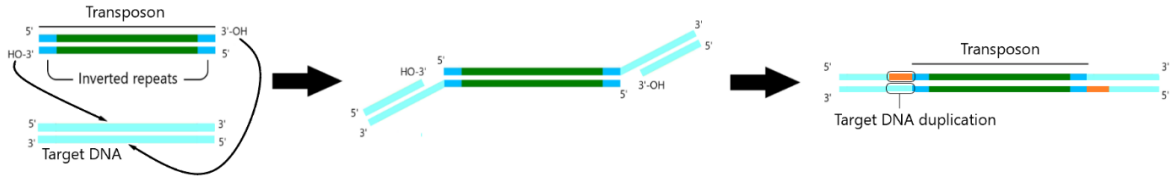

Figure 8: The OH-3' ends of the excised transposon are joined to the target DNA by transesterification leaving a gap on each strand initial to the 5' end of the transposon. DNA-repair polymerase fills the gap with nucleotides resulting in a duplication of the target site.

The DNA transposon Tn5 (Figure 6) includes terminal inverted repeats (IS50L and IS50R), recombination sites, as well as three genes coding for antibiotic resistance. The IS50R encodes for the Tnp transposase as well as the Inh inhibitor, while the IS50L codes for the inactive C-terminal of Tnp and Inh (Goryshin & Reznikoff 1998).

1.3.4 Development of hyperactive Tn5 transposase

It was discovered that only three parts of the Tn5 system, the Tnp protein, the 19 bp outside end (OE) of the IS50 elements, and the target DNA was necessary for transposition (Goryshin & Reznikoff 1998). Enhancing the activity of the Tn5 system is possible by mutation at three specific sites, LP372, MA56, and EK54. The LP372 mutation enhances Tnp activity, MA56 blocks synthesis of Inh and EK54 enhances the binding activity of the 19 bp OE of the IS50 elements (Goryshin & Reznikoff 1998).

1.4 Methods for Studying Epigenetic Mechanisms

1.4.1 Assays for studying DNA methylation

The classical and gold standard method for studying 5mC are bisulfite conversion followed by cloning and sequencing. The general idea with this technique is to deaminate all regular cytosine to uracil while leaving 5mC intact (Frommer et al. 1992).

1.4.2 Assays for histone PTMs

Specific histone PTMs can be globally detected with western blot (WB) enzyme-linked immunosorbent assay (ELISA) analysis, while immunohistochemistry (IHC) and immunofluorescence (IF) can provide higher resolution (Liang et al. 2018; Tamassia et al. 2018; Zhang et al. 2004). If the goal is to study the enrichment of specific histone marks on a specific genomic location, Chromatin Immunoprecipitation (ChIP) can be used (Lindeman et al. 2009).

1.4.2.1 Western Blot

Western Blot is used for global assessment of a protein in a sample. It is an immunological method using antibodies to bind specific epitopes. The proteins are separated and denaturated with Sodium Dodecyl Sulfate (SDS) - polyacrylamide gel electrophoresis. After the electrophoresis, the proteins are blotted onto a filter paper. The proteins of interest can then be bound to a specific primary antibody. A secondary antibody is used to visualize the proteins of interest by for instance fluorescence. The amount of proteins is quantified by measuring the signal intensity.

1.4.2.2 Chromatin Immunoprecipitation

The early studies of the localization of RNA polymerase II across the genome by making covalent bonds between DNA and a protein of interest (cross-linking) in a bacterial cell are considered to be the origins of present chromatin immunoprecipitation (ChIP) assays (Gilmour & Lis 1984; Gilmour & Lis 1985). The concept of ChIP is to identify a genomic position for a specific protein factor or DNA binding protein. The ChIP is a robust assay which has provided insight into molecular and epigenetic mechanisms for several decades.

1.4.3 Assays for measuring open chromatin

Epigenetic marks will together contribute to the modulation of chromatin and to investigate the cumulative effect, chromatin accessible assays have proven to be useful. Genome-wide chromatin accessibility can be measured with DNase hypersensitivity sequencing (DNase-seq) (Song & Crawford

2010), mapping of micrococcal nuclease activity coupled with sequencing (MNase-seq) (Pajoro et al. 2018) and assay for transpose-accessible chromatin (ATAC-seq) (Buenrostro et al. 2013). The three approaches all serve the same purpose in general, but DNase-seq and ATAC-seq are, because of a high signal to noise ratio, especially useful to investigate nucleosome positioning and transcription factor (TF) footprints in more detail. ATAC-seq is, arguably, a more elegant methodology than DNase-seq and MNase-seq for a couple of reasons. Firstly, the number of cells. DNase-seq and Mnase-seq require millions of cells, while ATAC-seq requires considerable less with only a few hundred. Secondly, ATAC-seq preparation is not time-consuming. The whole sample and library preparation can be done in just a few hours.

1.4.3.1 Assay for transposase-accessible chromatin

ATAC-seq utilizes a hyperactive Tn5 transposase such as the system depicted in "1.2.6 Development of hyperactive Tn5 transposase". The hyperactive transposase fragments DNA at accessible areas of chromatin with endonuclease activity, and simultaneously catalyze the ligation of 19 bp OE sequences to the fragmented DNA (Figure 9) (Buenrostro, J. et al. 2015; Buenrostro et al. 2013).

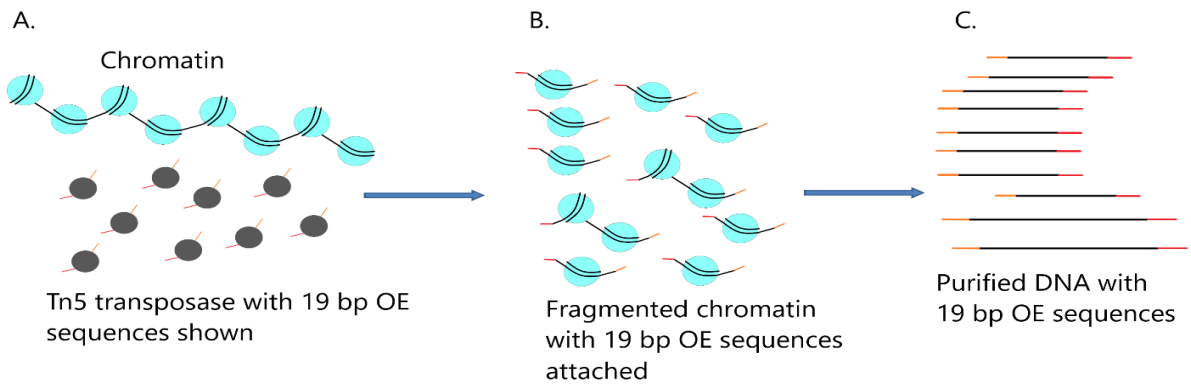

Figure 9: Schematic visualization of a tagmentation reaction with hyperactive Tn5 transposase, including purification. **A.** Non-fragmented chromatin as well as Tn5 transposases with 19 bp OE sequences (red and orange) attached. **B.** Fragmented chromatin, tagmented with 19 bp OE sequences. **C.** Purified tagmented DNA ready for library preparation.

The ATAC library is produced with PCR primers complementary to the 19 bp OE sequences, containing Illumina sequencing adapters and index sequences. The library preparation leads to a distinct pattern of fragments, separated by the size of a nucleosome. The pattern can be visualized by, e.g., gel-electrophoresis or an electropherogram (Figure 10). High-throughput sequencing is then utilized to analyze the nucleosomal pattern of the library.

On zebrafish, ATAC-seq has been used on the 256-cell stage and dome stage embryos (Chen et al. 2017), though without technical or biological replicates. On 24 hpf embryos, Bogdanovic et al. (2016) reported widespread DNA demethylation of enhancers during the phylotypic period while Quillien et al. (2017) used ATAC-seq for the identification of active tissue-specific enhancer elements. Detailed method protocols are available for small number zebrafish embryos (Doganli et al. 2017), pools of embryos (Fernandez-Minan et al. 2016), and for erythrocytes (Yang et al. 2016).

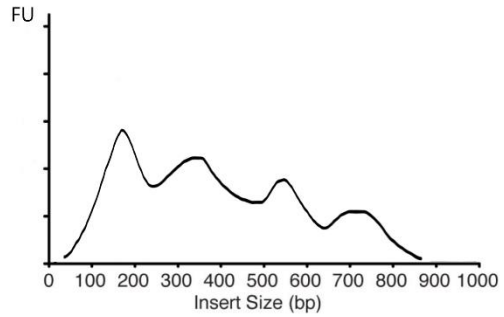

Figure 10: Fragment size distribution. An indication of a proper library preparation is the visualization of a pattern representing the separation of nucleosome-sized fragments. Note: Visualized fragments include 19 bp OE sequences and PCR primers.

1.5 Polycomb Group Proteins

Polycomb- and Trithorax group proteins (PcG and TrxG) are highly conserved from *Drosophila* to humans and were first described in *Drosophila* in the late 1970s for the involvement of epigenetic control of the hox genes (Ingham 1998; Lewis 1978), mainly involved in development (Bhatlekar et al. 2014). The two groups of epigenetic controllers are believed to regulate transcription by changing chromatin accessibility (Mahmoudi & Verrijzer 2001). PcG proteins are associated with gene silencing and cellular differentiation via transcriptional repression often mediated through trimethylation of H3K27 (Grossniklaus & Paro 2014; Tiwari et al. 2008), while TrxG proteins are in general associated with gene activation through, e.g., H3K4me3 (Kingston & Tamkun 2014). PcG and TrxG proteins recruit the PcG- and TrxG regulatory elements (PREs and TREs), responsible for the inheritance and maintenance of chromatin states throughout the development (Muller & Kassis 2006; Paro et al. 1998). Most of the PcG proteins divide into subcategories of repressive complexes, polycomb repressive complex 1 (PRC1) and polycomb repressive complex 2 (PRC2) (Grossniklaus & Paro 2014). Trimethylation of H3K27 is carried out by the catalytic subunit of PRC2, enhancer of zeste homolog 2 (Ezh2) (Grossniklaus & Paro 2014). PRC1 can potentially bind to H3K27me3 which may limit the access

of remodeling factors leading to lead to structural chromatin changes and transcriptional silencing (Cao & Zhang 2004; Grossniklaus & Paro 2014).

1.5.1 Inhibition of Ezh2 regulate malignity

The relationship between Ezh2 and cancer in humans was first discovered in the early 2000s when researchers found that the gene coding for Ezh2 was notably upregulated in metastatic prostate cancer (Varambally et al. 2002). Later research supports this relation showing genetic, transcriptional, and post-transcriptional dysregulation of Ezh2 in many cancer types (Yamagishi & Uchimaru 2017). Ezh2 has also lately been shown to alter gene expression of tumor suppressive genes, promoting tumorigenesis (Wassef et al. 2016), and Ezh2 initiates several molecular processes, characteristic for several cancer types (Yamagishi & Uchimaru 2017). Ezh2 encourages cancer research related to PRC2 and specifically the Ezh2 protein. The PRC2 is always dynamic regulated in all cellular processes during normal development, differentiation and reproduction making it hard to do a targeted study of Ezh2 (Yamagishi & Uchimaru 2017). Chromatin studies by analyzing the regulation of the chromatin associated with Ezh2 and developing ways to manipulate Ezh2 is, therefore, essential.

2 The aim of the study

The thesis is part of the Centre of Excellence (CoE) Centre for Environmental Radioactivity (CERAD) where effects of low-dose gamma-irradiation are elucidated, with the use of zebrafish as a model organism (Hurem et al. 2018; Kamstra et al. 2018). To understand these observed effects, CERAD wants to utilize open chromatin assays, ATAC-seq. As a controlled ATAC-seq study, ezh2 inhibition was chosen to obtain epigenetic effects detectable with ATAC-seq. The theory was an induced decline in H3K27me3 marks on chromatin, correlating with condensed chromatin and silent genes (Grossniklaus & Paro 2014).

We hypothesized genome-wide observable enrichment of accessible chromatin after treatment of PF-06726304 acetate, an Ezh2 inhibitor.

Initially, a published protocol on zebrafish embryo ATAC-seq (Doganli et al. 2017) was supposed to be utilized for this purpose. Unfortunately, due to shortcomings with this protocol, ATAC-seq was established at our laboratory with more robust technical solutions.

Subgoal for this project was,

1) Establish methods for manual dechorionation, a prerequisite for Doganli et al., (2017) method.
2) Establish more robust protocols for enzymatic dechorionation, deyolking and nuclei isolation.
3) Establish methods for library clean up.
4) Establish bioinformatic pipeline for initial analysis of ATAC-seq data.

3 Methods

3.1 Zebrafish husbandry

Embryos from the AB wild-type strain were obtained from the NMBU zebrafish facility and maintained according to standard operating procedures. Zebrafish were kept at 28±1 °C on a 14-10 hour light-dark cycle at a density of 5-10 fish/L. System water (SW) was prepared from a particle, and active charcoal filtrated tap water, deionized by reverse osmosis (RO) and kept sterile by UV irradiation. The water was conditioned to a conductivity of 500 µS/cm, general hardness (GH) 4-5 and pH 7.5 by addition of 155g synthetic sea salt (Instant Ocean, Blacksburg, USA), 53g sodium carbonate and 15g calcium chloride (Sigma-Aldrich) per liter RO water. Adult fish were fed with Gemma Micro 300 (Skretting, Stavanger, Norway) dry feed twice a day and live artemia (Scanbur, Karlslunde, Denmark) once a day.

3.1.1 Ethical considerations

All experiments were performed at the zebrafish facility of the Norwegian University of Life Sciences and according to Norwegian Animal Welfare Act (2009) and the EU Directive 2010/63, following appropriate guidelines.

3.2 Test compound

PF-06726304 acetate was obtained from Sigma (Sigma-Aldrich, prod nr. 1616287-82-1, purity ≥ 98%). The chemical was dissolved in DMSO to a stock solution of 10 mM (w/V) and immediately before treatment diluted in the culture medium.

3.3 Treatment of PF-06726304 acetate

Embryos were collected directly after fertilization and kept at 28±1 °C. Not viable embryos were discarded from the exposures. For ChIP-qPCR and Western Blot, 1 x 200 embryos were maintained in 9 mL of 5 µM PF-06726304 acetate and 1 x 200 embryos were kept in 9 ml control solution (4.5 µL DMSO). For ATAC-seq, 1 x 150 embryos were maintained in 9 mL of 5 µM PF-06726304 acetate and 1 x 150 embryos were maintained in 9 mL control solution. The exposure time for both experiments was from approximately 2-cell stage to 5.5 hpf.

3.4 Chromatin immunoprecipitation

3.4.1 Fixation of chromatin

For the fixation, 2 x 100 embryos from control exposure and 2 x 100 exposed embryos were transferred to and squeezed through to 21G needles and dissociated with 500 µL PBS. The needles were washed with 200 µL PBS. The volume was adjusted to 1 mL with PBS. The fixation was performed for 8 minutes after adding 27 µL formaldehyde (37%) to a final concentration of 1%. During the fixation, the samples were mixed a couple of times gently. The samples were incubated for 5 minutes, room temperature (RT), to stop the fixation, by adding 114 µL glycine (1.25 M) to a final concentration of 0.125 M. The samples were vortexed a couple of times during this step. Centrifugation at 500, 4 °C for 10 minutes was used to separate the chromatin from the lipids, followed by a wash with 1 mL ice-cold PBS. The samples were centrifuged again with the same specifications before all PBS was removed leaving a pellet containing intact cross-linked cells.

3.4.2 Fragmentation of chromatin

To lyse the cells, each pellet was resuspended in 100 µL lysis buffer (LB) (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% (wt/vol) SDS, 1% protease inhibitor (PI) (Sigma-Aldrich, cat. #P8340), 1% phenylmethanesulfonyl fluoride (PMSF)). The samples were then mixed until homogenous. For the fragmentation, the lysate was transferred to 0.6 ml Bioruptor tubes. The chromatin was then fragmented by sonication (Diagenode, Bioruptor), for 10 cycles (1 cycle: 30 seconds OFF, 30 seconds ON) at 4 °C. The tubes were centrifuged at 12000 × g, 4 °C for 10 minutes to separate high molecular weight, non-fragmented material, from the fragmented chromatin. 80 µL of the fragmented chromatin (the supernatant) was transferred into 4 x 1.5 mL tubes. The concentration (table 1) was measured on epoch (BioTek Instruments Inc.) and the samples were diluted by adding radioimmunoprecipitation assay (RIPA) buffer (10 mM Tris-HCl, pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% (vol/vol) Triton X-100, 0.1% (wt/vol) SDS, 0.1% (wt/vol), Na-deoxycholate, 1% PI, 1% PMSF) to a final concentration of 10 ng/µL chromatin (table 1). The samples were stored at -80°C.

Table 1: Volume [µL], and concentration [ng/µL] measured with epoch, for the two controls (C1 and C2) and the two 5 µM PF-06726304 acetate exposed (E1 and E2) samples. As well as the volume of RIPA added to each sample to obtain 10 ng/µL and a total volume of 1000 µL.

Sample	Concentration [ng/µL]	Volume [µL]	Volume RIPA added [µL]
Control 1	122	82	918
Control 2	353	28	972
5 µM PF-06726304 acetate 1	157	64	936
5 µM PF-06726304 acetate 2	211	47	953

3.4.3 Immunoprecipitation of H3K4me3, H3K27me3 and H3 pan

Duplicate immunoprecipitations of H3K4me3, H3K27me3, and H3 pan was performed on PF-06726304 acetate exposed chromatin as well as control chromatin. Duplicate controls without antibody were added to the setup, to consider unspecific binding. As a reference, 2 input controls were added for both exposed- as well as control chromatin. Before the immunoprecipitation, a stock solution of magnetic dynabeads (Invitrogen, Dynabeads Protein A, 30 mg/mL, 10002D) in RIPA was prepared. 180 µL of the dynabeads solution was transferred to a 1.5 ml tube and put on a pre-cooled magnetic rack to bind the beads. The original solution was replaced by 500 µL ice-cold RIPA and vortexed to wash the beads. The wash was performed 3 times, before adding RIPA to the original volume (180 µL). 10 µL of the dynabeads stock solution was then aliquoted with 90 µL RIPA. The stock solution was continuously vortexed during this procedure to maintain a homogenous solution. 2 µg of antibodies for H3K4me3 (Diagenode, C15410003) and H3K27me3 (Diagenode, C15410195), as well as 2 µL of antibody for H3 pan (Diagenode, C15310135), was added to the tubes. The reactions were incubated on a rotator at 40 rpm, 8 °C for 2 hours. After the incubation, 100 µL PF-06726304 acetate exposed and control chromatin was added to the respective reactions, before incubation on a rotator at 40 rpm, 8 °C, overnight (ON). After the incubation, the reactions were washed three times with ice-cold RIPA by using a magnetic rack to capture the beads. The beads were resuspended in 100 µL ice-cold TE (pH 8.0). The samples, including the dynabeads, were transferred to new 200 µL tubes. The original tubes were washed with 50 µL TE to collect potential remaining beads, resulting in a total volume of 150 µL bead solution in the new tubes.

3.4.4 Elution of chromatin from dynabeads

Before the DNA-isolation 3.5 mL, fresh complete elution buffer (CEB) (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% (wt/vol) SDS and 20 µg/mL proteinase K) was prepared. 150 µL EB was added to each solution and incubated on Thermomixer (Eppendorf, Thermomixer C) at 1200 rpm, 68 °C for 3 hours to eluate the chromatin-antibody complexes from the beads and for cross-linking reversal.

The beads were then captured using a magnetic rack and the supernatant, containing the immunoprecipitated chromatin, was transferred to new 1.5 mL tubes. The beads were resuspended in 150 µL CEB and incubated on Thermomixer at 1200 rpm, 68 °C for 15 minutes to eluate possibly remaining chromatin. The beads were captured, and the supernatant was pooled with the respective 1.5 mL tubes. 200 µL EB (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50 mM NaCl) was added to the samples, resulting in a total volume of 500 µL. 2 x 100 µL PF-06726304 acetate exposed and 2 x 100 µL control chromatin was added to 1.5 mL tubes to prepare input controls. 400 µL EB and 2 µL protease K (50 µg/mL) were added and incubated at 1200 rpm, 68 °C for 3 hours. The samples were maintained at 4 °C until DNA extraction.

3.4.5 DNA extraction

To extract DNA from the samples, 500 µL Phenol:Chloroform:Isoamyl alcohol (25:24:1; Invitrogen, cat. #15593-031), equilibrated in RT, was added to the samples before thorough vortexing. The samples were then centrifuged at 15000 × g, RT, for 5 minutes. 460 µL of the "water-phase" containing the DNA was then transferred to new 1.5 mL tubes. To rid the samples of phenol, 460 µL Chloroform:Isoamyl alcohol (24:1; Sigma-Aldrich, cat. #C0549), were added before thorough vortexing. The samples were then centrifuged at 15000 × g, RT, for 5 minutes. 400 µL of the "water-phase" was then transferred to new 1.5 mL tubes. 40 µL 3M Sodium acetate (pH 7.0) was added along with 10 µL linear acryl-amide carrier (Ambion, cat. #AM9520) as co-precipitant and 1 mL ethanol (96%) to precipitate the DNA. The samples were maintained at -80 °C until frozen solid, before thawing and centrifugation at 20000 × g, 4 °C for 20 minutes. The supernatant was removed leaving approximately 50 µL with the pellet. The samples were washed with 1 mL ice-cold ethanol (70%), before centrifugation with 20000 × g, 4 °C for 10 minutes. The wash was repeated once, before drying the pellets open-lid to remove the ethanol. The dry pellets were dissolved ON in 100 µL TE (pH 7.5).

3.4.6 Analysis with quantitative Polymerase Chain Reaction

Real-time PCR was performed on *βactin2* and *CCAAT/enhancer-binding protein α* (*c/ebpα*) on specific loci (table 4), using 2 µL template DNA and 10 µL qPCR master mix (1X FastStart Essential DNA Green Master (Roche, cat. #25595200), and 20 µM forward and reverse primer (Supplementary Table 3)). The temperature regime for the thermocycling was (1) 95 °C for 5 minutes, (2) 95 °C for 30 seconds, (3) 60 °C for 30 seconds, (4) 72 °C for 30 seconds. Step 2 - 4 was repeated 40 times. Amount of precipitated DNA relative to input was calculated. Relative precipitated H3K4me3 and H3K27me3 was visualized using the relative precipitated H3 pan as base-line.

3.4.7 Statistical Analysis

2-way Analysis of Variance (ANOVA) was performed with a 95% confidence interval, by comparing each treatment mean, for each dataset. The statistics were visualized with a standard error of the mean (SEM).

3.5 Open Chromatin Assay

3.5.1 Dechorionation with pronase treatment

The chorion was removed to avoid crude samples that would interfere with the experiment. The embryos from both conditions were incubated with 300 ng/mL pronase, RT, for 10 minutes. The dechorionated embryos were washed with system water. 3 x 30 intact embryos for each condition were collected and transferred to six 1.5 mL tubes using a P200 pipette.

3.5.2 Deyolking with Ginzburg deyolk buffer

The yolk was dissociated with 500 µL Ginzburg fish ringer deyolk buffer (55 mM NaCl, 1.8 mM KCl, 1.25 mM $NaHCO_3$). The embryos were incubated with shaking at 1100 rpm, RT, for 5 minutes before spinning down the cells at 500 × g, 4 °C for 5 minutes. The supernatant was removed, and the cell pellet was washed with 500 µL ice-cold PBS followed by a 500 × g centrifugation, at 4 °C, for 10 minutes.

3.5.3 Isolation of nuclei

The cell membrane were lysed with 50 µL lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Igepal CA630). The lysate was centrifuged at 500 × g, 4 °C, for 10 minutes. The supernatant was removed carefully, without touching the nuclei pellet.

3.5.4 Fragmentation and tagmentation

A 2X DNA tagmentation buffer (22 mM Tris-HCl pH 7.4, 10 mM MgCl2, 20% Dimethylformamide) was prepared. The isolated nuclei were then re-suspended in 50 µL transposition reaction mixture (25 µL DNA tagmentation buffer pH 7.5, 1.25 µL Tn5 (Illumina, 15027865), 23.75 µL H_2O). The samples were incubated on Thermomixer at 300 rpm, 37 °C for 30 minutes. The tagmented DNA was then purified with MinElute PCR Purification Kit (QIAGEN, cat. #28004) as following the manufacturers standard protocol. In short, the tagmented DNA was re-suspended in 250 µL buffer PB, pH adjusted with Na-acetate, mixed and transferred to spin columns (QIAGEN, cat. #1026476). All following centrifugations were performed at 17900 × g, RT, for 1 minute. The columns were centrifuged, and the flow-through

was discarded. 750 µl buffer PE containing ethanol was added to the column before new centrifugation. The flow-through was discarded, and the columns were centrifuged once more to rid the columns for ethanol. The columns were placed in new 1.5 µL tubes and eluted in 10 µL elution buffer (QIAGEN, cat. #28004) was carefully added to the center of the columns to ensure proper elution. The samples were maintained on ice until library preparation.

3.5.5 Library preparation

The Nextera Index Kit (Illumina, FC-121-1030) was used for the library preparation. The exact indexes used are included in table 3. Real-time PCR was performed using 1 µL of the purified transposed DNA and 9 µL qPCR master mix (1X NEBNext High-Fidelity PCR Master Mix (New England Labs, cat. #M0541), 1X SybrGreen, 0.25 nM indexed primers) to determine the number of PCR cycles needed for the library preparation. The temperature regime for the thermocycling was (1) 72 °C for 5 minutes, (2) 98 °C for 30 seconds, (3) 98 °C for 10 seconds, (4) 63 °C for 30 seconds, and (5) 72 °C for 30 seconds. Step 3 - 5 was repeated 20 times. The samples were maintained at 4 °C after qPCR. The preferred number of amplification cycles for the libraries was decided to be between 9 and 10 cycles. The cycle number was calculated by adding 2 cycles to the ct value (table 2).

Table 2: Ct values, as well as the bases in the adapters for the common and specific indices, used for the three controls (C1, C2, and C3) and the 5 µM PF-06726304 acetate exposed (E1, E2, E3) samples.

Sample	Ct value	Common index	Specific index
Control 1	7.98	GCGTAAGA (N517)	TCGCCTTA (N701)
Control 2	8.07	GCGTAAGA (N517)	CTAGTACG (N702)
Control 3	7.56	GCGTAAGA (N517)	TTCTGCCT (N703)
5 µM PF-06726304 acetate	7.80	GCGTAAGA (N517)	GCTCAGGA (N704)
5 µM PF-06726304 acetate	7.72	GCGTAAGA (N517)	AGGAGTCC (N705)
5 µM PF-06726304 acetate	7.40	GCGTAAGA (N517)	CATGCCTA (N706)

The remaining 9 µL of purified transposed DNA was PCR amplified with 41 µL PCR master mix (1X NEB Next High-Fidelity PCR Mastermix (New England Labs, cat #M0541), 5 µM common and specific index (table 3)). The temperature regime for the thermocycling was (1) 72 °C for 5 minutes, (2) 98 °C for 30 seconds, (3) 98 °C for 10 seconds, (4) 63 °C for 30 seconds, and (5) 72 °C for 60 seconds. Step 3-5 was repeated 9 times. The samples were maintained at 4 °C after the PCR. Finally, the amplified library was purified using a PCR Purification Kit (QIAGEN, cat. #28004) with the same procedure as above (see 3.5.5 Fragmentation and tagmentation), and eluted in 20 µL of elution buffer.

3.5.6 Visualization on high sensitivity bioanalyzer

For the bioanalyzer, a gel-dye mix was prepared by mixing 25 µL dye-concentrate with the gel matrix vial. The solution was applied to a spin filter (provided in the kit) and spun down at 2300 × g, RT, for 15 minutes. The gel-dye mix was kept in the dark. For the analysis, the microfluidic chip was first prepared using a priming station. The base plate was positioned at position C, the syringe clip at the lowest position and the plunger was at 1 mL before starting. 9 µL of the gel-dye mix was added to the well marked Ⓖ on the chip, followed by closing the priming station and pressing the plunger down until it was held by the syringe clip. This position was held for exactly one minute before the plunger was released. After five seconds it was carefully set back to the 1 mL position before the priming station was opened. 9 µL of the gel-dye mix was further added to the next wells marked G. 5 µL of the marker was added every sample well, including the ladder well. 1 µL ladder was added to the ladder well, and 1 µL of each library was added to each respective well. The chip was vortexed at 2400 rpm, RT, for one minute using an IKA vortex mixer. The wells were visually inspected to have no bubbles before the analysis. The analysis was performed using Agilent 2100 Expert software and a high-sensitivity bioanalyzer assay.

3.5.7 AMPure beads removal of primers

The sample (20 µL) was mixed with 28 µL AMPure XP beads solution (Beckman Coulter, cat. #A63880) and incubated, RT, for 10 minutes. The beads would bind fragment >100 bp to the beads. The tubes were placed on a magnetic rack to capture the beads. The smaller non-bound fragments were removed with a volume of 5 µL was left in the tubes. The beads were washed with 200 µL freshly made 80% (vol/vol) ethanol by rotating the tubes 180° to cause the beads to flow through the solution. The rotating was performed eight times to ensure proper washing before removing the supernatant. The wash was repeated one more time. The tubes were spun down to collect remaining ethanol and placed back on the magnetic rack. The remaining ethanol was removed and the tubes incubated, RT, for 5 minutes to dry the beads. The tubes were removed from the magnetic rack to elute >100 bp DNA with 40 µL EB (Qiagen, cat. #28004), RT, for 10 minutes. The beads were again captured, and the supernatant containing the library was transferred to new 1.5 mL tubes.

3.5.8 Library clean up on XE e-gel

The ATAC libraries were cleaned up to avoid additional peaks containing possible adapter fragments. The complete libraries were separated on 2% agarose gel (Xe, dry gel system, Invitrogen). The gel cassette was opened, and the agarose was shortly exposed to blue light. During blue light excitation, the 100 bp and 15 kb localizations were marked with a scalpel. The gel was removed from the blue

light source, and the DNA libraries of interest were extruded from the gel into pre-weighed 1.5 mL tubes. The gel was documented with pictures before and after library clean up. The DNA was extracted from the gel matrix using Invitrogen PureLink Quick Gel Extraction Kit (Invitrogen, cat. #K2100). In short, the gel fragment was solubilized in an average volume of 300 uL Gel Solubilization Buffer (L3) corresponding to 3:1 of the weight of the gel piece. The samples were incubated on Thermomixer at 50°C, for 20 minutes and inverted several times until complete solubilization was observed. The DNA was bound on a Quick Gel Extraction Column at 12000 × g, for 1 minute, washed with Wash Buffer (W1) and eluted in a Recovery Tube in a volume of 50 µL. The DNA concentration was measured with Qubit High Sensitivity assay. Due to low DNA concentration, the DNA was up concentrated using Qiagen MinElute system, eluting samples in 15 µL elution buffer.

3.5.9 Sequencing

Paired-end NGS with a read length of 150 bases and a read depth of 50 million raw reads was outsourced to Novogene, Hong Kong, China.

3.5.10 Raw Data Processing and Statistical Analysis

Fastq files were processed with GUAVA (A **GU**I tool for the **A**nalysis and **V**isualization of **A**TAC-seq data) version 1 (Divate, Mayur & Cheung, Edwin 2018), a software manager with a graphical user interface to perform a complete standardized ATAC-seq bioinformatic analysis from raw data to results. GUAVA managed the following steps: Adapter trimming (minimum read length = 30, error rate = 0.1), alignment (Bowtie2, maximum insert size = 2000, No. of genomic hits (m) = 1, genome = GRCz10), Filtering of mitochondrial- as well as unknown reads, MACS2 peak calling (q = 0.05) and alignment correction as the Tn5 transposase induces and uneven cut of 9 bp (Buenrostro et al. 2013). Comparative analysis was performed on the processed samples with the following parameters, DESeq2, log2(FC) = 1.5, p = 0.05. Peaks within 5000 bp upstream and 3000 bp downstream of transcriptional start site (TSS) was associated with the respective gene. ATAC-seq dataset published by Chen et al. (2015) was processed through the same GUAVA pipeline. The processed datasets were visualized and compared with Interactive Genomic Viewer (IGV) (Thorvaldsdóttir et al. 2013).

3.6 Western Blot

3.6.1 Dechorionation and deyolking

At 5.5 hpf embryos were dechorionated as described. The dechorionated embryos were washed with system water. 3 x 30 intact embryos for each treatment were collected and transferred to six 1.5 mL tubes using a P200 pipette. The embryos were then deyolked as described in «3.5.2 Deyolking with Ginzburg deyolk buffer». The supernatant was removed, and the cell pellet was washed with 500 µL ice-cold PBS followed by centrifugation at 500 × g, 4 °C, for 10 minutes. The cells were immediately stored in -80 °C.

3.6.2 Isolation and sonication of nuclei

The frozen cell pellet was lysed in 30 µL lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 0.25% DOC, 1% Igepal CA630, 1x PI) followed by centrifugation at 3500 × g, 4 °C, for 5 minutes to separate nuclei from cytosol. The cytosol fraction was removed, and the nuclei were sonicated (Diagenode, Bioruptor), for 10 cycles (1 cycle: 30 seconds OFF, 30 seconds ON) at 4 °C.

3.6.3 Denaturation

The samples were denatured with 1X NuPAGE™ LDS Sample Buffer (Thermo Fisher Scientific, cat. #NP0007) and 1X NuPAGE™ Sample Reducing Agent (Thermo Fisher Scientific, cat. #NP0004), at 90 °C, for 5 minutes, followed by a quick vortex and 13000 rpm centrifugation (Heraeus, Biofuge Pico), RT, for 5 minutes.

3.6.4 Gel-electrophoresis

20 µL of the samples were applied to the gel (Bio-Rad, cat. #3450119). 2 µL size benchmark was applied on each side of the samples. Gel-electrophoresis ran with 200 V for 1 hour. The gel electrophoresis was performed with 1X XT MOPS running buffer (Bio-Rad, cat. #161-0788).

3.6.5 Blotting

An Immunoblot PVDF membrane (Bio-Rad, cat. #162-0255) was first soaked in methanol for 2 minutes. The methanol was poured off and the membrane rinsed with H_2O. The PVDF membrane was then soaked in 1X Tris/CAPS Buffer for Semi-Dry Blotting (Bio-Rad, cat. #161-0778) for 5 minutes. A couple of filter paper was also soaked in anode/cathode buffer respectively (Bio-Rad, cat. #161-0778). The proteins were blot from the gel to the PVDF membrane on a transfer cell (Bio-Rad, Trans-Blot SD Semi-Dry Transfer Cell). The filter papers from the anode buffer were placed on the anode of the transfer

cell with the PVDF membrane on top trying to avoid air bubbles. The gel was then placed on top of the PVDF membrane, and the filter paper from the cathode buffer was placed at the very top of this sandwich. The transfer cell cathode was attached at the lid was put on. The blotting ran with 25 V for 1 hour. After the blotting, the sandwich was dismantled and the filter blotted with the proteins was quickly transferred to a container with a blocking solution (5% dry milk (Bio-Rad, cat. #170-6404) in 1X TBST and incubated, RT, for 60 minutes. After blocking, the filter was split into three parts, one for each antibody.

3.6.6 Addition of primary and secondary antibodies

Each filter was transferred to a 50 mL Falcon tube, with the proteins facing the middle, and incubated with 2 mL of primary antibody binding solution (1% Blotting-Grade Blocker (Bio-Rad, cat. #170-6404)) on a roller mixer (Stuart, Roller Mixer, SRT9D), 8 °C, ON. The proteins were incubated with 5.8 µg of antibodies for H3K27me3 (Diagenode, C15410195), 5.6 µg H3K27ac (Diagenode, C15410196), and 2 µg H4 (Abcam, ab10158) was added. The filters were washed with a substantial amount of 1X TBST for 3 x 15 seconds, following 3 x 15 minutes. The secondary antibody was bound by incubating the filters with 2 mL secondary antibody binding solution (1% Blotting-Grade Blocker), RT, for 90 minutes. The proteins were incubated 3 µg of antibodies goat anti-rabbit (Novex, cat. #A24534) was used. The filters were washed as before.

3.6.7 ECF printing and Thypoon Scanning

ECF substrate was evenly spread out on a glass plate corresponding to the area of the filters. The filters were incubated on the ECF substrate, with the proteins facing down, RT, for 5 minutes before drying the filters completely. The filters were placed in a Typhoon scanner (Amersham Biosciences, Typhoon 9200) with the proteins down, covered by a glass plate. The filters were scanned using the Typhoon Scanner Control 3.0 software with the following settings, Acquisition Mode: Fluorescence, Emission filter: Fluorescence, PMT: 500, Laser: Green (532), Sensitivity: Normal, Pixel size: 200.

3.6.8 Quantification of Typhoon Scanning

The typhoon scan was quantified with ImageJ by selecting the bands of interest and measuring the density of each band relative to the background.

3.6.9 Statistical Analysis

2way Analysis of Variance (ANOVA), was performed with a 95% confidence interval, by comparing each treatment mean, for each dataset. The statistics were visualized with a standard error of the mean (SEM).

4 Results

4.1 Treatment of PF-06726304 acetate reveal epigenetic effects

4.1.1 Visualized fragment size distribution

High-throughput sequencing determines the fragment size distribution as visualized below (Figure 11). The distinct pattern reveals mono-, di-, and tri-nucleosomal fragments without adapters included. The distribution also reveals a peak smaller than one nucleosome (peak at approximately 100 bp). The distribution is compared to published fragment size distribution from Buenrostro et al. (2015).

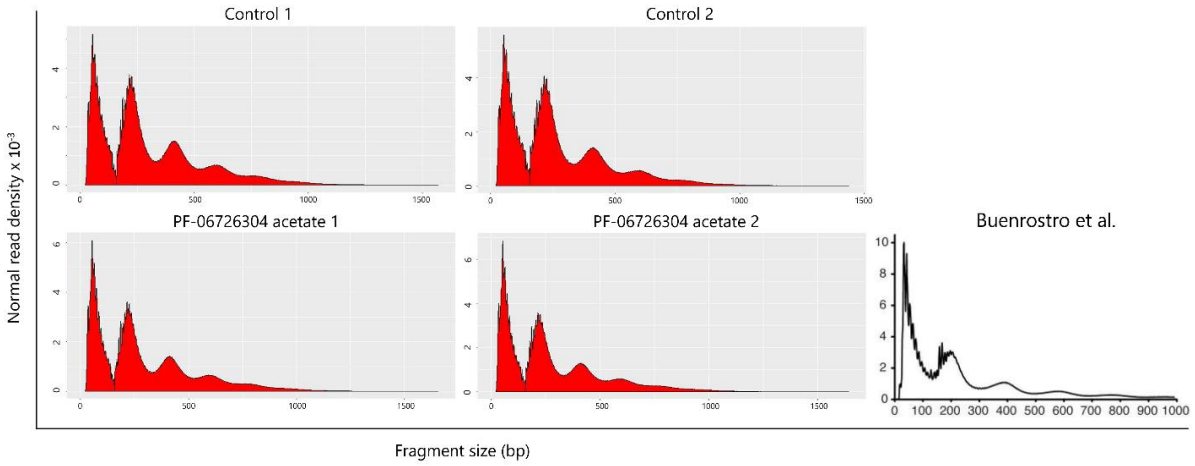

Figure 11: Fragment size distribution determined by high-throughput sequencing showing read density (y-axis) of the library fragment sizes (x-axis), excluding adapters. As a comparison, we show the size distribution of Buenrostro et al. (2015), which is the developer of ATAC.

4.1.2 ATAC-seq differential analysis reveals changes to chromatin

Principle Component Analysis reveals 72% variance between controls and treated samples, PC1 (x-axis) as well as 16% variance within the replicates, PC2 (y-axis) (Figure 12A) after treatment of 5 µM PF-06726304 acetate. The GUAVA pipeline reveals 63 "gained-close" and 32 "gained-open" regions (Figure 12B). Out of the differentially annotated accessibility peaks, 60% are located in promoter regions, 20% immediate downstream, and 20% in intergenic regions (Figure 12C). Genes associated with differential change of > log2(FC)=1.45 is listed in supplementary table 4.

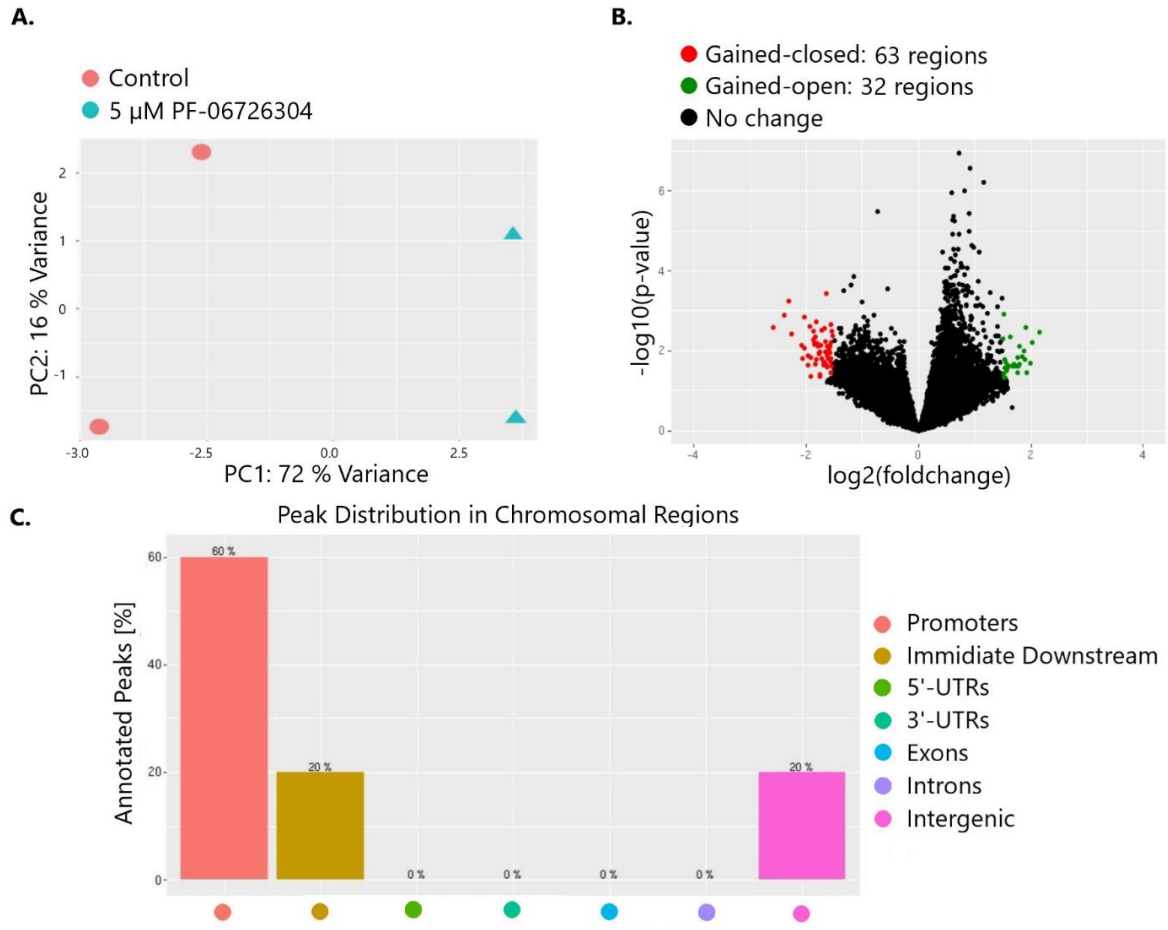

Figure 12: Differential analysis output from the GUAVA software. **A.** Principle Component Analysis. Illustration of the variance between controls and treated samples. Revealing 72% variance between controls and treated samples, PC1 (x-axis), as well as 16% variance within the replicates, PC2 (y-axis). **B.** Volcano plot. Log2(FC)=1.5. The plot reveals 63 gained-closed regions (red) and 32 gained-open regions (green). The majority is regions with no change (black). **C.** Peak distribution of differentially enriched ATAC-seq peaks, in the promoter, immediately downstream, 5'-UTR, 3'-UTR, exons, introns, and intergenic regions.

4.1.3 Manual interogation of GUAVA differential analysis

The examples of associated genes listet in supplementary table 4 were interrogated closer with IGV. 2179 bp upstream for *homeobox B1b* (*hoxb1b*) (p = 0.043), and 718 bp upstream of *interleukin enhancer binding factor 2* (*ilf2*) (p = 0.016) are gained-open (Figure 13), while 4222 bp upstream for *aspartoacylase* (*aspa)* (p = 0.003), at TSS of *solute carrier family 12 (sodium/chloride transporter), member 3* (*slc12a3*) (p = 0.013), and 4858 bp upstream for *methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1b* (*mthfd1b*) (p = 0.015) are gained-closed (Figure 14).

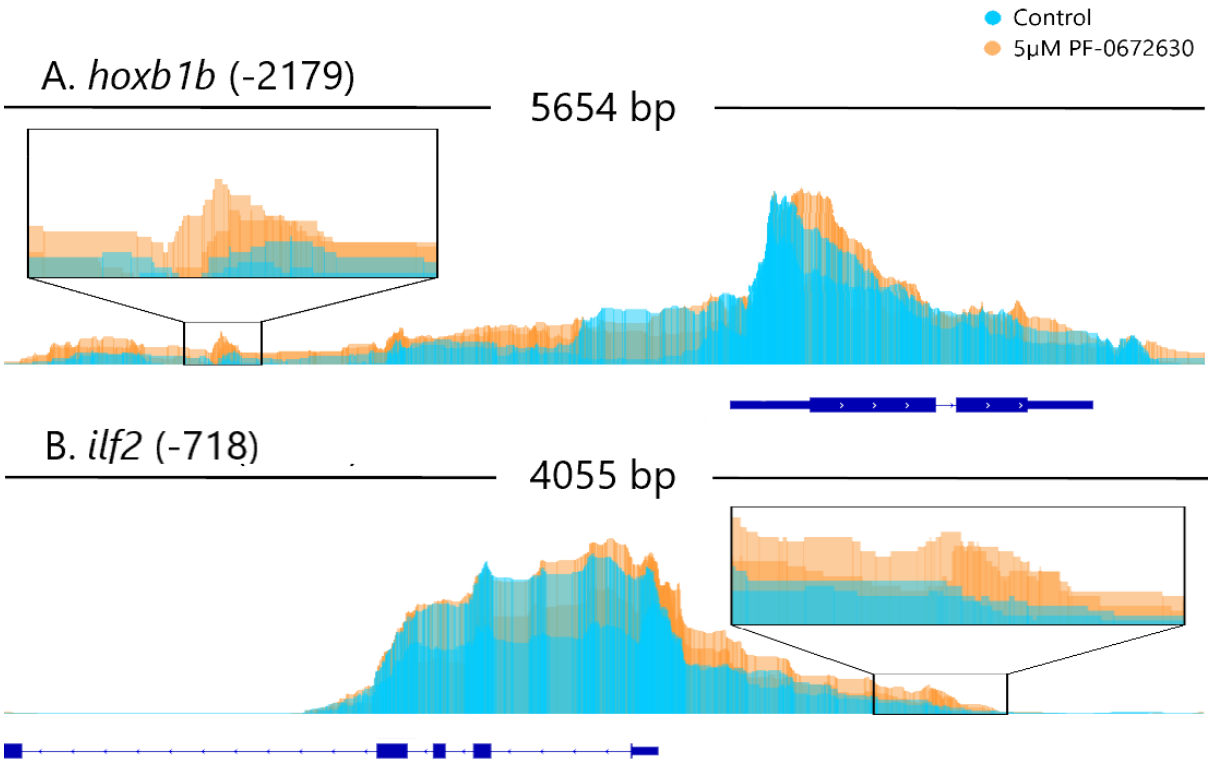

Figure 13: Visualization of specific chromatin regions with IGV, revealing gained-open regions. **A.** 2179 bp upstream *hoxb1b* (p = 0.043). **B.** 718 bp upstream of *ilf2* (p = 0.016). Every profile is visualized with the same y-axis.

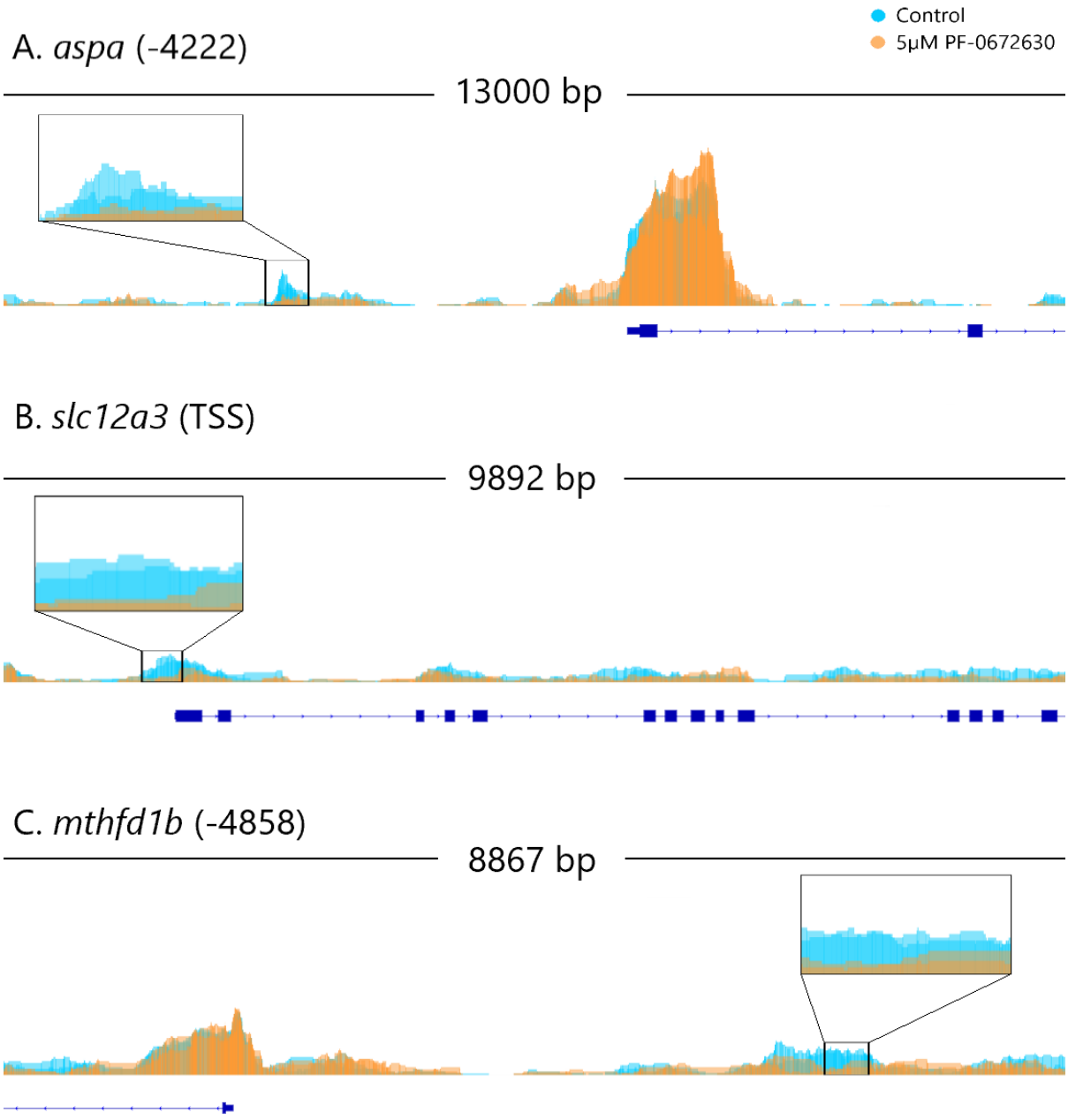

Figure 14: Visualization of specific chromatin regions with IGV revealing gained-closed regions. **A.** 4222 bp upstream for *aspa* (p = 0.003), **B.** TSS of *slc12a3* (p = 0.013), and **C.** 4858 bp upstream for *mthfd1b* (p = 0.015). Every profile are visualized with the same y-axis.

4.1.4 Comparison with published ATAC-seq datasets

Published ATAC-seq dataset from dome stage embryos by Chen et al. (2017) was processed through the GUAVA version 1 pipeline and compared with our controls and exposed samples. Profiles of two loci with accessible chromatin (*βactin2*, *c/ebpα*) and one profile indicating condensed chromatin *hepatocyte nuclear factor 4, α* (*hnf4α*) are visualized in Figure 15.

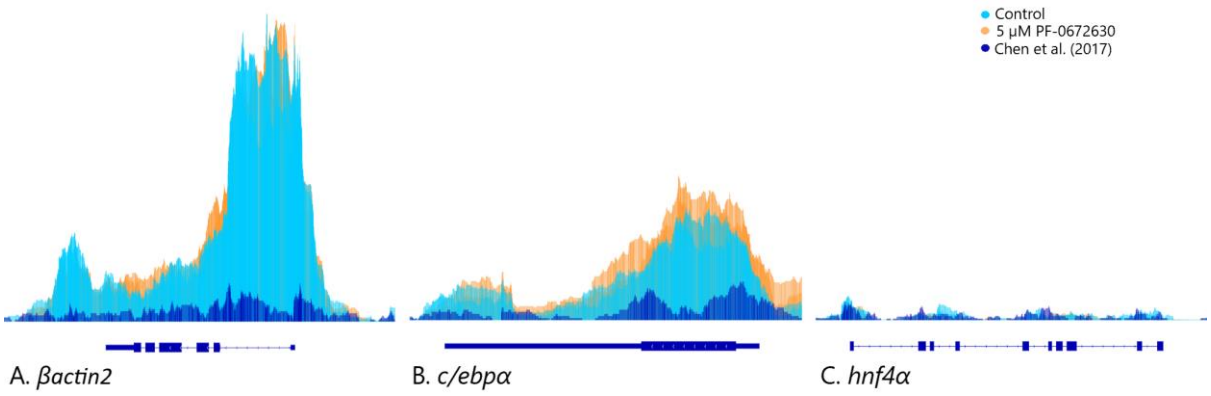

Figure 15: Visual comparison between Chen et al. (2017) ATAC-seq on dome stage zebrafish embryos, and this thesis' control and treated samples at 5.5 hpf embryos. **A.** *βactin2*, **B.** *c/ebpα*, and **C.** *hnf4α*. The y-axis has the same scale for every profile.

4.2 Validation of PF-06726304 acetate treatment

4.2.1 Western blot

After treatment with 5 µM PF-06726304 acetate, western blot analysis of H3K27ac and H3K27me3 reveal a global rise in H3K27ac, while no significant change in H3K27me3 (Figure 16).

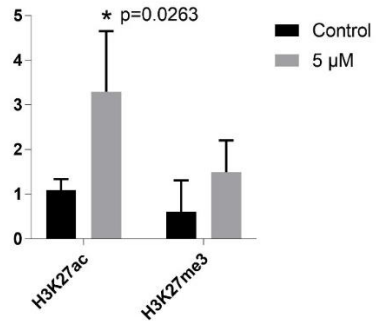

Figure 16: Western Blot relative to the background of H3K27ac and H3K27me3 after treatment with 5 µM PF-06726304 acetate, normalized to H4. Revealing the global rise of H3K27ac, but no significant change in H3K27me3.

4.2.2 Chromatin Immunoprecipitation

After treatment with 5 µM PF-06726304 acetate, H3K4me3 ChIP reveal locus-specific decrease on *βactin2* (gene body) and no significant change on *c/ebpα* (upstream, TSS and gene body). Immunoprecipitation of H3K27me3 reveals no significant change on *βactin2* and *c/ebpα* (upstream), while a significant decrease on *c/ebpα* (TSS and gene body) (Figure 17).

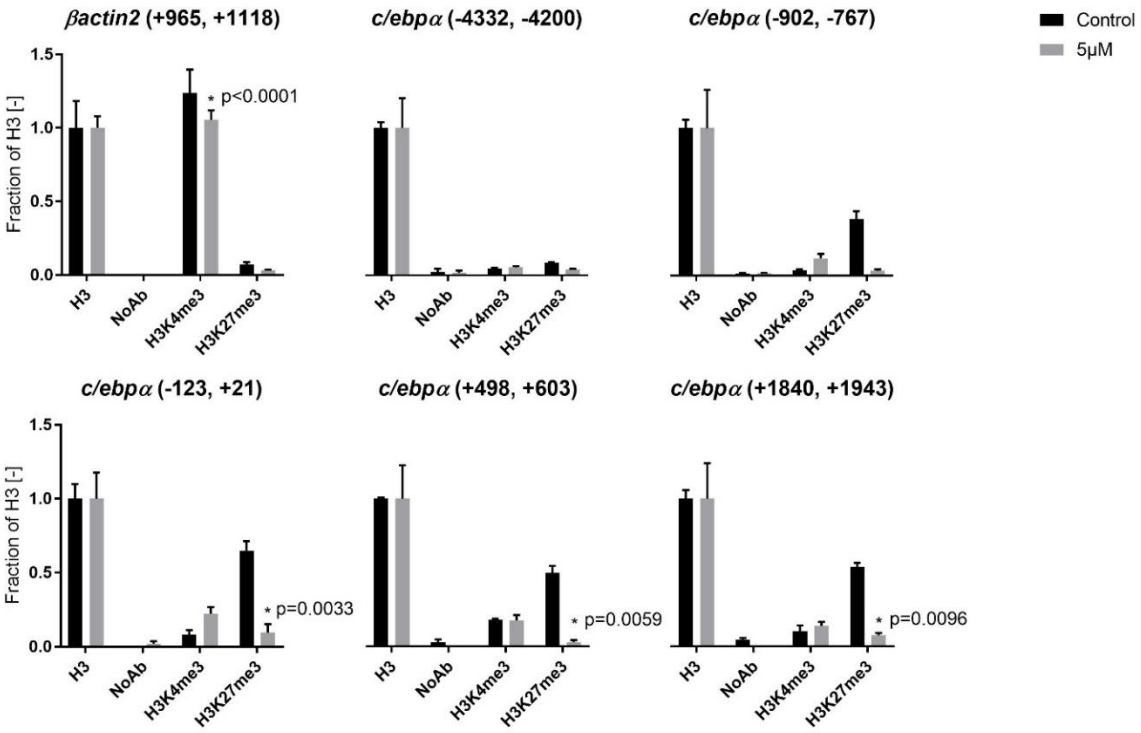

Figure 17: Chromatin Immunoprecipitation of H3K4me3 and H3K27me3 on zebrafish embryos (5.5 hpf), untreated versus exposed to PF-06726304 acetate. The enrichment is calculated as a percent of input and normalized to histone H3 enrichment. Numbers indicate amplicon location relative to TSS.

5 Discussion

The goal of this project was to visualize chromatin accessibility after low-dose treatment of a novel Ezh2 inhibitor, PF-0672630 acetate, in zebrafish from zygote to early gastrula stage. An Ezh2 inhibition causes downregulation of H3K27me3, an epigenetic repressive mark (Wiles & Selker 2017). The hypothesis was that enrichment of accessible chromatin would be demonstrated by ATAC-seq after treatment with the inhibitor PF-06726304 acetate. In a broader perspective, this thesis is part of introducing ATAC-seq as a routine assay for research using zebrafish and other model species for radiation biology effects on the epigenome, within the Centre of Excellence Centre for Environmental Radioactivity (CoE CERAD.

5.1 Main Findings

5.1.1 Interpretation of GUAVA differential analysis

The hypothesized result after PF-0672630 treatment was the genome-wide gain of open chromatin regions due to the inhibition of H3K27me3. The PCA plot (Figure 12A) shows that there is variance (76%) between the treated samples and the controls. At the same time, the variance within the replicates is 13%. The controls and treated samples were sequenced on separate lanes which could lead to lane-to-lane differences. The lane-to-lane differences were checked with a new differential analysis comparing replicates of control-treated with control-treated. The new differential analysis revealed no hits because of the independence, confirming that the variance between and within the replicates shown in figure 12A is acceptable. The volcano plot (Figure 12B) revealed 63 gained-open and 32 gained-closed regions with a cut-off at p=0.05 and log2(FC)=1.5, and figure 12C shows that there indeed are differences in promotor regions after treatment.

A closer look at the GUAVA differential analysis reveals changes in chromatin regions associated with *hoxb1b, ilf2*, *mthfd1b*, *aspa*, and *slc12a3* (Supplementary Table 4). Gained-open chromatin is evident around 2179 bp upstream for *hoxb1b*, and 718 bp upstream of *ilf2*. In general, the hox genes are highly conserved regulators of embryonic development, and *hoxb1b* is associated with the organization of nucleosomes (Weicksel et al. 2014). The increase of accessible chromatin in these regions may be explained by the increase of H3K27ac due to a decrease in H3K27me3 (Pasini et al. 2010). Interestingly, treatment of PF-0672630 also revealed significantly gained-closed regions around 4222 bp upstream for *aspa*, 4858 bp upstream for *mthfd1b*, as well as *slc12a3 TSS* which is especially interesting. The gained-closed regions were not expected from our initial hypothesis, but it is not really a contradiction given the complexity of the epigenome.

5.1.2 Biological interpretation of the differential analysis

Manual interpretation of the dataset in IGV points to larger areas with gained-open chromatin. GUAVA only deemed parts of these regions as significant (Figure 12 and 13) due to either the observable variation within the replicates or the fact that DESeq2 is not optimized for ATAC-seq datasets (discussed in paragraph 5.2.1).

A comparison with previously published datasets by Chen et al. (2017), shows the same tendencies with enriched profiles for *βactin2*, *c/ebpα* (Figure 15) and a repressed profile for *hnf4α* although our dataset shows a higher signal to noise ratio for the open chromatin profiles.

5.2 ATAC-seq methodology

The ATAC-seq protocol is described to be performed in 3h with a few laboratory steps (Buenrostro, J. et al. 2015). The protocol, describing ATAC-seq on low number zebrafish embryos (Doganli et al. 2017) was used in the initial experiments of this thesis, but in our hands to the production of proper libraries proved to be more difficult than initially thought. Interestingly, a highly similar protocol exists for Xenopus embryos (Bright & Veenstra 2018), but no published dataset is available for either. Based on the protocol of (Fernandez-Minan et al. 2016), we revised it at several steps: dechorionation, deyolking, and isolation of nuclei. The embryos were dechorionated with pronase and thoroughly rinsed after the treatment to avoid effects from traces of pronase on the results. The yolk contains a substantial amount of material that could affect the library preparation negatively, and removal of the yolk was added to the protocol to improve the library preparation further. The amount of nuclei relative to the amount of Tn5 transposase used during a tagmentation of chromatin is essential to be able to produce proper libraries of nucleosomal fragments (Buenrostro, J. et al. 2015). Counting the nuclei after lysis is an excellent way to be sure there is no loss of material in the first steps, and that the nuclei are adequately isolated.

The goal is to achiece a complex library of fragments, ranging from 200 to 1000 bp revealing enrichment of mono-, di-, and tri-nucleosomal lengths. It is essential to produce an unbiased library, more or less evenly amplified over the different fragment sizes. Roughly, the library preparation should not be amplified for more than 11 - 12 cycles of PCR, in total (Buenrostro, J. et al. 2015; Doganli et al. 2017). The fragment size distribution determined by high-throughput sequencing reveal the expected mono-, di-, and tri-nucleosomal pattern (Figure 11). The distribution also reveals smaller fragments around 100 bp (Figure 11). The smaller fragment is hypothesized internucleosomal

fragments from linker DNA (Divate, M. & Cheung, E. 2018; Szerlong & Hansen 2011), still representing euchromatin.

The tagmentation procedure utilized in this thesis was performed at 37 °C for 30 minutes, optimized for human cells (Buenrostro, J. et al. 2015). Although the protocol worked seemingly well, since zebrafish are kept at 28 °C rather than 37 °C, there might be potential for further optimization of the method for zebrafish.

5.2.1 GUAVA

GUAVA is an easy to use program that can process ATAC-seq datasets from adapter trimming to a finished result file. It has, however, some limitations. For example, each sample in this project was aligned with the zebrafish GRCz10 genome as this was the newest zebrafish genome available in GUAVA, although the further refined GRCz11 assembly is available. For the differential analysis method, the only option is the DESeq2 algorithm (Love et al. 2014). DESeq2 serves its purpose and gives a preliminary result, but it is not necessarily the best option for ATAC-seq data since the algorithm is initially developed for differential transcriptomic analysis. As for log2(FC), GUAVA only offers the choice between 1, 1.5 or 2.0. In situations where log2(FC)=1 is too low and log2(FC)=1.5 too high, it would be beneficial to have more fine tuning setting possibilities.

Although the experimental design and the exposures and sequencing were carried out in triplicates, shortcomings of the GUAVA software however caused a reduction in power of the bioinformatic analysis to be performed as duplicates of control and treated samples. For the comparative analysis differentially changed regions of chromatin within 5000 bp upstream and 3000 bp downstream of a gene was associated with the respective gene. The basis of this choice was previously performed ATAC-seq experiments by Bogdanović et al. (2016).

5.3 Validation of PF-0672630 treatment

5.3.1 Global effects

Western blot analysis revealed a global rise in H3K27ac after treatment of 5 µM PF-0672630. The global rise in H3K27ac is logic after exposure with a histone demethylase inhibitor as the epitope now would be accessible for histone acetyltransferases (Wiles & Selker 2017). In theory, the H3K27me3 levels should decrease as a result of the rise in H3K27ac (Wiles & Selker 2017), but the western blot data revealed no significant switch between the H3K27 marks.

5.3.2 Locus-specific effects on *βactin2* and *c/ebpα*

The result from ChIP-PCR revealed a locus-specific decrease ($p < 0.0001$) in H3K4me3 on *βactin2* (Figure 17). *βactin2* is a housekeeping gene often used as a normalization control for gene expression with a substantially high enrichment of activating marks compared to other loci. The result revealed enrichment in H3K4me3 to the same level as pan H3 on the *βactin2* gene body. Immunoprecipitation of *c/ebpα* assessed to validate that the 5 µM PF-0672630 used in the treatment induced a decrease of H3K27me3 enrichment in all loci investigated related to *c/ebpα*, both in putative regulatory sequences and in the gene body. Immunoprecipitation of H3K4me3 revealed no significant effects at the *c/ebpα* loci. The GUAVA differential analysis of the ATAC-seq dataset revealed no significant change in chromatin accessibility on the *c/ebpα* locus, but manual interpretation with IGV points to a gained-open accessible profile for one of the treated replicates (Figure 15) and would be interesting to investigate further with more replicates.

6 Concluding remarks

This thesis shows that low-dose treatment of zebrafish embryos with PF-06726304 acetate induces specific effects on chromatin accessibility. The analysis should be repeated with the triplicate datasets to produce event clearer results with increased statistical power. The thesis demonstrates that GUAVA can be used as a primary analysis tool for ATAC-seq datasets, but that more in-depth bioinformatic analysis is necessary to generate a more mechanistic understanding of how Ezh2 interferes with the epigenetic control of gene expression. Still, GUAVA produced a quick differential analysis with a few genes mapped to regions with changes in chromatin accessibility, *hoxb1b, ilf2*, *hes2.1*, *mthfd1b*, *aspa*, and *slc12a3,* that could be interesting for further analysis.

6.1 Further work

This thesis has introduced ATAC-seq as an assay for research in zebrafish embryo epigenetics, and have now also been performed in gamma irradiation experiments for zebrafish to study effects on the epigenome (personal communication). As a priority, further work would be to establish a more thorough bioinformatic analysis pipeline for ATAC-seq to be, e.g., able to map datasets to the GRCz11 zebrafish genome assembly. Next is to correlate ATAC datasets to relevant transcriptomes, and eventually methylomes. Finally, the next step will probably be single-cell ATAC-seq (Buenrostro, J. D. et al. 2015) with corresponding related epigenomics.

7 Supplementary material

Table 3: PCR primers used for ChIP analysis. Gene, forward and reverse primer with start and end position relative to TSS, as well as annealing temperature.

Gene	Start	End	Forward primer	Reverse primer	Temp. (°C)
βactin2	+965	+1118	ACTATGAACTGAACCGACTG	CTGCGATCAATTACACAACC	60
c/ebpα	-4332	-4200	TCTCTGGCGCAACTTCCAAT	TCTGGGTCGAACAAATGGGT	60
c/ebpα	-902	-767	CATTCTTTAGGTGCCAGGCC	CCCCATAGTGCGAGAAAAGC	60
c/ebpα	-123	+21	TAGGTCTATCAGTGCGTCCG	ACTTGCAACCTCAGTGTGTG	60
c/ebpα	+498	+603	GATGTATGGCTGCCTGAACG	TCCCGAGGCTCTTGTTTGAT	60
c/ebpα	+1840	+1943	GCAGTGAAGTCCTGTCTTGC	CACTGAAACCATCTCTGCTCG	60

Table 4: Differentially changed regions associated with genes, the location of the region, log2(FC), p-value, and type of change.

Gene	Location	log2(FC)	p-value	Type of change
hoxb1b	Upstream	1,45	0,043	gained-open
slc12a3	TSS	-1,96	0,013	gained-close
ilf2	Upstream	1,89	0,016	gained-open
aspa	Upstream	-1,87	0,003	gained-close
mthfd1b	Upstream	-1,56	0,015	gained-close

8 References

Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K. & Walter, P. (2015). *Molecular Biology of the Cell*. 6th ed.: Garland Science.

Aleström, P. & Winther-Larsen, H. C. (2016). 7 - Zebrafish offer aquaculture research their services. In MacKenzie, S. & Jentoft, S. (eds) *Genomics in Aquaculture*, pp. 165-194. San Diego: Academic Press.

Altucci, L. & Rots, M. G. (2016). Epigenetic drugs: from chemistry via biology to medicine and back. *Clinical Epigenetics*, 8: 56.

Alvarenga, E. M., Rodrigues, V. L., Moraes, A. S., Naves, L. S., Mondin, M., Felisbino, M. B. & Mello, M. L. (2016). Histone epigenetic marks in heterochromatin and euchromatin of the Chagas' disease vector, Triatoma infestans. *Acta Histochem*, 118 (4): 401-12.

Ambros, V. (2004). The functions of animal microRNAs. *Nature*, 431 (7006): 350-5.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116 (2): 281-97.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, 447 (7143): 407-12.

Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. (2009). An operational definition of epigenetics. *Genes & Development*, 23 (7): 781-783.

Bhatlekar, S., Fields, J. Z. & Boman, B. M. (2014). HOX genes and their role in the development of human cancers. *J Mol Med (Berl)*, 92 (8): 811-23.

Bodi, Z., Button, J. D., Grierson, D. & Fray, R. G. (2010). Yeast targets for mRNA methylation. *Nucleic Acids Res*, 38 (16): 5327-35.

Bogdanović, O., Smits, A. H., de la Calle Mustienes, E., Tena, J. J., Ford, E., Williams, R., Senanayake, U., Schultz, M. D., Hontelez, S., van Kruijsbergen, I., et al. (2016). Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nature genetics*, 48 (4): 417-426.

Bright, A. R. & Veenstra, G. J. C. (2018). Assay for Transposase-Accessible Chromatin-Sequencing Using Xenopus Embryos. *Cold Spring Harb Protoc*.

Buenrostro, J., Wu, B., Chang, H. & Greenleaf, W. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol*, 109: 21.29.1-9.

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*, 10 (12): 1213-8.

Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., Chang, H. Y. & Greenleaf, W. J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, 523: 486.

Busam, K. J., Pulitzer, M. P., Coit, D. C., Arcila, M., Leng, D., Jungbluth, A. A. & Wiesner, T. (2017). Reduced H3K27me3 expression in Merkel cell polyoma virus-positive tumors. *Modern Pathology*, 30: 877.

Cao, R. & Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Current Opinion in Genetics & Development*, 14 (2): 155-164.

Castel, S. E. & Martienssen, R. A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature reviews. Genetics*, 14 (2): 100-112.

Cavalieri, V. & Spinelli, G. (2017). Environmental epigenetics in zebrafish. *Epigenetics & chromatin*, 10 (1): 46-46.

Chang, P., Gohain, M., Yen, M.-R. & Chen, P.-Y. (2018). Computational Methods for Assessing Chromatin Hierarchy. *Computational and Structural Biotechnology Journal*, 16: 43-53.

Chen, Y., Zeng, S., Hu, R., Wang, X., Huang, W., Liu, J., Wang, L., Liu, G., Cao, Y. & Zhang, Y. (2017). Using local chromatin structure to improve CRISPR/Cas9 efficiency in zebrafish. *PLoS ONE*, 12 (8): e0182528.

Crick, F. H. (1958). On protein synthesis. *Symp Soc Exp Biol*, 12: 138-63.

Divate, M. & Cheung, E. (2018). GUAVA: A Graphical User Interface for the Analysis and Visualization of ATAC-seq Data. *Front Genet*, 9: 250.

Divate, M. & Cheung, E. (2018). GUAVA: A Graphical User Interface for the Analysis and Visualization of ATAC-seq Data. *Frontiers in genetics*, 9: 250-250.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S. & Ren, B. (2012). Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature*, 485 (7398): 376-380.

Doganli, C., Sandoval, M., Thomas, S. & Hart, D. (2017). Assay for Transposase-Accessible Chromatin with High-Throughput Sequencing (ATAC-Seq) Protocol for Zebrafish Embryos. *Methods Mol Biol*, 1507: 59-66.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, 485 (7397): 201-6.

Eccleston, A., Cesari, F. & Skipper, M. (2013). Transcription and epigenetics. *Nature*, 502: 461.

Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. *Oncogene*, 21: 5400.

Feil, R. & Fraga, M. F. (2012). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, 13: 97.

Fernandez-Minan, A., Bessa, J., Tena, J. J. & Gomez-Skarmeta, J. L. (2016). Assay for transposase-accessible chromatin and circularized chromosome conformation capture, two methods to explore the regulatory landscapes of genes in zebrafish. *Methods Cell Biol*, 135: 413-30.

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L. & Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A*, 89 (5): 1827-31.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N. & Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell reports*, 15 (9): 2038-2049.

Genome Reference Consortium. (2017). *Zebrafish Genome Overview*. Available at: https://www.ncbi.nlm.nih.gov/grc/zebrafish (accessed: 5/8).

Gilmour, D. S. & Lis, J. T. (1984). Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proc Natl Acad Sci U S A*, 81 (14): 4275-9.

Gilmour, D. S. & Lis, J. T. (1985). In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. *Mol Cell Biol*, 5 (8): 2009-18.

Giraldez, A. J., Cinalli, R. M., Glasner, M. E., Enright, A. J., Thomson, J. M., Baskerville, S., Hammond, S. M., Bartel, D. P. & Schier, A. F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. *Science*, 308 (5723): 833-8.

Goryshin, I. Y. & Reznikoff, W. S. (1998). Tn5 in vitro transposition. *J Biol Chem*, 273 (13): 7367-74.

Grossniklaus, U. & Paro, R. (2014). Transcriptional Silencing by Polycomb-Group Proteins. *Cold Spring Harbor Perspectives in Biology*, 6 (11): a019331.

Gutierrez-Lovera, C., Vazquez-Rios, A. J., Guerra-Varela, J., Sanchez, L. & de la Fuente, M. (2017). The Potential of Zebrafish as a Model Organism for Improving the Translation of Genetic Anticancer Nanomedicines. *Genes (Basel)*, 8 (12): 349.

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, 14 (10): 3156.

Hurem, S., Martin, L. M., Lindeman, L., Brede, D. A., Salbu, B., Lyche, J. L., Alestrom, P. & Kamstra, J. H. (2018). Parental exposure to gamma radiation causes progressively altered transcriptomes linked to adverse effects in zebrafish offspring. *Environ Pollut*, 234: 855-863.

Ingham, P. W. (1998). trithorax and the regulation of homeotic gene expression in Drosophila: a historical perspective. *Int J Dev Biol*, 42 (3): 423-9.

James D. Watson, T. A. B., Stephen P. Bell, Alexander Gann, Michael Levine, Richard Losick. (2008). *Molecular Biology of the Gene*. Sixth ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Jenuwein, T. & Allis, C. D. (2001). Translating the Histone Code. *Science*, 293 (5532): 1074.

Kamstra, J. H., Hurem, S., Martin, L. M., Lindeman, L. C., Legler, J., Oughton, D., Salbu, B., Brede, D. A., Lyche, J. L. & Aleström, P. (2018). Ionizing radiation induces transgenerational effects of DNA methylation in zebrafish. *Scientific reports*, 8 (1): 15373-15373.

Karsli-Ceppioglu, S., Dagdemir, A., Judes, G., Lebert, A., Penault-Llorca, F., Bignon, Y.-J. & Bernard-Gallon, D. (2017). The Epigenetic Landscape of Promoter Genome-wide Analysis in Breast Cancer. *Scientific Reports*, 7 (1): 6597.

Khalil, A. M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B. E., van Oudenaarden, A., et al. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A*, 106 (28): 11667-72.

Kimmel, C. B., Sessions, S. K. & Kimmel, R. J. (1981). Morphogenesis and synaptogenesis of the zebrafish Mauthner neuron. *J Comp Neurol*, 198 (1): 101-20.

Kimmel, C. B., Sepich, D. S. & Trevarrow, B. (1988). Development of segmentation in zebrafish. *Development*, 104 Suppl: 197-207.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev Dyn*, 203 (3): 253-310.

Kingston, R. E. & Tamkun, J. W. (2014). Transcriptional Regulation by Trithorax-Group Proteins. *Cold Spring Harbor Perspectives in Biology*, 6 (10): a019349.

Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science*, 184 (4139): 868-71.

Le Dily, F., Baù, D., Pohl, A., Vicent, G. P., Serra, F., Soronellas, D., Castellano, G., Wright, R. H. G., Ballare, C., Filion, G., et al. (2014). Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes & Development*, 28 (19): 2151-2162.

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature*, 276 (5688): 565-70.

Li, G. & Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. *Current opinion in genetics & development*, 21 (2): 175-186.

Li, X., Egervari, G., Wang, Y., Berger, S. L. & Lu, Z. (2018). Regulation of chromatin and gene expression by metabolic enzymes and metabolites. *Nature Reviews Molecular Cell Biology*.

Liang, Z. L., Wu, D. D., Yao, Y., Yu, F. Y., Yang, L., Tan, H. W., Hylkema, M. N., Rots, M. G., Xu, Y. M. & Lau, A. T. Y. (2018). Epiproteome profiling of cadmium-transformed human bronchial epithelial cells by quantitative histone post-translational modification-enzyme-linked immunosorbent assay. *J Appl Toxicol*, 38 (6): 888-895.

Lindeman, L. C., Vogt-Kielland, L. T., Alestrom, P. & Collas, P. (2009). Fish'n ChIPs: chromatin immunoprecipitation in the zebrafish embryo. *Methods Mol Biol*, 567: 75-86.

Liu, N. & Pan, T. (2015). RNA epigenetics. *Transl Res*, 165 (1): 28-35.

Love, M. I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 15 (12): 550.

MacDonald, W. A. (2012). Epigenetic Mechanisms of Genomic Imprinting: Common Themes in the Regulation of Imprinted Regions in Mammals, Plants, and Insects. *Genetics Research International*, 2012: 585024.

Mahmoudi, T. & Verrijzer, C. P. (2001). Chromatin silencing and activation by Polycomb and trithorax group proteins. *Oncogene*, 20 (24): 3055-66.

Meyer, K. D. & Jaffrey, S. R. (2016). Expanding the diversity of DNA base modifications with N(6)-methyldeoxyadenosine. *Genome Biology*, 17: 5.

Moss, B., Gershowitz, A., Stringer, J. R., Holland, L. E. & Wagner, E. K. (1977). 5'-Terminal and internal methylated nucleosides in herpes simplex virus type 1 mRNA. *J Virol*, 23 (2): 234-9.

Muller, J. & Kassis, J. A. (2006). Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Curr Opin Genet Dev*, 16 (5): 476-84.

Nichols, J. L. (1979). N6-methyladenosine in maize poly(A)-containing RNA. *Plant Science Letters*, 15 (4): 357-361.

Ntziachristos, P., Tsirigos, A., Vlierberghe, P. V., Nedjic, J., Trimarchi, T., Flaherty, M. S., Ferres-Marco, D., da Ros, V., Tang, Z., Siegle, J., et al. (2012). Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nature Medicine*, 18: 298.

Pajoro, A., Muiño, J. M., Angenent, G. C. & Kaufmann, K. (2018). Profiling Nucleosome Occupancy by MNase-seq: Experimental Protocol and Computational Analysis. In Bemer, M. & Baroux, C. (eds) *Plant Chromatin Dynamics: Methods and Protocols*, pp. 167-181. New York, NY: Springer New York.

Paro, R., Strutt, H. & Cavalli, G. (1998). Heritable chromatin states induced by the Polycomb and trithorax group genes. *Novartis Found Symp*, 214: 51-61; discussion 61-6, 104-13.

Pasini, D., Malatesta, M., Jung, H. R., Walfridsson, J., Willer, A., Olsson, L., Skotte, J., Wutz, A., Porse, B., Jensen, O. N., et al. (2010). Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic acids research*, 38 (15): 4958-4969.

Payer, B., Lee, J. T. & Namekawa, S. H. (2011). X-inactivation and X-reactivation: epigenetic hallmarks of mammalian reproduction and pluripotent stem cells. *Human genetics*, 130 (2): 265-280.

Peer, E., Rechavi, G. & Dominissini, D. (2017). Epitranscriptomics: regulation of mRNA metabolism through modifications. *Curr Opin Chem Biol*, 41: 93-98.

Pombo, A. & Dillon, N. (2015). Three-dimensional genome architecture: players and mechanisms. *Nature Reviews Molecular Cell Biology*, 16: 245.

Pourakbar, S., Pluard, T. J., Accurso, A. D. & Farassati, F. (2017). Ezh2, a novel target in detection and therapy of breast cancer. *OncoTargets and therapy*, 10: 2685-2687.

Quillien, A., Abdalla, M., Yu, J., Ou, J., Zhu, L. J. & Lawson, N. D. (2017). Robust Identification of Developmentally Active Endothelial Enhancers in Zebrafish Using FANS-Assisted ATAC-Seq. *Cell Rep*, 20 (3): 709-720.

Rao, Suhas S. P., Huntley, Miriam H., Durand, Neva C., Stamenova, Elena K., Bochkov, Ivan D., Robinson, James T., Sanborn, Adrian L., Machol, I., Omer, Arina D., Lander, Eric S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*, 159 (7): 1665-1680.

Rodenhiser, D. & Mann, M. (2006). Epigenetics and human disease: translating basic biology into clinical applications. *Cmaj*, 174 (3): 341-8.

Saksouk, N., Simboeck, E. & Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin*, 8: 3.

Schaefer, I.-M., Fletcher, C. D. M. & Hornick, J. L. (2015). Loss of H3K27 trimethylation distinguishes malignant peripheral nerve sheath tumors from histologic mimics. *Modern Pathology*, 29: 4.

Scharf, A. N. & Imhof, A. (2011). Every methyl counts--epigenetic calculus. *FEBS Lett*, 585 (13): 2001-7.

Song, L. & Crawford, G. E. (2010). DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor protocols*, 2010 (2): pdb.prot5384-pdb.prot5384.

Szerlong, H. J. & Hansen, J. C. (2011). Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 89 (1): 24-34.

Tamassia, N., Arruda-Silva, F., Calzetti, F., Lonardi, S., Gasperini, S., Gardiman, E., Bianchetto-Aguilera, F., Gatta, L. B., Girolomoni, G., Mantovani, A., et al. (2018). A Reappraisal on the Potential Ability of Human Neutrophils to Express and Produce IL-17 Family Members In Vitro: Failure to Reproducibly Detect It. *Front Immunol*, 9: 795.

The RNA Institute. (2018). *The RNA Modification Database*. College of Arts and Sciences, State University of New York at Albany. Available at: http://mods.rna.albany.edu/mods/.

Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*, 14 (2): 178-192.

Tiwari, V. K., McGarvey, K. M., Licchesi, J. D. F., Ohm, J. E., Herman, J. G., Schübeler, D. & Baylin, S. B. (2008). PcG Proteins, DNA Methylation, and Gene Repression by Chromatin Looping. *PLoS Biology*, 6 (12): e306.

Trojer, P. & Reinberg, D. (2007). Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Molecular Cell*, 28 (1): 1-13.

Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., Shi, Y., Segal, E. & Chang, H. Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science (New York, N.Y.)*, 329 (5992): 689-693.

Tsompana, M. & Buck, M. J. (2014). Chromatin accessibility: a window into the genome. *Epigenetics Chromatin*, 7.

Turner, B. M. (2007). Defining an epigenetic code. *Nat Cell Biol*, 9 (1): 2-6.

Varambally, S., Dhanasekaran, S. M., Zhou, M., Barrette, T. R., Kumar-Sinha, C., Sanda, M. G., Ghosh, D., Pienta, K. J., Sewalt, R. G., Otte, A. P., et al. (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature*, 419 (6907): 624-9.

Waddington, C. H. (1942). The epigenotype. 1942. *Int J Epidemiol*, 41 (1): 10-3.

Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G. & Klein, A. M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science*, 360 (6392): 981-987.

Wassef, M., Michaud, A. & Margueron, R. (2016). Association between EZH2 expression, silencing of tumor suppressors and disease outcome in solid tumors. *Cell Cycle*, 15 (17): 2256-2262.

Weicksel, S. E., Gupta, A., Zannino, D. A., Wolfe, S. A. & Sagerstrom, C. G. (2014). Targeted germ line disruptions reveal general and species-specific roles for paralog group 1 hox genes in zebrafish. *BMC Dev Biol*, 14: 25.

Wiles, E. T. & Selker, E. U. (2017). H3K27 methylation: a promiscuous repressive chromatin mark. *Current opinion in genetics & development*, 43: 31-37.

Wu, J., Xu, J., Liu, B., Yao, G., Wang, P., Lin, Z., Huang, B., Wang, X., Li, T., Shi, S., et al. (2018). Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature*, 557 (7704): 256-260.

Wu, T. P., Wang, T., Seetin, M. G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S. D., Mackintosh, S. G., Zhong, M., et al. (2016). DNA methylation on N6-adenine in mammalian embryonic stem cells. *Nature*, 532: 329.

Yamagishi, M. & Uchimaru, K. (2017). Targeting EZH2 in cancer therapy. *Curr Opin Oncol*, 29 (5): 375-381.

Yang, S., Ott, C. J., Rossmann, M. P., Superdock, M., Zon, L. I. & Zhou, Y. (2016). Chromatin immunoprecipitation and an open chromatin assay in zebrafish erythrocytes. *Methods Cell Biol*, 135: 387-412.

Yoo, K. H. & Hennighausen, L. (2012). EZH2 methyltransferase and H3K27 methylation in breast cancer. *Int J Biol Sci*, 8 (1): 59-65.

Yu, X. & Li, Z. (2015). Long non-coding RNA HOTAIR: A novel oncogene (Review). *Mol Med Rep*, 12 (4): 5611-8.

Zhang, K., Siino, J. S., Jones, P. R., Yau, P. M. & Bradbury, E. M. (2004). A mass spectrometric "Western blot" to evaluate the correlations between histone methylation and histone acetylation. *Proteomics*, 4 (12): 3765-75.

Zhu, J., Zhu, W. & Wu, W. (2018). MicroRNAs Change the Landscape of Cancer Resistance. *Methods Mol Biol*, 1699: 83-89.

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway