The novel function of CFP-1 in *Caenorhabditis elegans* development

Bharat Pokhrel

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Molecular and Cellular Biology

August 2019

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Part of the experimental work described in Chapter 3, Chapter 4 and Chapter 6 of the thesis are included in the publication below and is directly attributable to the candidate.

Pokhrel, B., Chen, Y. & Biro, J. J. (2019) CFP-1 interacts with HDAC1/2 complexes in *C. elegans* development. FEBS J, 286: 2490-2504. doi:10.1111/febs.14833.

Candidate performed most of the work, devised the layout and wrote the manuscript. Yannic Chen performed the RNAi sensitivity assay and Jonathan Biro performed the heat shock reporter assay.

Acknowledgements

I am grateful to the University of Leeds for funding my PhD. I would like to thank Dr. Ron Chen and Dr. Patricija van Oosten-Hawle for the guidance and constant support throughout the project. I would also like to thank Dr. Andrew Macdonald for constructive feedback during my PhD and for comments and suggestions for the thesis.

I am grateful to Dr. Amit Kumar for all the encouragements, constant support and for critically reading my thesis. My solemn thanks to my friends and colleagues Jonathan Biro, Yannic Chen and Dovile Milonaityte for critically reading my thesis chapters and being great friends. My sincere thanks to Dr Laura Jones for guidance during my final year and, Alhanouf Alhabri and James Cain for the technical support and for being wonderful lab mates and friends. I would also like to thank Laura Warwick, Hann Ng and Rosamund Clifford for making my working environment enjoyable. I am grateful to Dr. Rupesh Paudyal and Dr. Gurmeet Kaur for motivating and supporting this journey and for comments on my thesis. My sincere thanks to my friend Ajinkya Sunil Rao who always supported me throughout this journey.

I am extremely grateful to my father (Ishori Prasad Pokhrel), mother (Kamala Pokhrel), brother (Saroj Pokhrel), sister in law (Prativa Acharya), daughter (Aayusha Pokhrel) and all the family members for providing me with continued support and encouragement throughout my life. Finally, I am highly grateful to my wife Tulaja Dahal for all the love, motivation, encouragement and support. Without her support, this would have never been possible.

Abstract

CxxC finger protein 1 (Cfp1) is an evolutionally conserved CpG binding protein that binds to unmethylated CpG-rich promoters. This conserved epigenetic regulator is a part of the complex of proteins associated with Set1 (COMPASS) complex that contains Set1 in mammals and SET-2 in *C. elegans* as a histone 3 lysine 4 trimethylation (H3K4me3) methyltransferase. The canonical function of Cfp1 is to link the COMPASS complex to CpG island promoters to subsequently deposit H3K4me3 marks at the 5′ end of genes. Previous studies have indicated the importance of Cfp1 in embryonic stem cell differentiation and cell fate specification. However, neither the function nor the mechanism of action of Cfp1 is well understood at the organismal level.

To further investigate the conserved function of Cfp1, I used *Caenorhabditis elegans cfp-1(tm6369)* and *set-2(bn129)* mutants. I observed that deletion of *cfp-1* or *set-2* results in a drastic reduction of H3K4me3 levels and the stronger expression of heat shock and salt-inducible genes. That suggests H3K4me3 could play a repressive role in gene induction. Surprisingly, I found that despite both genes being essential for H3K4me3 deposition and gene induction, only loss of CFP-1 (but not SET-2) function can suppress an organogenesis defect caused by aberrant epidermal development. I next carried out the small genetic screen to identify the genes that can enhance or suppress CFP-1 function in the epidermal development. Results indicated that CFP-1 could interact genetically with Histone deacetylases (HDACs) to regulate epidermal development.

To further investigate the genetic interaction between CFP-1 and HDACs, I conducted RNAi of HDACs on the *cfp-1(tm6369)* mutant and measured the effects on fertility. It was observed that CFP-1 but not SET-2 genetically interacts with HDAC1/2 complexes to promote fertility. In summary, this study suggests that apart from the function of CFP-1 in the H3K4me3 deposition, CFP-1 could interact with HDAC1/2 complexes in *C. elegans* development.

Acknowledgements ii

Abstract iii

List of figures x

List of tables xii

Introduction 1

1.1 Tightly controlled gene expression programme at the organismal level 1

1.1.1 Transcription 1

1.1.2 Transcription factors and gene expression 5

1.1.3 Role of TFs in development 6

1.1.4 Gene regulation and development in the context of chromatin 7

1.2 H3K4me3 13

1.3 Histone lysine methyltransferase (HMT) 14

1.3.1 Yeast Set1/COMPASS complex and its orthologues 17

1.3.2 Trx and its orthologues MLL1 and MLL2 19

1.3.3 Trithorax related: the orthologue of MLL3/MLL4 20

1.3.4 Deposition of H3K4me3 21

1.3.5 Recruitment of COMPASS complex on chromatin 21

1.4 Role of H3K4me3 in the regulation of gene expression 25

1.5 H3K4me3 in gene repression 26

1.6 H3K4me3 in genomic stability 27

1.7 Cfp1 28

1.7.1 Characterisation of Cfp1 29

1.7.2 Cfp1 and Set1 interaction 32

1.7.3 Cfp1 and cytosine methylation 32

1.7.4 Role of Cfp1 in development 33

1.8 Histone acetylation 34

1.8.1 HATs 34

1.8.2 Histone deacetylases (HDACs) 36

1.9 Crosstalk between H3K4me3 and acetylation 40

1.10 *Caenorhabditis elegans:* a powerful genetic model 41

1.11 Vulval development in *C. elegans* 44

1.11.1 Receptor Tyrosine Kinase (RTK)/Ras GTPase/MAP kinase (MAPK), Notch and Wnt signalling promotes vulval development 46

1.11.2 Determination of Vulval fate 48

1.11.3 Synthetic Multivulva (SynMuv) genes 51

1.12 Outline of the thesis 53

Materials and Methods 55

2.1 Strains and their Maintenance 55

2.1.1 Preparation of NGM 56

2.2 Genotyping by single-worm PCR 56

2.3 Genetic crosses 59

2.3.1 Double mutants generated in this study 61

2.4 Outcrossing 61

2.5 Synchronization of *C. elegans* by bleaching 63

2.6 Western Blot 63

2.7 Brood size assay 64

2.7.1 Brood size at 15 °C and 20 °C 64

2.7.2 Brood size at 25 °C 66

2.8 Development assay 68

2.9 Analysis of chromosomal abnormalities 70

2.10 Agar pad preparation 70

2.11 SynMuv suppressor screen 70

2.12 RNAi plates 71

2.12.1 Fertility assay at 20 °C 71

2.12.2 Fertility assay at 15 ° C. 71
2.12.3 Fertility assay at 25 ° C 71
2.12.4 SynMuv suppressor screens 73
2.12.5 Egg laying defective assay 73
2.12.6 Development assay 73

2.13 Fertility assay of TSA treated *C. elegans* 74

2.14 T-test for fertility assays 74

2.15 Fisher's exact-test for multivulva assays 75

2.16 Heat Shock experiments 75

2.17 Salt induction experiments 75
2.17.1 Salt induction reporter assay 76
2.17.2 Fluorescence Image Quantification 76
2.17.3 Salt induction experiment for qPCR 76

2.18 RNAi efficiency experiment for qPCR 77

2.19 RNA extraction and qPCR 77
2.19.1 RNA extraction 77
2.19.2 cDNA synthesis 78
2.19.3 qPCR 78

2.20 Efficiency of primers 81

2.21 The composition of media and buffers 81
2.21.1 The composition of NGM media 81
2.21.2 The composition of Luria-Bertani (LB) medium 81
2.21.3 The composition of M9 buffer 82
2.21.4 The composition of the bleach solution 82
2.21.5 The composition of TBST buffer solution 82
2.21.6 The composition of PK lysis buffer 82

Investigating the role of CFP-1 in *C. elegans* development 83

3.1 Introduction 83

3.2 Results 85

3.2.1 The global level of H3K4me3 is drastically reduced in the *cfp-1(tm6369)* mutant 85

3.2.2 CFP-1 promotes fertility 87

3.2.3 Loss of CFP-1 function results in temperature-sensitive sterility 90

3.2.4 CFP-1 promotes *C. elegans* development 92

3.2.5 Loss of CFP-1 function results in gross chromosomal changes in germ cells. 95

3.2.6 Reduced fertility in *cfp-1(tm6369)* and *set-2(bn129)* mutants is not due to increased apoptosis in germ cells 97

3.3 Discussion 100

Investigating the role of CFP-1 and SET-2 in gene induction 103

4.1 Introduction 103

4.2 Results 106

4.2.1 A salt-inducible reporter gene is strongly expressed in *cfp-1(tm6369)* and *set-2(bn129)* mutants 106

4.2.2 The *gpdh-1* gene is strongly induced in *cfp-1(tm6369)* and *set-2(bn129)* mutants 109

4.2.3 Loss of CFP-1 or SET-2 function results in increased transcription of heat inducible genes 111

4.3 Discussion 114

SET-2/COMPASS independent function of CFP-1 in vulval development 119

5.1 Introduction 119

5.2 Results 122

5.2.1 CFP-1 promotes vulval induction in a SynMuv mutant 122

5.2.2 The *cfp-1(tm6369)* genetic mutant confirmed the role of CFP-1 in the SynMuv phenotype. 125

5.2.3 Role of CFP-1 in the RTK/Ras/MAPK pathway during vulval development 127

5.2.4 Role of CFP-1 in the Notch pathway during vulval development 129

5.2.5 Suppressor/enhancer screen 129
5.2.6 PBRM-1 suppresses the function of CFP-1 in vulval induction 133
5.2.7 GTBP-1, SIN-3 and HDA-1 enhance the function of CFP-1 in the SynMuv phenotype of the *lin-15AB(n765)* mutant 133
5.2.8 GTBP-1 and RTK/Ras/MAPK signalling 135
5.2.9 CFP-1 and GTBP-1 genetically interact with each other in *C. elegans* development 137

5.3 Discussion 139
5.3.1 Role of CFP-1 in vulval cell fates 139
5.3.2 Chromatin regulators regulate the role of CFP-1 in vulval development 140
5.3.3 GTBP-1 interacts genetically with CFP-1 in *C. elegans* development 141
5.3.4 SET-2/COMPASS independent function of CFP-1 142

CFP-1 interacts with HDAC1/2 complexes in *C. elegans* development 143

6.1 Introduction 143

6.2 Results 145
6.2.1 RNAi-mediated knockdown of *hda-1* enhances the poor fertility phenotype of *cfp-1(tm6369)* 145
6.2.2 CFP-1 cooperates with class I HDACs to promote fertility 148
6.2.3 CFP-1 interacts genetically with SIN-3, CHD-3 and SPR-1 complexes 150
6.2.4 Genetic mutants confirmed the genetic interaction between CFP-1 and HDAC1/2 complexes 152
6.2.5 TSA reduces the fertility of the *cfp-1(tm6369)* mutant 154
6.2.6 CFP-1 does not interact genetically with class II, III and IV HDACs 156
6.2.7 Investigating the genetic interaction between CFP-1 and HATs 159
6.2.8 The poor fertility and developmental delay of *cfp-1(tm6369)* or *set-2(bn129)* mutants is enhanced in *cfp-1(tm6369);set-2(bn129)* double mutants 164

6.3 Discussion 166
6.3.1 CFP-1 interacts with HDAC complexes 166

6.3.2 The interaction between CFP-1 and HAT-1 167
6.3.3 SET-2 and HATs..... 168
6.3.4 The interaction between CFP-1 and HDACs is independent of SET-2 168

Discussion and future work..... 171

7.1 Discussion..... 171
7.1.1 Role of H3K4me3 in *C. elegans* development. 171
7.1.2 Role of H3K4me3 in gene expression 172
7.1.3 Functions of CFP-1 independent of the SET-2/COMPASS complex 173
7.1.4 CFP-1 plays a role in organogenesis..... 173
7.1.5 CFP-1, an epigenetic regulator..... 174

7.2 Future direction 177

7.3 Conclusion 180

References 181

Appendix 228

9.1 Appendix: Gel images depicting genotype of mutants 228

9.2 Appendix: Western blot analysis..... 229

9.3 Appendix: CT values of *gpdh-1*, *pmp-3* gene and *tba-1* genes 230

9.4 Appendix: CT values of heat inducible and housekeeping genes..... 231

List of figures

Figure 1.1: Transcription regulation from core promoters..... 3

Figure 1.2: Chromatin and histone modifications..... 12

Figure 1.3: Diagrammatic representation of the COMPASS complex indicating the conservative nature 15

Figure 1.4: Recruitment of Set1/COMPASS complex 23

Figure 1.5: Diagrammatic representation of the mammalian Cfp1 protein 31

Figure 1.6: HDAC1/HDAC2 multi-protein complexes 39

Figure 1.7: Life Cycle of *C. elegans* 43

Figure 1.8: Vulva development at the third larval stage 45

Figure 1.9: Determination of Vulval fate..... 50

Figure 2.1: Genetic cross 60

Figure 2.2: Outcrossing of single mutants 62

Figure 2.3: Schematic representation of worm transfer for fertility assessment at 20 °C and 15 °C 65

Figure 2.4: Schematic representation of worm transfer for fertility assessment at 25 °C 67

Figure 2.5: Scoring scheme for development assay..... 69

Figure 2.6: Schematic representation of worm transfer for fertility screen 72

Figure 3.1: Molecular characterization of the *cfp-1(tm6369)* allele 86

Figure 3.2: CFP-1 is required for fertility..... 88

Figure 3.3: *cfp-1(tm6369)* and *set-2(bn129)* display temperature-sensitive sterility 91

Figure 3.4: *cfp-1(tm6369)* and *set-2(bn129)* mutants development is delayed..... 93

Figure 3.5: CFP-1 deficiency results in defects in chromosome 96

Figure 3.6: Knockdown of key apoptotic genes did not rescue the poor fertility of *cfp-1(tm6369)* and *set-2(bn129)* mutants..... 99

Figure 4.1: Loss of CFP-1 or SET-2 function results in stronger expression of GFP protein 108

Figure 4.2: The expression of *gpdh-1* is higher in *cfp-1(tm6369)* and *set-2(bn129)* mutants 110

Figure 4.3: Heat shock response genes are strongly expressed in *cfp-1(tm6369)* and *set-2(bn129)* mutants..... 113

Figure 4.4: Model showing regulation of gene induction by H3K4me3 118

Figure 5.1: Synthetic multivulva phenotype 121

Figure 5.2 : Expression levels of the *cfp-1* mRNA upon RNAi negative control and *cfp-1* RNAi. 124

Figure 5.3: CFP-1 and GTBP-1 genetically interact to promote fertility and developmental rate 138

Figure 6.1: RNAi-mediated knockdown of *hda-1* reduces the brood size of *cfp-1(tm6369)* 147

Figure 6.2: CFP-1 interacts genetically with HDAC 1/2 to promote fertility..... 149

Figure 6.3: CFP-1 interacts with key subunits of HDAC 1/2 complexes to promote fertility..... 151

Figure 6.4: Genetic interaction between CFP-1 and HDAC1/2 subunits 153

Figure 6.5: TSA treatment further reduces the brood size of *cfp-1(tm6369)*..... 155

Figure 6.6: CFP-1 does not interact with class II, III or IV HDACs 158

Figure 6.7: Genetic interaction between CFP-1 and HATs 162

Figure 6.8: Loss of CFP-1 and SET-2 resulted in a further reduction of fertility and enhanced the developmental delay phenotype. 165

Figure 6.9: Proposed model indicating that CFP-1 cooperates with SET-2/COMPASS and/or with HDACs in a context-dependent manner 170

List of tables

Table 1.1: Conserved subunits of the COMPASS and COMPASS-like complexes 16

Table 2.1 List of strains used in this study 55

Table 2.2: PK digestion program 57

Table 2.3: Primers used for genotyping 57

Table 2.4: PCR program 58

Table 2.5: Primers used for qPCR 79

Table 2.6: qPCR program 80

Table 5.1: Percentage of *lin-15AB(n765)* mutant animals with the SynMuv phenotype during *cfp-1* RNAi. 123

Table 5.2: Percentage of animals with the SynMuv phenotype and the average number of protrusions 126

Table 5.3: Percentage of *let-60(n1046), let-23(sa62)* and *lin-12(n137n460)* mutants exhibiting the Muv phenotype during RNAi-mediated knockdown of *cfp-1* 128

Table 5.4: RNAi screen identified three genes that could enhance and one gene that could suppress the role of *cfp-1* in the vulval development 131

Table 5.5: Number of protrusions in *lin-AB(n765)* and *cfp-1(tm6369);lin-15AB(n765)* in RNAi knockdown of *pbrm-1, gtbp-1, sin-3* or *hda-1* 134

Table 5.6: Percentage of *let-23(sa62)* mutants with the Muv phenotype in RNAi of *cfp-1*, *gtbp-1* and *cfp-1*+*gtbp-1* RNAi 136

Table 6.1: Egg-laying defective (Egl) phenotype of the *cbp-1(ku258)* mutant 163

List of Abbreviations

5mC	5-methylcytosine
AC	Anchor cell
CBP	CREB-binding protein
CFP1	CxxC finger protein 1
CGIs	CpG islands
ChIP-seq	Chromatin immunoprecipitation sequencing
COMPASS	Complex of proteins associated with Set1
CoREST	Corepressor of REST
Cps	COMPASS
CSL	CBF1/Su(H)/LAG-1
CTD	C-terminal domain
DAPI	4′,6-diamidino-2-phenylindole
DBD	DNA binding domain
DCC	Dosage compensation complex
DIC	Differential interference contrast
DMSO	Dimethyl sulfoxide
Dnmt	DNA methyltransferase
Dpc	Days post conception
DREAM	Dimerization partner, Retinoblastoma-like, E2F, and MuvB
DSB	Double-Strand DNA Break
DSIF	DRB-sensitivity-inducing factor
DSL	Delta/Serrate/LAG-2
EcR	Ecdysone receptor
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ESCs	Embryonic stem cell
FOXO	Forkhead box type O transcription factor
GAP	GTPase-activating protein
Gcn5	General control non-repressible 5
GFP	Green fluorescent protein

GNAT	Gcn5-related histone *N*-acetyltransferases
GPDH-1	Gycerol- 3- phosphate dehydrogenase
GPS2	G protein pathway suppressor 2
GTBP1	GAP binding protein 1
GTFs	General transcription factors
HAT	Histone acetyltransferase (HAT)
HCF1	Host cell factor 1(HCF1)
HDAC	Histone deacetylase
HP1	Heterochromatin protein 1
HSE	Heat shock element
HSF1	Heat shock transcription factor 1
HSPs	Heat shock proteins
HSR	Heat shock response
H3K4me3	Histone 3 lysine 4 trimethylation
ING1	Inhibitors of growth 1
ING2	Inhibitors of growth 2
ING4	Inhibitor of Growth 4
ISWI	Imitation Switch
KMT	Lysine methyltransferase
LB	Luria-Bertani
lncRNA	Long non-coding RNA
LSD1	Lysine-specific demethylase 1
MAPK	MAP kinase
MBD	Methyl-CpG-binding domain
MEF2	Myocyte-specific enhancer factor 2
MLL	Mixed Lineage Leukaemia
Mrt	Mortal germline
MTA	Metastasis-associated protein
Muv	Multivulva
MYST	MOZ, Ybf2, Sas2, and Tip60
NCOA6	Nuclear receptor coactivator
NCoR	Nuclear receptor corepressor
NELF	Negative elongation factor

NFκB	Nuclear factor-κB
NGM	Nematode growth media
NLS	Nuclear localisation signal
NMIs	Non-methylated islands
NuRD	Nucleosome remodelling and histone deacetylase
NURF	Nucleosome Remodeling Factor
PA1	PTIP associated 1
Paf1	Polymerase association factor 1
PBRM1	Polybromo 1
PCAF	P300/CBP-associated factor
PcG	PolyComb group
PCR	Polymerase chain reaction
PHD	Plant homeodomain
PIC	Pre-initiation complex
Pol II	Polymerase II
PPARγ	Peroxisome proliferator activated receptor-γ
PTIP	Pax transactivation domain interacting protein
PTMs	Post-translational modifications
Pvul	Protruded vulva
qPCR	quantitative PCR
RbAp46/48	Retinoblastoma-associated protein 46/48
REST	RE-1 Silencing Factor
RNAi	RNA interference
RTK	Receptor Tyrosine Kinase
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel Electrophoresis
SEM	Standard error of the mean
SET	*Su(var)3–9*, *Enhancer of zeste* (*E(z)*), and *trx*
SID	SET1 interacting domain
Sir	Silent information regulator
SiRNA	Small interfering RNA

SMRT	Silencing mediator of retinoic acid and thyroid hormone receptor
SWI/SNF	Switch/Sucrose non-fermentable
SynMuv	Synthetic Multivulva
TBE	Tris Borate EDTA
TBL1	Transducin β-like 1
TBLR1	TBL-related 1
TBST	Tris-buffered saline with tween
TFII	Transcription factor II
TFs	Transcription factors
TGFα	Transforming growth factor α
Trr	Trithorax-related
Trx	Trithorax
TS	Targeting sequence
TSA	Trichostatin-A
TSS	Transcription start site
VPCs	Vulval precursor cells
VU	Ventral uterine
Vul	Vulvaless
Wt	Wild-type

Chapter 1

Introduction

1.1 Tightly controlled gene expression programme at the organismal level

Genes encode proteins that dictate cellular activity. Therefore, genes must be expressed correctly in a cell to determine the appropriate function of a cell. Gene expression is a process by which information stored in genes are transcribed to produce mRNA which is later translated to proteins (Shandilya and Roberts, 2012, Buratowski, 2009). Higher organisms are composed of diverse cell types that perform specialised and specific functions, however, each of these cells contains the entirety of the genome. In order to have specialized and specific functions, each cell must employ a tightly regulated gene expression. In addition, these cells are continuously exposed to extracellular and intracellular cues, leading to a change in the gene expression programme to provide an appropriate response. Expression of the gene is tightly regulated at two interconnected levels. Transcription factors (TFs) and the cellular transcription apparatus are involved at the first level, and chromatin and chromatin regulators are involved at the second level (Lee and Young, 2013).

1.1.1 Transcription

RNA polymerase II (Pol II) catalyses the transcription of all the protein-coding genes. It uses DNA as a template to synthesise mRNA. Various TFs help to recruit Pol II at promoter regions to initiate transcription. As shown in **Figure 1.1**, transcription is initiated with the binding of the transcription factor IID (TFIID) complex to DNA after which TFIIA and TFIIB are recruited subsequently. TFIIA and TFIIB stabilise the TFIID complex and recruit Pol II and TFIIF. TFIIF acts as a bridge between Pol II and the assembling transcriptional complex. Once Pol II is recruited, TFIIE joins the complex and recruits TFIIH (Shandilya and Roberts, 2012). The Pol II contains a C-terminal domain (CTD), which consists of multiple repeats of the ^{1}Tyr-Ser-Pro-Thr-Ser-Pro-Ser^{7} sequences. Once, transcription is initiated, serine 5 of the CTD tails is phosphorylated by TFIIH (Shandilya and

Roberts, 2012, Buratowski, 2009). The Pol II synthesises small oligonucleotides of 4-5 nt before it forms longer hybrids (Shandilya and Roberts, 2012). The P-TEFP complex, which contains CDK9 kinase, phosphorylates CTD serine 2 of heptapeptide repeat sequences. This causes Pol II to escape from initiation and enter into elongation (Shandilya and Roberts, 2012, Sikorski and Buratowski, 2009). The phosphorylation at serine 2 is also needed to release Pol II from a paused state (Adelman and Lis, 2012). Negative elongation factor (NELF) along with DSIF (DRB-sensitivity-inducing factor containing SPT4/SPT5) prevents the forward movement of Pol II and keeps it in a paused state downstream of the TSS (**Figure 1.1**) (Adelman and Lis, 2012). The CDK9 kinase phosphorylates the NELF and DSIF. After phosphorylation NELF and DSIF dissociate and Pol II is released from the promoter (Adelman and Lis, 2012). Enhancers can promote the release of Pol II by recruiting cofactors such as p300 and bromodomain-containing protein 4 (BRD4) that either recruit and stimulate CDK9 or directly affect the release of paused Pol II. Subsequently, other elongation factors such as ELL, TFIIS, Paf complex and super elongation complex (SEC) bind to establish a steady rate of transcription (Levine, 2011, Smith et al., 2011, Zhou et al., 2012, Luo et al., 2012). The transition of CTD state from unphosphorylated to phosphorylated state also allows Pol II to interact with factors that are required for downstream events such as termination, cleavage, mRNA splicing and 3′ polyadeadenylation (Orphanides and Reinberg, 2002). During elongation, Pol II moves in a 5′ to 3′ direction along the target DNA. Termination occurs when Pol II reaches the end of the encoding DNA template (Orphanides and Reinberg, 2002). The mRNA transcript is then exported from the nucleus to the cytoplasm. In the cytoplasm, mRNA is translated into a functional protein (Orphanides and Reinberg, 2002).

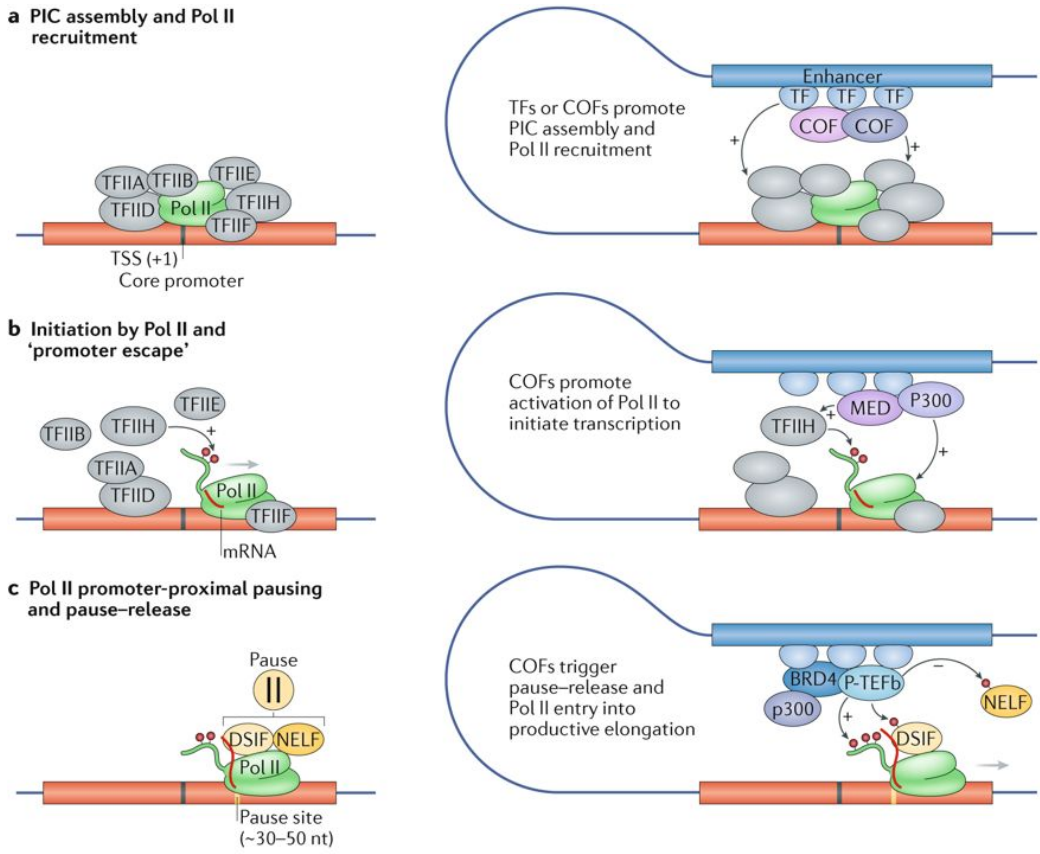


Figure 1.1: Transcription regulation from core promoters

(a) Pre-initiation complex (PIC) assembly and Pol II recruitment. PIC contains Pol II and general transcription factors (GTFs): TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH (image on the left). Enhancers can promote the assembly of the PIC by recruiting TFs and co-factors (COF) that interact with GTFs or Pol II (right image). (b) Transcription initiation. After PIC assembly, transcription is initiated by the Pol II at the transcription start site (TSS). Once, transcription is initiated, serine 5 and 7 of the CTD tails of Pol II is phosphorylated by TFIIH (image on the left). Enhancers can promote this process by recruiting COFs such as histone acetyltransferase p300 or the Mediator complex (MED) (image on the right). (C) Promoter pausing and elongation. Negative elongation factor (NELF) along with DSIF (DRB-sensitivity-inducing factor containing SPT4/SPT5) prevents the forward movement of Pol II and keeps it in a paused state downstream of the TSS. The paused Pol II is released by the phosphorylation of DSIF and NELF by the CDK9 kinase of the P-TEFP complex (image on the left). Enhancers can promote the release of Pol II by recruiting cofactors such as p300 and

bromodomain-containing protein 4 (BRD4) that either recruit and stimulate CDK9 or directly affect the paused Pol II release (image on the right). The picture was taken from Haberle and Stark (2018).

1.1.2 Transcription factors and gene expression

Pol II cannot independently recognise or bind to promoter regions and therefore requires the TFs. TFs bind to regulatory elements upstream to the transcription initiation site and facilitates the recruitment of Pol II in a context-dependent manner (Lee and Young, 2013). TFs make sequence specific contact with the target DNA by using their DNA-binding domain and make protein-protein interaction with other factors via a regulatory domain. These interactions allow TFs to make specific contacts with DNA and protein, therefore regulating the expression of cell-type specific genes. Each TF binds to a 6–12 bp-long specific DNA sequence. However, the specific DNA sequence is often not enough to identify the target sites of TFs. For example, in the mouse genome, if one TF recognises specific 8 bp sequences, a number of potential binding sites could be ~45000 (Niwa, 2018). This suggests that another mechanism is required for the target gene activation. One of the mechanisms is combinatorial occupancy. In the enhancer regions with clusters of different TF binding sites, TFs interact with each other and regulate discrete and precise patterns of transcriptional activity (Spitz and Furlong, 2012). For example, the myocyte-specific enhancer factor 2 (Mef2) TFs family of MADS-box proteins act as a central regulator of cell survival, proliferation, differentiation and apoptosis in a range of cell types, including cardiac, skeletal, and smooth muscle, brain, bone and lymphocytes (Potthoff and Olson, 2007, Cunha et al., 2010). The diverse function of Mef2 is mediated by cooperative activity with specific regulators and extracellular signals (Potthoff and Olson, 2007, Cunha et al., 2010). Another TF, Yin Yang 1, can act as a transcriptional initiator, activator or repressor depending on the presence or absence of additional factors (Shi et al., 1997). Similarly, in *Drosophila melanogaster*, pMAD (the phosphorylated form of MAD) which is the effector of bone morphogenetic protein and decapentaplegic signalling, regulates particular cell fates by interacting with cell type-specific TFs. pMAD interacts with Scalloped in the wing imaginal disc and Tinman in the dorsal mesoderm, thereby restricting a specific cell fate to only those cells which have both pMAD and the second TF (Lee and Frasch, 2005, Guss et al., 2001, Spitz and Furlong, 2012).

Depending upon the developmental stages or conditions, TFs often occupy diverse sets of enhancers. Some of the TFs are exclusively expressed at specific stages whereas some are always present but only bind during specific developmental stages. For example, during *D. melanogaster* embryogenesis, some mesoderm-specific TFs bind to enhancers at either early or late developmental stages, even though they are always present and available (Spitz and Furlong, 2012). The concentration of TFs also plays a crucial role in target gene expression. For example, in *D. melanogaster*, Dorsal, a Rel-containing sequence-specific TF, controls dorsal–ventral patterning and gastrulation by regulating target genes in a concentration-dependent manner (Papatsenko and Levine, 2005, Stathopoulos et al., 2002). Low level of Dorsal regulates the neurogenetic gene in ventrolateral regions of the embryo, whereas high Dorsal levels activate enhancers in the ventral region (Spitz and Furlong, 2012, Zinzen et al., 2006). Additionally, post-translational modifications (PTMs) of TFs also play a role in gene regulation. For example, during environmental changes, transcription of specific genes may need to be altered, this could be mediated by ubiquitination of TFs for degradation (Salghetti et al., 2001). Similarly, additional PTMs of TFs, such as acetylation, phosphorylation and methylation could alter the activity, subcellular localization, or ability to recognize binding sites on DNA (Zhang and Reinberg, 2001).

The timing of expression, the timing of DNA occupancy, PTM of TFs, the concentration of the individual TF, interaction with other TFs or proteins control the genetic network that drives developmental progression (Spitz and Furlong, 2012).

1.1.3 Role of TFs in development

TFs regulate important parts of development; therefore, the deletion of TF genes results in developmental abnormalities. For example, deletion of the Gap gene, *buttonhead*, in *D. melanogaster* results in loss of mandibular, intercalary, and antennal head segments (Schock et al., 1999, Wimmer et al., 1996). Similarly, mutation of a hox gene, *antennapedia*, results in the development of an antennal imaginal disc into legs instead of antennae (Beira and Paro, 2016). Another hox gene *Ultrabithorax (Ubx)* functions as a major determinant of segment identity in

the abdominal and thoracic region. Ubx represses appendage development and initiates abdominal development (Castelli-Gair and Akam, 1995). Mutation of the antennapedia homologue, *mab-5,* in *Caenorhabditis elegans* results in cell death, alterations in a number of neuronal cell fates and cell migrations (Kenyon, 1986). Deletion of the muscle-specific helix-loop-helix TF, *hlh-1* results in defects in muscle and mutant animals die as larvae or young adults (Chen et al., 1994). Similar to *D. melanogaster* and *C. elegans*, in mammals, Hox genes specify developmental boundaries and are required for cell fate determination during morphogenesis (Innis, 1997). Pioneer TFs FoxA and GATA, play important roles in gut development and are evolutionarily conserved in all metazoans (Zaret and Carroll, 2011).

1.1.4 Gene regulation and development in the context of chromatin

Chromatin and chromatin regulators also play a crucial role in gene regulation.

1.1.4.1 Chromatin

In higher organisms, the organisation of DNA is complex. DNA associates with histone proteins to form a complex structure called chromatin (Kornberg, 1974, Kornberg and Lorch, 1999). Histones provide a base to organise DNA into structures called nucleosomes (**Figure 1.2**) (Kornberg, 1974, Kornberg and Lorch, 1999). A nucleosome consists of less than two turns of DNA (146 bp) wrapped around an octameric protein which is made up of two molecules each of Histone 2A (H2A), Histone (H2B), Histone 3 (H3) and Histone 4 (H4) (Kornberg, 1974, Kornberg and Lorch, 1999). Histone 1 proteins link the nucleosomes together to form a compact structure (Kornberg, 1974, Kornberg and Lorch, 1999). The classical model suggests that an 11 nm nucleosome makes a polymer and folds into a 30 nm structure, which further folds into chromosomes (Baumann, 2017). However, a recent study using chromatin electron microscopy tomography imaging suggests that the chromatin is organised into a flexible and disordered chain of 5 nm to 24 nm lengths (Baumann, 2017, Ou et al., 2017). Flexible fibres bend at various lengths and contact between and within themselves to achieve structural compaction (Baumann, 2017, Ou et al., 2017). Based on the structure and function, chromatin is

divided into two territories viz., heterochromatin and euchromatin. Heterochromatin is highly condensed, and genes in this region are transcriptionally silent. In contrast, euchromatin is less condensed, and genes in this region are transcriptionally active (Allshire and Madhani, 2018).

1.1.4.1.1 Heterochromatin

Heterochromatin regions are highly condensed, and when stained with DNA binding dyes, they appear as dark irregular particles (Allshire and Madhani, 2018, Holmquist et al., 1982). The DNA in heterochromatin regions is less accessible to proteins such as TFs (Allshire and Madhani, 2018). Heterochromatin is further divided into constitutive heterochromatin and facultative heterochromatin (Grewal and Jia, 2007). Constitutive heterochromatin is transcriptionally inactive throughout the cell cycle. It encompasses the region containing highly repetitive DNA sequences such as satellite DNA, centromeres, telomeres, and the mating type loci (Grewal and Jia, 2007). On the other hand, facultative heterochromatin encompasses DNA that is inactive in particular cells or at a phase in development but may become active on certain cell types. For example, during the X chromosome inactivation in female mammals, one X chromosome is active while another X chromosome is inactivated (Grewal and Jia, 2007, Allshire and Madhani, 2018).

1.1.4.1.2 Euchromatin

Euchromatin is less condensed and forms light bands when stained with DNA-binding dyes. The less condensed structure allows different proteins, such as RNA polymerase and TFs, to bind and initiate transcription (Annelie Strålfors, 2011). In contrast to heterochromatin, euchromatin regions are enriched in transcriptionally active genes (Annelie Strålfors, 2011). However, unlike heterochromatin, not all euchromatin regions are similar. Euchromatin regions are enriched for different types of modifications (Bannister and Kouzarides, 2011). For example, histone 3 lysine 4 (H3K4) mono-methylation (-me1) modification is enriched at enhancer regions, whereas H3K4 trimethylation (-me3) is highly enriched at promoter regions (Santos-Rosa et al., 2002, Barski et al., 2007).

1.1.4.1.3 Role of chromatin in gene expression and development.

For the transcription to occur, DNA must be accessible to TFs. Therefore, nucleosomes must be displaced to open the chromatin structure and give access to TFs. Occupancy of TFs is highly corelated with nucleosome-depleted regions (NDRs) (Li et al., 2011, Degner et al., 2012, John et al., 2011, Liu et al., 2006). This suggests that the position of nucleosome helps to determine the TFs occupancy sites. Displacement of nucleosomes is particularly required for the occupancy of low-affinity motifs, therefore in some cases, chromatin remodelling is required before TF binding. For example, in mammals, chromatin remodelling by SWI/SNF enzyme BRG1 is required for the recruitment of glucocorticoid receptor at a number of sites (John et al., 2008). BRG1 is also essential for the expression of the β-globin gene during embryonic erythropoiesis (Ho and Crabtree, 2010). Similarly, hormone-responsive elements require the chromatin remodelling by Brahma-associated factor (BAF) and P300/CBP-associated factor (PCAF) to facilitate receptor binding (Vicent et al., 2009, Spitz and Furlong, 2012). In *D. melanogaster*, chromatin remodelling by BAP complex is required for specifying segmentation and depletion of components of the BAP complex in zygote results in multiple defects in organ formation and is lethal at late embryogenesis (Ho and Crabtree, 2010).

PTMs of histone tails can also influence the TFs occupancy. The X-ray crystal structure of the nucleosome shows that the tails of histones protrude outside of the core particle to contact neighbouring histones (**Figure 1.2**) (Luger et al., 1997). Amino acids present in these protruded histone tails can be post-translationally modified by methylation on lysine or arginine, phosphorylation on threonine or serine residues, ubiquitylation and sumoylation of lysines, and acetylation of lysines (Bannister and Kouzarides, 2011). Some of the histone modifications can alter the chromatin structure by changing the histone charge and regulate gene expression. For example, histone acetylation neutralises the positive charge present on the lysine residues of histone, thereby decreasing their affinity to DNA (Kouzarides, 2007). Histone modifications can also act as binding sites for a diverse array of chromatin remodelers and DNA-interacting proteins that contribute to transcriptional control

(Kouzarides, 2007). For example, trimethylation of histone 3 lysine 9 (H3K9me3) which is considered to be a hallmark of heterochromatin helps in the maintenance and formation of heterochromatin by recruiting heterochromatin protein 1 (HP1) (Lehnertz et al., 2003, Bannister et al., 2001). Loss of function of histone 3 lysine 9 (H3K9) methyltransferase Suv39h1 and Suv39h2 in mice results in the death of some embryos, and the inactivation of another histone H3K9 methyltransferase, G9a, leads to early embryonic lethality (Li, 2002). Similar to H3K9me3, histone 3 lysine 27 trimethylation (H3K27me3) is also involved in gene repression. H3K27me3 is mainly found in facultative heterochromatin regions, and it represses the expression of key transcriptional regulators during development (Bernstein et al., 2006). In *D. melanogaster*, Polycomb (Pc) genes, which deposit H3K27me modifications, contribute to Hox gene expression and is required for proper segmentation. Inactivation of Pc genes results in developmental defects (Pirrotta, 1998, Ingham and Whittle, 1980, Shilatifard, 2012). In *C. elegans,* PcG genes play a role in the regulation of gene expression, the specification of germ cells, vulval development and development of embryos (Reinke et al., 2013, Ahringer and Gasser, 2018).

In contrast to H3K9me3 and H3K27me3, Histone 3 lysine 36 trimethylation (H3K36me3) is associated with gene activation and is enriched in actively transcribed genes (Bannister et al., 2005). H3K36me3 prevents erroneous transcription initiation by recruiting histone deacetylase (HDAC) complexes (Carrozza et al., 2005, Joshi and Struhl, 2005). Another histone modification that is associated with active transcription is H3K4me (Shilatifard, 2008, Shilatifard, 2012). H3K4 can be mono-, di- or tri-methylated and these different methylation levels are enriched at different regions of the genes (Barski et al., 2007, Wang et al., 2009). For example, H3K4me3 are enriched at promoters, and H3K4me2 are present at promoters regions downstream of H3K4me3, whereas H3K4me1 are enriched at enhancer regions (Santos-Rosa et al., 2002, Barski et al., 2007, Shilatifard, 2012, Bernstein et al., 2002). H3K4me3 can recruit plant homeodomain (PHD) containing proteins, such as inhibitors of growth 2 (ING2) and bromodomain and PHD finger transcription factor (BPTF). ING2 is a native subunit of repressive Sin3a-HDAC1 histone deacetylase complex. Binding of ING2 helps in the recruitment of Sin3a-

HDAC1 histone deacetylase complex to regulate proliferation-related genes during DNA damage (Shi et al., 2006). BPTF is a subunit of nucleosome remodelling factor (NURF), an ISWI-containing chromatin-remodelling complex. Binding of BPTF helps in the recruitment of the NURF complex to regulate Hox gene expression (Wysocka et al., 2006). Taken together these findings suggest that chromatin regulators and chromatin modifications can indeed contribute to gene expression. My project investigated the role of CxxC finger protein 1 (Cfp1) in *C. elegans* development. Cfp1 is required for H3K4me3 modifications. Thus, I will next describe H3K4me3 and Cfp1 in detail.

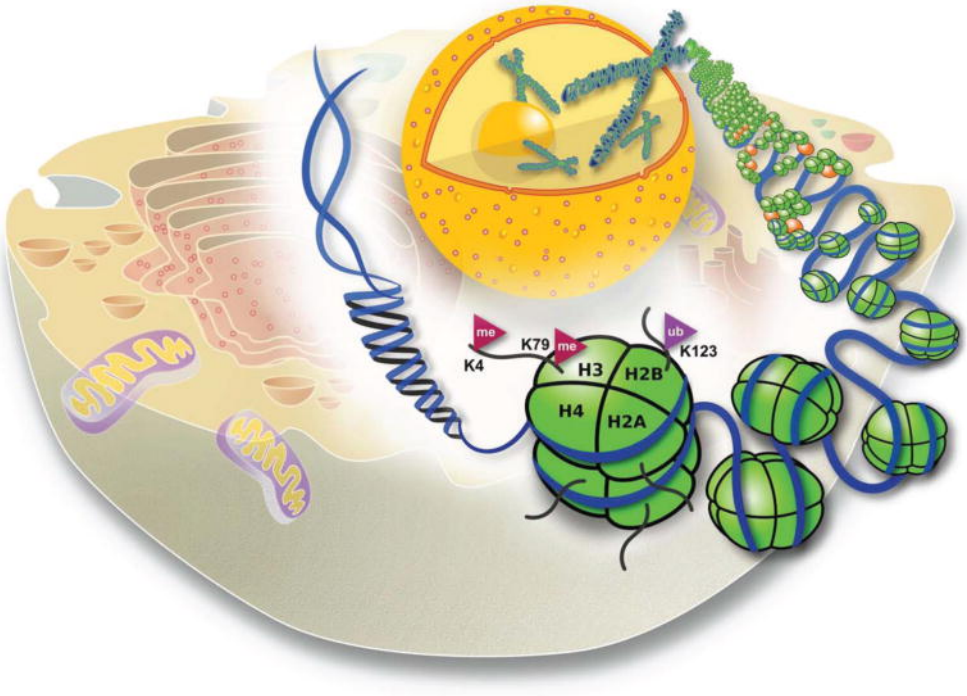


Figure 1.2: Chromatin and histone modifications

Eukaryotic DNA forms a compacted structure with histone proteins. Structural studies of nucleosomes show that the tails of histones protrude outside of a core particle to contact neighbouring histones. Amino acids present in these protruded histone tails can be post-translationally modified. The figure shows H3 and H2B modifications. Taken from Shilatifard (2012).

1.2 H3K4me3

H3K4me3 is one of the most studied histone marks and is considered to be a hallmark of active genes (Santos-Rosa et al., 2002, Barski et al., 2007, Shilatifard, 2012, Bernstein et al., 2002). H3K4me3 is distributed around the transcription start sites and is conserved in yeast, *D. melanogaster*, *C. elegans* and higher eukaryotes (Santos-Rosa et al., 2002, Barski et al., 2007, Shilatifard, 2012, Bernstein et al., 2002, Xiao et al., 2011, Ardehali et al., 2011). H3K4me3 peaks around the transcription start sites are not specific to the sense strand, as H3K4me3 peaks are also present in the promoters of antisense transcripts at the 3' ends of genes in yeast and mammals (Margaritis et al., 2012, Lavender et al., 2016).

Correlation studies suggest that the amount of H3K4me3 and the levels of nascent sense transcription is co-related, and that highly expressed genes have strong H3K4me3 peaks (distribution of H3K4me3 signals are narrow and higher) (Howe et al., 2017, Guillemette et al., 2011, Dong and Weng, 2013, Araki et al., 2009). Similarly, the breadth of H3K4me3 is also positively correlated with gene expression (Benayoun et al., 2014, Chen et al., 2015). For example, the broad H3K4me3 levels (H3K4me3 signals covering at least 500 bp to 3,500 bp) are associated with an increase in transcription elongation and enhancer activity (Benayoun et al., 2014, Chen et al., 2015). The broad H3K4me3 levels are also an epigenetic signature of tumour suppressor genes, and shortening of broad H3K4me3 peaks is associated with reduced expression of tumour suppressor genes (Chen et al., 2015, Howe et al., 2017). While H3K4me3 is predominantly associated with active genes, leading to a strong correlation with gene expression, H3K4me3 also plays an essential role in maintaining the genomic stability and development of an organism (Pena et al., 2008, Adam et al., 2018, Acquaviva et al., 2013, Sommermeyer et al., 2013). Perturbation of the H3K4me3 mark is associated with developmental defects, changes in cell fate and different diseases such as Huntington’s disease, chronic glomerular diseases, colorectal cancer (Salz et al., 2014, Shilatifard, 2012, Bertero et al., 2015, Lefevre et al., 2010, Dong et al., 2015). However, neither the function nor the mechanism of action of H3K4me3 is well understood.

1.3 Histone lysine methyltransferase (HMT)

An evolutionarily conserved complex of proteins associated with the Set1 (COMPASS) complex deposits the H3K4me3 mark at 5′ sites of the target genes (Mohan et al., 2011, Shilatifard, 2012, Krogan et al., 2002, Miller et al., 2001). Set1 is the key catalytic subunit, responsible for lysine methylation, while the associated proteins help in assembly and regulation of H3K4 methylation (Miller et al., 2001, Lee et al., 2007a, Ardehali et al., 2011, Xiao et al., 2011, Qu et al., 2018, Hsu et al., 2018). Set1 contains a 130- to 140-amino-acid motif known as the SET domain (*Su(var)3–9*, *Enhancer of zeste* (*E(z)*), and *trx*), which is essential for methyltransferase (KMT) activity (Kim et al., 2009). There are several SET domain-containing KMTs which are divided into eight different classes (Tschiersch et al., 1994, Stassen et al., 1995, Allis et al., 2007). Set1 belongs to the KMT2 class and was first identified in yeast (Miller et al., 2001, Shilatifard, 2012). As shown in **Figure 1.3,** Set1 is conserved from yeast to mammals (Miller et al., 2001, Lee et al., 2007a, Ardehali et al., 2011, Xiao et al., 2011).

In yeast, there is only one COMPASS complex, whereas there are three orthologues of the Set1/COMPASS complex: dSet1, Trx and Trr, in *D. melanogaster* (**Figure 1.3**) (Ardehali et al., 2011). In humans, there are six of these complexes present: SET1A, SET1B and four Mixed Lineage Leukaemia protein complexes (MLL1, MLL2, MLL3 and MLL4) (Shilatifard, 2012, Lee et al., 2007a, Hughes et al., 2004). In *C. elegans* there are two COMPASS complexes SET-2 and SET-16 (**Figure 1.3**) (Li and Kelly, 2011, Xiao et al., 2011, Simonet et al., 2007). Set1 in yeast, dSet1 in *D. melanogaster*, SET1A/B in mammals and SET-2 in *C. elegans* are responsible for the majority of the H3K4me3 modification observed, whereas the Trx and Trr in *D. melanogaster,* MLL1, MLL2, MLL3 and MLL4 in mammals, SET-16 in *C. elegans* are responsible for H3K4me3 modification of only a subset of genes (Shilatifard, 2012). Trx, Trr, MLL1, MLL2, MLL3, MLL4 and SET-16 are also called COMPASS-like complexes (Shilatifard, 2012). In addition to Set1, its associated subunits are also important for H3K4 methylation and are conserved from yeast to human (**Figure 1.3** and **Table 1.1**) (Miller et al., 2001, Lee et al., 2007a, Ardehali et al., 2011, Xiao et al., 2011, Hughes et al., 2004, Simonet et al., 2007).

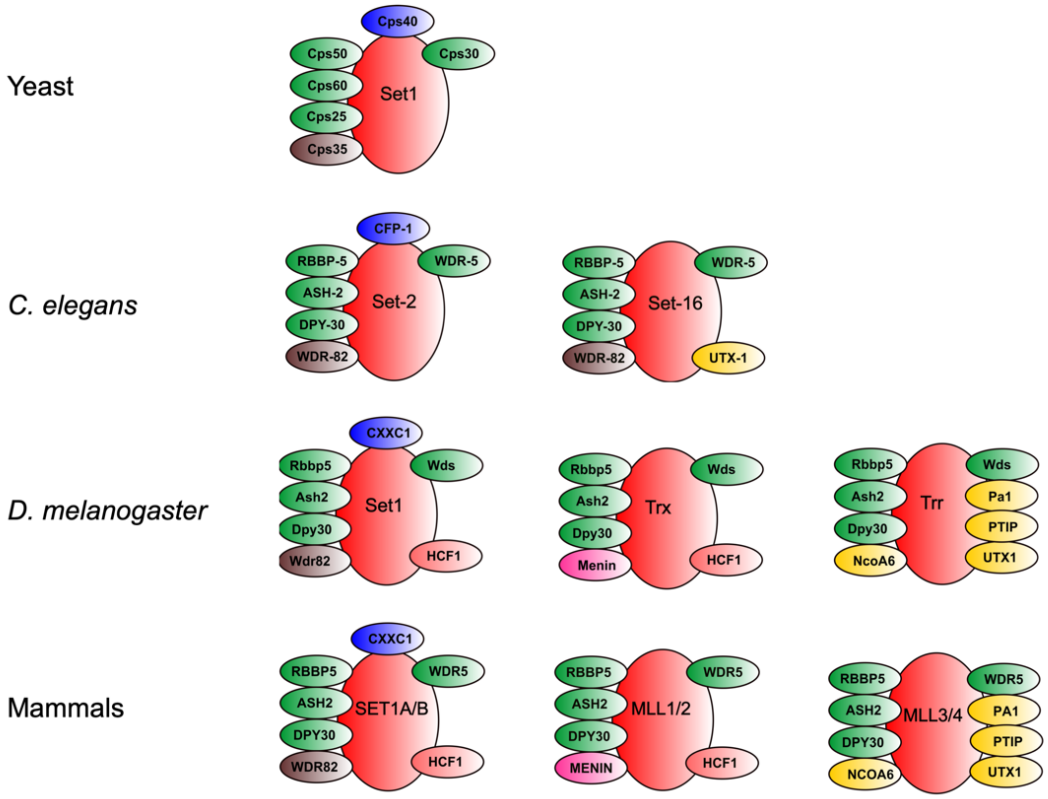


Figure 1.3: Diagrammatic representation of the COMPASS complex indicating the conservative nature

The SET domain-containing enzymes are shown in red, and conserved subunits are shown in the same colour. Set1 COMPASS was first found in yeast, and it is a founding member of the COMPASS family. In *C. elegans* there are two COMPASS complexes; SET-2 and SET-16. In *D. melanogaster,* there are three COMPASS complexes; Set1, Trx and Trr. In humans, there are 6 COMPASS complexes SET1A/B and MLL1/2/3/4. SET1A/B are orthologues of yeast Set1, and MLL1/2/3/4 are orthologues of *D. melanogaster* Trx and Trr like complexes.

Table 1.1: Conserved subunits of the COMPASS and COMPASS-like complexes

Conserved subunits of the COMPASS and COMPASS-like complexes. Adapted from Shilatifard (2012).

Yeast	Drosophila	Mammalian	Function
Set1	Set1, Trx, Trr	SET1A/B, MLL1–4	It is the major catalytic subunit of COMPASS.
Cps60 (Bre2)	Ash2	ASH2L	Required for H3K4me3/ H4mek4me2
Cps50 (Swd1)	RbBP5	RBBP5	Required for assembly
Cps40 (Spp1)	CxxC1 (dCfp1)	CxxC1 (CFP1)	Required for major H3K4me3
Cps35 (Swd2)	Wdr82	WDR82	Required for proper H3K4 di- and Trimethylation
Cps30 (Swd3)	Wds	WDR5	Required for assembly
Cps25 (Sdc1)	Dpy30	DPY30	Required for H3K4me2/3
	HCF1	HCF1	Components of Set1 complexes
	Menin	Menin	Trx and MLL1/MLL2 specific
	PTIP, PA1, NCOA6, UTX	PTIP, PA1, NCOA6	Trr and MLL3/MLL4 specific

1.3.1 Yeast Set1/COMPASS complex and its orthologues

Yeast Set1/COMPASS complex is the founding member of the COMPASS complex. Set1 is the main catalytic enzyme of the complex, and its deletion results in the loss of H3K4 mono-, di- and tri-methylation (Miller et al., 2001, Krogan et al., 2002, Roguev et al., 2001). Even though Set1 is a key enzyme, other subunits are require for the enzymatic activity and to stabilise Set1 (Schneider et al., 2005, Schlichter and Cairns, 2005, Dehe et al., 2006, Shilatifard, 2008). The other subunits of the Set1/COMPASS complex include COMPASS (Cps) 25, Cps30, Cps35, Cps40, Cps50 and Cps60, respectively (**Table 1.1**) (Miller et al., 2001). The Cps25, Cps35 and Cps60 subunits are required for H3K4 di- and tri-methylation, whereas Cps40 subunit is only required for H3K4 tri-methylation. Cps30 and Cps50 are responsible for maintaining the stability of the COMPASS complex (**Table 1.1**) (Schneider et al., 2005, Schlichter and Cairns, 2005, Dehe et al., 2006, Shilatifard, 2008).

Mammalian SET1A and SET1B COMPASS complexes are the orthologues of the yeast Set1/COMPASS complex (Shilatifard, 2012, Lee et al., 2007a, Hughes et al., 2004). Along with Set1, all the associated subunits of the Set1/COMPASS are conserved in mammals. SET1A and SET1B COMPASS complexes contain unique subunits, CxxC-finger protein-1 (CFP1) and WDR82, which are not present in MLL1/2/3/4 complexes (Shilatifard, 2012, Lee et al., 2007a, Hughes et al., 2004). SET1A and SET1B COMPASS complexes are responsible for the majority of H3K4me3 modification as knockdown of SET1A and SET1B result in bulk loss of H3K4me3 modification (Wu et al., 2008, Lee et al., 2007a, Lee and Skalnik, 2005). Even though both SET1A and SET1B are required for H3K4methylation, they also have non-overlapping functions and are required at different stages of development. In mice, the *SET1A* orthologue, *Setd1a* deleted embryos fail to gastrulate suggesting that it is required for early embryonic development (Bledau et al., 2014, Fang et al., 2016). *Setd1a* is also required for embryonic stem cells (ESCs) viability and maintenance of the self-renewal circuit (Fang et al., 2016, Cao et al., 2017). Knockdown of *Setd1a* increases the apoptosis of ESCs as well as severely affecting the proliferation of ESCs (Cao et al., 2017, Fang et al., 2016, Bledau et al., 2014,

Sze et al., 2017). *Setd1b* knockout embryos can gastrulate and grow as normal until 7.5 days post conception (dpc) but show abnormal growth after 7.5 dpc and die before 11.5 dpc (Bledau et al., 2014).

SET-2 and SET-16 COMPASS complexes are the *C. elegans* orthologues of the COMPASS complex (Li and Kelly, 2011, Xiao et al., 2011, Simonet et al., 2007). SET-2 is the orthologue of yeast Set1, while SET-16 is the orthologue of mammalian MLL family (Li and Kelly, 2011, Xiao et al., 2011, Simonet et al., 2007). SET-2 plays a major role in the deposition of H3K4me3 modification (Li and Kelly, 2011, Xiao et al., 2011, Simonet et al., 2007, Robert et al., 2014). SET-2 antagonises heterochromatin protein 1 (HP1) which is associated with gene repression, suggesting that SET-2 plays a role in establishing and/or maintaining euchromatin in *C. elegans* (Simonet et al., 2007). Similar to yeast and mammalian Set1/COMPASS, SET-2/COMPASS also contains conserved core subunits: CFP-1 (Cps40), WDR-5 (Cps30), RBBP-5 (Cps50), DPY-30 (Cps25) and ASH-2 (Cps60) (**Figure 1.3**). These subunits play an important role in the deposition of H3K4me2/me3 in embryos (Li and Kelly, 2011, Wang et al., 2011b, Xiao et al., 2011). SET-2 and core subunits of COMPASS complex also play a role in longevity as suggested by the findings that *wdr-5* and *set-2* mutants live longer whereas RNAi knockdown of *ash-2* and *rbbp-1* shortens the lifespan of worms (Greer et al., 2010). Some of the core subunits of the COMPASS complex also function within other complexes. For example, DPY-30 also functions within the dosage compensation complex (DCC) to target DCC to both X chromosomes of hermaphrodites to repress gene expression (Pferdehirt et al., 2011). WDR-5 and its paralog WDR-5.2 play a role in spermatogenesis-oogenesis transformation which is independent of their role in the COMPASS complex (Li and Kelly, 2014).

D. melanogaster dSet1 is the orthologue of yeast SET1 and plays a significant role in the deposition of H3K4me2/3 marks (Ardehali et al., 2011, Mohan et al., 2011, Hallson et al., 2012). Similar to other organisms, dSet1 also exists within a macromolecular complex, dSet1/COMPASS, which enhances its enzymatic activity (Ardehali et al., 2011). The core subunits of the Set1/COMPASS complex are also

conserved in the dSet1/COMPASS complex (**Figure 1.3**) (Ardehali et al., 2011, Hallson et al., 2012).

1.3.2 Trx and its orthologues MLL1 and MLL2

D. melanogaster Trx COMPASS complex shares four core components (Wds, Rbbp5, Ash2 and Dpy30) with the dSet1/COMPASS complex but lacks the Wdr82 and Cfp1 subunits (**Figure 1.3**) (Mohan et al., 2011). Trx plays a regulatory role in homeotic gene expression and is known to antagonise the action of the PolyComb group (PcG) complex (Klymenko and Müller, 2004). In *D. melanogaster, trx* mutation results in partial transformations of halteres into wings (Philip Ingham, 1980, Ingham, 1981, Shearn, 1989, Shearn et al., 1987, Shilatifard, 2012). Halteres are a pair of lobe-like projections that lie behind the wings. They inform the insects about the position of the body during a flight (Philip Ingham, 1980, Ingham, 1981, Shearn, 1989, Shearn et al., 1987). Trx also plays a role in apoptosis, X chromosome inactivation, cell fate specification and survival (Gonzalez and Busturia, 2009, Siebold et al., 2010). Trx COMPASS complex deposits H3K4me3 on specific gene loci, such as Hox genes and enhancer regions (Shilatifard, 2012).

The structure and function of Trx are also conserved in mammals. MLL1 and MLL2 are the mammalian orthologues of Trx and play an important role in the homeotic phenotype (**Figure 1.3**) (Shilatifard, 2012). MLL1 plays an essential role in the deposition of H3K4me3 in a subset of *Hox* genes (Guenther et al., 2005). However, deletion of menin, a subunit presents in both MLL1 and MLL2 complexes, resulted in a reduction of the majority of H3K4 methylation at entire *Hox* loci. This suggests that both MLL1 and MLL2 are require for H3K4methylation at *Hox loci* (Thiel et al., 2012, Dreijerink et al., 2009, Thiel et al., 2013, Shilatifard, 2012). Recent findings have suggested that MLL2 deposits broad H3K4me3 in the untranscribed region during oogenesis and is responsible for bulk H3K4me3 in oocytes (Dahl et al., 2016, Hanna et al., 2018).

1.3.3 Trithorax related: the orthologue of MLL3/MLL4

D. melanogaster Trithorax-related (*trr*) gene is a homologue of *trx* (Sedkov et al., 1999). *trr* encodes a SET domain, but does not play a major role in the regulation of *Hox* genes. Even though *trr* is a homologue of *trx*, deletion of the *trr* gene results in embryonic lethality suggesting that *trr* and *trx* have a non-redundant role in development (Eissenberg and Shilatifard, 2010, Sedkov et al., 1999). Trr colocalises with ecdysone receptor (EcR), a major receptor for steroid hormone in *D. melanogaster* and helps in the deposition of H3K4me3 (Sedkov et al., 2003). Deletion of the SET domain of Trr resulted in a decrease in H3K4me3 marks and transcription rate (Sedkov et al., 2003). Trr also deposits H3K4me1 at enhancer regions (Herz et al., 2012).

MLL3/4, the mammalian orthologues of Trr, also exists in complex and share core units with Set1/A/B and MLL1/2 (Shilatifard, 2012). They lack WDR82 and CFP1 but contain specific subunits: H3K27 demethylase, the pax transactivation domain interacting protein (PTIP), PTIP associated 1 (PA1), and nuclear receptor coactivator (NCOA6) (**Figure 1.3**) (Cho et al., 2007, Wu et al., 2008). MLL3/4 co-precipitate with the estrogen receptor, and their knockdown by small interfering RNA (siRNA) results in reduced receptor-dependent transcription (Goo et al., 2003, Lee et al., 2006, Mo et al., 2006). Like Trr, MLL3/4 can also methylate H3K4. RNAi of both MLL3 and MLL4 reduces H3K4me3 of the retinoic acid receptor target genes (Lee et al., 2006). MLL3/4 regulates enhancer activity by depositing H3K4me1 and promoting H3K27ac (Agger et al., 2007, Herz et al., 2012, Tie et al., 2012, Pasini et al., 2010). The SET domain of MLL3/4 helps in the deposition of H3K4me1, and H3K27 demethylase subunit removes methylation from H3K27. Removal of methylation by H3K27 demethylase facilitates the acetylation of H3K27 (Agger et al., 2007, Herz et al., 2012, Tie et al., 2012, Pasini et al., 2010).

C. elegans SET-16 is the orthologue of Trr/MLL3/MLL4, and it also exists in a similar type of complex called COMPASS-like complex (Fisher et al., 2010). Similar to the mammalian MLL3/4 complex, SET-16 has a specific subunit not found in the SET-2/COMPASS complex (**Figure 1.3**). It contains UTX-1, a JmjC

domain-containing protein, which demethylates H3K27me3/H3K27me2 (Fisher et al., 2010, Vandamme et al., 2012). UTX plays a vital role in the development: loss of function of UTX-1 results in sterility, development defects and embryonic lethality (Vandamme et al., 2012, Fisher et al., 2010, Vandamme and Salcini, 2013). A recent study suggested that the role of UTX-1 in embryonic and postembryonic development is independent of its demethylase activity, and it can be mediated by its function within the SET-16/COMPASS complex (Vandamme and Salcini, 2013, Vandamme et al., 2012).

1.3.4 Deposition of H3K4me3

Global proteomic analysis in yeast suggested that monoubiquitylation of H2B by Rad6 is essential for proper H3K4 methylation (Schneider et al., 2004, Dover et al., 2002). Rad6 is recruited to the chromatin by Bre1 (C3HC4 RING finger-containing protein), an E3 ligase (Schneider et al., 2004, Hwang et al., 2003). Cps35, a subunit of Set1/COMPASS, mediates the crosstalk between H2B monoubiquitination and H3K4 methylation (Zheng et al., 2010, Vitaliano-Prunier et al., 2008, Lee et al., 2007b). Similar to yeast, the role of H2B monoubiquitination in H3K4me3 is also conserved in *D. melanogaster* and mammals (Zhu et al., 2005, Wang et al., 2008, Mohan et al., 2010).

1.3.5 Recruitment of COMPASS complex on chromatin

The exact mechanism of how the COMPASS complex is recruited to chromatin is still not clear, and it is a major area of current research. Based on previous studies in yeast, mammals and *D. melanogaster,* the COMPASS complex could be recruited by four different mechanisms: (i) DNA sequence, (ii) long non-coding RNA (lncRNA), (iii) basal transcription factors, and (iv) nuclear hormones and DNA damage pathways.

1.3.5.1 Recruitment of the COMPASS complex by DNA sequence

Most promoter regions contain high CpG-rich DNA sequences called CpG islands (CGIs) (Deaton and Bird, 2010). Introduction of CpG rich elements without promoter sequences in the genome showed acquisition of H3K4me3 modifications suggesting that CGIs can act as a specific DNA sequence to recruit COMPASS and COMPASS-like complexes (Thomson et al., 2010, Brown et al., 2017, Clouaire et al., 2012). One of the key subunits of the COMPASS complex is Cfp1, it contains a DNA-binding domain and can bind to unmethylated CGIs (Voo et al., 2000, Carlone et al., 2002)**.** Cfp1 is located at 81% of unmethylated CpG-rich DNA sequences, and most of these sites are methylated at H3K4 (Thomson et al., 2010, Brown et al., 2017, Clouaire et al., 2012). This suggests that Cfp1 can aid the recruitment of COMPASS complexes for deposition of H3K4 methylation (**Figure 1.4**). Similarly, MLL1 and MLL2 also contain a CxxC domain and could recruit MLL1 and MLL2 complexes (Cierpicki et al., 2010, Birke et al., 2002, Ayton et al., 2004).

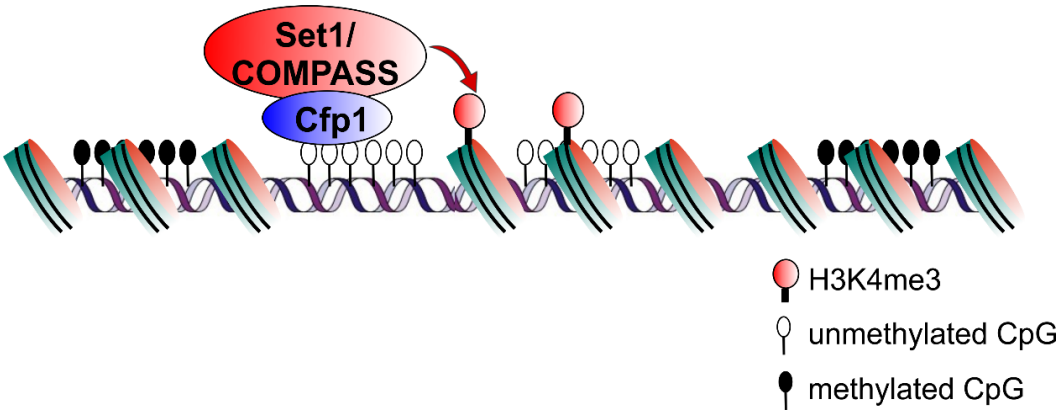


Figure 1.4: Recruitment of Set1/COMPASS complex

Most promoter regions contain unmethylated CpG-rich DNA sequences. Cfp1 contains a DNA binding domain that can bind to unmethylated CpG islands. Cfp1 binds to the unmethylated CpG islands and tethers the Set1/COMPASS complex.

1.3.5.2 Recruitment of the COMPASS complex by long non-coding RNA (lncRNA)

LncRNA can also be involved in recruiting the COMPASS complex. The WDR5 subunit of COMPASS and COMPASS-like complexes contain RNA binding domains and can bind to lncRNA to recruit the COMPASS complex (Malek et al., 2017, Wang et al., 2011a, Yang et al., 2014, Bochynska et al., 2018). In human fibroblasts, it has been observed that during *HOXA6* and *HOXA7* gene expression, the lncRNA *HOTTIP*, which is transcribed from the *HOXA* locus, interacts with WDR5 (Wang et al., 2011a). Perturbation of this interaction impairs global H3K4me3 levels, suggesting that lncRNA could interact with WDR5 subunits to recruit the COMPASS complex to chromatin (Malek et al., 2017, Wang et al., 2011a, Yang et al., 2014, Bochynska et al., 2018). Similarly, in mice, *HoxBlinc,* a lncRNA encoded by a gene in the *Hoxb* cluster helps in the recruitment of COMPASS complex at *Hoxb* cluster. Loss of *HoxBlinc* results in reduced recruitment of the COMPASS complex in *Hoxb* cluster (Deng et al., 2016).

1.3.5.3 Recruitment of COMPASS complexes by TFs

In yeast, TFs have been suggested to be involved in the recruitment of the COMPASS complex. Global proteome analysis of yeast indicated that the polymerase association factor 1 (Paf1) complex, which is associated with the elongating Pol II, can recruit the COMPASS complex (Krogan et al., 2003, Wood et al., 2003, Shi et al., 1996). Furthermore, studies show that the Paf1 complex plays a dual role in the deposition of H3K4me3. They serve as a landing pad for the Set1/COMPASS on RNA Pol II, and promote H2B monoubiquitination through Rad6/Bre1, hence promoting H3K4me (Krogan et al., 2003, Wood et al., 2003, Shi et al., 1996, Shilatifard, 2012).

In mouse ESCs, the COMPASS subunit Wdr5 interacts with TFs such as Oct4, Sox2 and Nanog (Ang et al., 2011). These TFs play an essential role in the maintenance of self-renewal or pluripotency of ESCs (Ang et al., 2011). Upon differentiation, levels of Oct4, Sox2 and Nanog decrease, and this is accompanied by a reduction in levels

of H3K4me3. This suggests that Oct4, Sox2 and Nanog could recruit the COMPASS complex to deposit H3K4me3 by binding with WDR5 (Ang et al., 2011). Moreover, Oct4 also interacts with the C-terminal region of Set1a and this interaction is independent of the interaction mediated *via* Wdr5. This suggests that Oct4 could recruit the COMPASS complex by two different mechanisms (Fang et al., 2016).

1.3.5.4 Recruitment of COMPASS complex by nuclear hormones and DNA damage pathways

Signalling pathways, such as the nuclear receptor signalling pathways and DNA damage signalling pathways, can also mediate the recruitment of the COMPASS complex (Sedkov et al., 2003, Lee et al., 2009a, Lee et al., 2009b). For example, recruitment of MLL3/4/COMPASS complexes can be regulated by the nuclear hormone receptor signalling pathway. The ecdysone receptor, in the presence of ecdysone hormone, forms a heterodimeric complex. This complex translocates to the nucleus with MLL3 and MLL4 COMPASS complexes to activate ecdysone-dependent gene expression (Sedkov et al., 2003). Similarly, MLL3 and MLL4 COMPASS complexes can also be recruited by p53-mediated DNA damage signalling. During the DNA damage response, the NCOA6 subunit of MLL3 and MLL4 COMPASS complex binds to p53 and recruits the complexes to the promoters of p53 target genes (Lee et al., 2009a).

1.4 Role of H3K4me3 in the regulation of gene expression

H3K4me3 has emerged as an important player in the regulation of gene expression (Pena et al., 2008, Wysocka et al., 2006, Doyon et al., 2006, Li et al., 2006). A strong correlation between H3K4me3 modification and gene expression levels suggests that H3K4me3 promotes gene expression. It is believed that H3K4me3 contributes to the regulation of gene expression by functioning as a binding site for TFs and chromatin modifiers (Pena et al., 2008, Wysocka et al., 2006, Doyon et al., 2006, Li et al., 2006). For example, the nucleosome remodelling factor (NURF) complex contains the H3K4me3 binding domain and can bind to H3K4me3 *in vitro*.

The NURF complex increases the accessibility of the chromatin to the transcriptional machinery by re-modulating the histone proteins (Wysocka et al., 2006). Similarly, studies in yeast and mammals suggest that H3K4me3 can act as binding sites for the recruitment of histone acetyltransferases (HATs) such as p300 (Clouaire et al., 2014, Taverna et al., 2006, Doyon et al., 2004, Hung et al., 2009). HATs increase the accessibility of the chromatin by acetylating histones (Clouaire et al., 2014, Taverna et al., 2006, Doyon et al., 2004, Hung et al., 2009). H3K4me3 can also act as a binding platform for some histone demethylases such as JMJD2C also known as KDM4C/GASC1, which bind to H3K4me3 and contribute to transcriptional regulation by removing the H3K9me3 mark (Canovas et al., 2012, Huang et al., 2006, Pedersen et al., 2014).

H3K4me3 could also play a role in gene expression by directly recruiting basal transcription factors and post-transcriptional machinery. H3K4me3 interacts with TAF3/TFIID to enhance p53-dependent transcription by directly stimulating pre-initiation complex (PIC) formation and facilitates selective gene activation in response to DNA damage (Lauberth et al., 2013). Similarly, H3K4me3 also contributes to gene expression by promoting post-transcriptional gene regulation. It could act as a binding site for the pre-mRNA splicing complex. Sims et al. (2007) found that in humans, CHD1 interacts with H3K4me3 to facilitate the recruitment of SF3a, a sub-complex of U2 snRNP, for pre-mRNA splicing. siRNA knockdown of ASH2, a subunit of SET1/COMPASS complex, and CHD1 reduced the association of U2 snRNP components, which decreases the pre-mRNA splicing on active genes (Sims et al., 2007).

1.5 H3K4me3 in gene repression

Despite H3K4me3 being associated with active gene expression, growing evidence has provided a convincing argument that Set1/COMPASS-dependent H3K4me3 can also contribute to gene silencing. Yeast Set1 has also been reported to play a regulatory role by repressing Ty1 retrotransposons and cryptic transcription (Lorenz et al., 2014, Mikheyeva et al., 2014). The gene expression profile of *SET1* mutants with low levels of H3K4me3 in yeast has shown the de-repression of subsets of

stress response genes, *TF2* retrotransposons, lncRNA and pericentromeric repeats (Lorenz et al., 2014). In mammals, H3K4me3 marks act as binding sites for inhibitors of growth 2 (ING2) which recruit the Sin3/HDAC complex to silence transcription of genes responsible for promoting cell proliferation (Pena et al., 2008). In *C. elegans*, the loss of the key H3K4me3 methyltransferase, SET-2, leads to the ectopic expression of somatic genes in the germline (Robert et al., 2014).

These findings are generated from genetic model organisms and strongly suggest that H3K4me3 can play a crucial role in gene repression under physiological conditions. It is also possible that the repressive role of H3K4me3 is context-dependent at the organismal level. Nevertheless, the mechanistic action of H3K4me3 in regulating gene expression remains to be better elucidated.

1.6 H3K4me3 in genomic stability

Over the past decade, various studies have emphasised the role of H3K4me3 in maintaining genomic stability. For example, the loss of H3K4me3 has been associated with increased DNA breaks, delay in DNA repair, inability to promote DNA-damage-induced apoptosis, and increased chromosomal abnormalities (Pena et al., 2008, Doyon et al., 2006, Palacios et al., 2010, Faucher and Wellinger, 2010, Shilatifard, 2012). H3K4me3 helps in maintaining genomic stability by acting as a binding site for different proteins or complexes. For example, H3K4me3 interacts with the C-terminal plant homeodomain (PHD) finger of inhibitors of growth 1(ING1) and helps in its recruitment to activate the p53 gene (Pena et al., 2008). Perturbation of the interaction between the C-terminal PHD of ING1 and H3K4me3 prevents cells from DNA repair or to promote DNA-damage-induced apoptosis (Pena et al., 2008).

Similarly, in mice, during DNA damage, an ING2 subunit of the repressive Sin3a-HDAC histone deacetylase complex binds to H3K4me3 and helps in the recruitment of the Sin3a-HDAC complex at the promoters of genes responsible for cell cycle progression (Shi et al., 2006). This complex deacetylates the histone and shuts off transcription of genes responsible for promoting cell proliferation to prevent the

propagation of cells with damaged DNA (Shi et al., 2006, Pena et al., 2008). H3K4me3 also helps in the formation of the PIC by interacting with TAF3/TFIID and facilitates selective gene activation by p53 in response to DNA damage (Lauberth et al., 2013).

In yeast, Set1 is the only enzyme responsible for H3K4me3 modification and loss of function of Set1 results in increased sensitivity to DNA damage-inducing agents (Li et al., 2006). Set1-deficient fission yeast cannot respond to replication stress and are deficient in non-homologous end joining repair (Pan et al., 2012). Similarly, in *C. elegans* deletion of *set-2* which is responsible for the majority of H3K4me3 modifications, results in increased sensitivity to DNA damage-inducing agents, defects in double-strand DNA break (DSB) repair and chromosome fragmentation (Herbette et al., 2017). These findings reflect that H3K4me3 plays an important role in DSB repair and maintaining the genomic stability.

1.7 Cfp1

Cfp1 is one of the major subunits of the COMPASS complex (Lee and Skalnik, 2005, Ardehali et al., 2011, Shilatifard, 2012). Cfp1 is evolutionarily conserved and is essential for the deposition of H3K4me3 marks at promoter regions of active genes (Lee and Skalnik, 2005, Ardehali et al., 2011, Shilatifard, 2012, Yu et al., 2017, Clouaire et al., 2014, Clouaire et al., 2012, Van De Lagemaat et al., 2018, Sha et al., 2018). Loss of *Cfp1* function results in a significant reduction of H3K4me3 levels at active promoters (Lee and Skalnik, 2005, Ardehali et al., 2011, Shilatifard, 2012, Yu et al., 2017, Clouaire et al., 2014, Clouaire et al., 2012). Cfp1 contains a DNA binding domain (CxxC domain), which binds to unmethylated CGIs (Voo et al., 2000, Lee and Skalnik, 2005, Brown et al., 2017, Clouaire et al., 2012). Cfp1 binds to the unmethylated CGI and targets the Set1/COMPASS complex at the promoter regions of active genes (Voo et al., 2000, Lee and Skalnik, 2005, Brown et al., 2017, Clouaire et al., 2012). Cfp1 helps to restrict the Set1/COMPASS complex at the promoter regions of active genes and also prevents the aberrant accumulation of H3K4me3 at ectopic regions (Clouaire et al., 2012).

1.7.1 Characterisation of Cfp1

Cfp1 was first identified in a ligand screen to find effectors that are specific for unmethylated CpG dinucleotides (Lee et al., 2001, Voo et al., 2000). The presence of methylation on DNA ablates the binding of Cfp1 to DNA (Lee et al., 2001, Voo et al., 2000). Selection of Cfp1 binding sites from a pool of random oligonucleotide sequences revealed that the CpG motif is necessary and sufficient to act as a binding site for Cfp1. The presence of cytosine or adenine nucleotides adjacent to the CpG dinucleotide strengthens the binding (Lee et al., 2001). Chromatin immunoprecipitation sequencing (ChIP-seq) revealed that most of the Cfp1 (91.5%) peaks are present at non-methylated islands (NMIs) suggesting that Cfp1 binds to NMI. It occupies 38.1% of total NMIs and most of those occupied NMI are annotated TSSs and had elevated RNA Pol II and H3K4me3 (Brown et al., 2017). Consistently, Cfp1 is enriched at 81% of unmethylated CpG islands *in vivo* (Thomson et al., 2010).

Cfp1 contains two copies of the PHD, basic and acidic stretches, a coiled-coil domain, the Set1 interacting domain (SID), and a highly conserved CXXC domain (**Figure 1.5**) (Lee et al., 2001). The CXXC DNA binding domain is required for the Cfp1-DNA interaction, as a mutation in cysteine (C169A) residue within the CXXC domain ablates the interaction between Cfp1 and DNA (Lee et al., 2001). The SID domain at the C-terminal end is required for the interaction between Cfp1 and the Set1, as a mutation in the cysteine (C375A) residue ablates the interaction between Cfp1 and the Set1 complex (Butler et al., 2008). The PHD fingers bind to H3K4me3 and are required for appropriate enrichment of Cfp1 at NMIs region (Brown et al., 2017).

Cfp1 is localised to the nucleus and exhibits a punctate or speckled distribution which is highly concentrated at euchromatin (Lee and Skalnik, 2002). Furthermore, it is co-localised with acetylated H3 and acetylated H4. Both, H3 and H4 acetylation are markers of euchromatin. Multiple domains that include acidic, basic, and coiled-coil domains contribute cooperative signals for the speckled nuclear distribution of

Cfp1 (Lee and Skalnik, 2002). Interestingly, the CXXC domain is required for DNA binding but is not needed for the localisation of Cfp1 at euchromatin regions suggesting that this localisation is mediated by protein-protein interactions (Lee and Skalnik, 2002, Skalnik, 2010).

Cfp1 is highly mobile in the nucleoplasm (Brown et al., 2017). Interestingly, a two amino acid substitution mutations in an alpha helix of the SID domain results in a dramatic increase in the mobility of Cfp1, which is further enhanced by a combined mutation on PHD, CXXC and SID domains (Brown et al., 2017). These suggest that Set1 restricts the mobility of Cfp1.

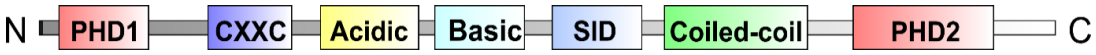


Figure 1.5: Diagrammatic representation of the mammalian Cfp1 protein
Different domains are shown in different colours. The plant homeodomains 1/2 (PHD1/2) help in binding with H3Kme3 and are required for appropriate enrichment of Cfp1 at NMIs region. The DNA-binding domain (CXXC) helps in binding at unmethylated CGIs. The SID helps to interact with Set1. Acidic, basic, and coiled-coil domains contribute cooperative signals for the speckled nuclear distribution of Cfp1

1.7.2 Cfp1 and Set1 interaction

Cfp1 plays a vital role in the stability of Setd1a (Tate et al., 2010). Loss of Cfp1 function in mice ESCs (*Cfp1*$^{-/-}$) results in a significant decrease in the expression levels of *Setd1a*, which can be restored by the expression of the full length or N/C-terminal domains of *Cfp1* (Tate et al., 2010). SID, which lies between the conserved basic and coiled-coil domains of Cfp1, is required for the interaction of Cfp1 with Setd1a/Setd1b. However, the SID domain alone is not sufficient for the interaction and requires the C-terminal end with the PHD2 domain for the interaction (Butler et al., 2008). Mutations in the SID domain such as a single alanine substitution at cysteine 375 (C375A) or a triple alanine substitution at tyrosine 390, cysteine 391, and serine 392 (YCS→AAA), ablates the interaction with the Setd1 complexes. Similarly, a single alanine substitution at cysteine 580 (C580A) within the PHD domain also ablates the interaction with the Setd1 complexes (Butler et al., 2008).

Cfp1 is required for the localisation of the Setd1a at euchromatin as deletion of Cfp1 results in a significant increase in the localisation of Setd1a and H3K4me3 at heterochromatin (Tate et al., 2010). Only full-length expression of Cfp1 restores the localisation of Setd1a and H3K4me3 (Tate et al., 2010). Consistent with this, chromatin immunoprecipitation showed an ectopic gain of H3K4me3 in the *Cfp1*$^{-/-}$ ESCs (Clouaire et al., 2012). Loss of *Cfp1* in mice ESCs resulted in a major reduction on the occupancy of Set1 at TSS (Clouaire et al., 2012, Clouaire et al., 2014). Similarly, deletion of the Cfp1 orthologue in *D. melanogaster* results in a drastic reduction of dSet1 on the polytene chromosomes (Ardehali et al., 2011). These studies suggest that Cfp1 functions as an unmethylated CpG reader, which recruits the SET1/COMPASS complex and prevents inappropriate drift of Set1 from unmethylated CpG regions (**Figure 1.4**).

1.7.3 Cfp1 and cytosine methylation

Cfp1 also plays an essential role in cytosine methylation as *Cfp1*$^{-/-}$ ESCs exhibited a ~60% decrease in global cytosine methylation levels (Carlone et al., 2005).

Cytosine methylation deficiency was observed in all genes tested, including repetitive DNA elements, imprinted genes and single copy genes (Carlone et al., 2005). Similarly, deletion of the Cfp1 orthologue in zebrafish resulted in a 60% reduction of global cytosine methylation (Young et al., 2006). The deficiency in cytosine methylation in *Cfp1*$^{-/-}$ ESCs is rescued by the introduction of Cfp1 (Carlone et al., 2005).

Co-immunoprecipitation experiments indicate that Cfp1 physically interacts with DNA methyltransferase 1 (Dnmt1) (Butler et al., 2008). Dnmt1 can interact with Cfp1 through its N-terminal and C-terminal domains. The N-terminal (aa 169–493) protein fragment which contains a nuclear localisation signal (NLS) and a large portion of the targeting sequence (TS) that directs the Dnmt1 to heterochromatin region, mediates the interaction with Cfp1 (Butler et al., 2008). Several conserved domains of the Cfp1 mediate the interaction with the Dnmt1 (Butler et al., 2008). In murine *Cfp1*$^{-/-}$ ESCs, levels of the *Dnmt1* were reduced by 50% (Carlone et al., 2005). Similarly, in Cfp1 depleted oocytes, DNA methylation levels were significantly decreased (Yu et al., 2017). These findings suggest that Cfp1 plays an important role in DNA methylation.

1.7.4 Role of Cfp1 in development

Cfp1 plays an essential role in the development of an organism (Carlone and Skalnik, 2001, Carlone et al., 2005, Young et al., 2006). In mice, *Cfp1* depletion results in embryonic lethality between 3.5 – 6.5 dpc stage (Carlone and Skalnik, 2001). *Cfp1* deleted blastocysts appear normal, but these embryos fail to gastrulate (Carlone and Skalnik, 2001, Carlone et al., 2005). Surprisingly, ESCs that lack the *Cfp1* gene are viable, but unable to differentiate (Young and Skalnik, 2007). *Cfp1*$^{-/-}$ ESCs exhibited a longer doubling time compared to wild-type ESCs, and the number of apoptotic cells is significantly higher (Carlone et al., 2005). Similarly, RNAi knockdown of the *CFP1* in a human *PLB-985* myeloid cell line resulted in reduced clonal survival, reduced cell growth, and impaired cellular differentiation (Young and Skalnik, 2007). Conditional knockdown of the *Cfp1* in adult mice leads

to failure of haematopoiesis, a dramatic loss of lineage-committed progenitors, and death within 14 days (Chun et al., 2014). Furthermore, injection of the *cfp1* transcript-specific antisense morpholino oligonucleotides in zebrafish results in embryo runting, failure of primitive haematopoiesis, cardiac oedema, incomplete vasculature formation, increased apoptosis, and death (Young et al., 2006). These findings suggest that *Cfp1* is required for early and post-embryonic development.

1.8 Histone acetylation

Histone acetylation, one of the most common chromatin marks, plays an important role in gene regulation. Lysine residues on the N-terminus of histones are positively charged, and they interact with the negatively charged phosphate group of DNA (Bannister and Kouzarides, 2011). HATs acetylate the lysine residues in histones. Lysine acetylation by the HATs neutralises the positive charge of histones, thereby reducing the affinity of histones for DNA. As a consequence, the closed chromatin is transformed into an open state which can facilitate the transcription process (Carrozza et al., 2003). Acetylated histones can act as binding sites for various chromatin remodulators which regulate gene expression. Hyper-acetylation of histone is often associated with transcriptionally active genes whereas hypo-acetylation of histones is associated with transcriptionally silent genes (Marmorstein and Zhou, 2014).

1.8.1 HATs

Histone acetyltransferase 1 (HAT1) and the general control non-repressible 5 (Gcn5) HAT are the first HATs which were shown to acetylate histones (Kleff et al., 1995, Brownell et al., 1996). Since the isolation of the HAT1 and Gcn5, many other HATs have been identified. Some of these HATs, such as the PCAF and HAT1 have conserved sequences with the Gcn5 and many other HATs such as MYST (MOZ, Ybf2, Sas2, and Tip60), and CREB-binding protein (CBP)/p300 have limited sequences conserved (CBP)/p300 (Marmorstein and Zhou, 2014). Based on localisation, HATs are divided into two classes, Type A and Type B (Lee and

Workman, 2007, Chakravarty et al., 2014). Type A HATs are localised in a nucleus and are responsible for acetylation of nucleosomal histones. Type B HATs are localised in the cytoplasm and are responsible for the acetylation of free histones (Chakravarty et al., 2014). Based on the structure and sequence homology, HATs are classified into different families. Out of these families, four HAT families have been studied extensively, which includes Gcn5-related histone *N*-acetyltransferases (GNAT) family, p300/CBP family, MYST family and Rtt109 family (Marmorstein and Zhou, 2014). GNAT and MYST families are conserved from yeast to human but the p300/CBP family members are only found in metazoans, and the Rtt109 family is fungal-specific (Marmorstein and Zhou, 2014). There are other HATs subfamilies such as the steroid receptor coactivators (ACTR/AIB1, SRC1, TAF250, and CLOCK), whose HAT activities have not been studied as extensively as compared to the four major families (Spencer et al., 1997, Kawasaki et al., 2000, Mizzen et al., 1996, Doi et al., 2006).

The GNAT families consist of HATs such as PCAF, Hat1, Elp3, Hpa2, Hpa3, Nut1 and ATF-2 which have high structural and sequence similarity to Gcn5 (Roth et al., 2001, Lee and Workman, 2007). These HATs acetylate Histone H2B, H3 and H4 (Roth et al., 2001, Lee and Workman, 2007). MYST is the largest family and consists of KAT6A (MOZ), KAT6B (MORF), KAT7 (HBO1), KAT8h (MOF) and KAT5 (Tip60) (Marmorstein and Zhou, 2014, Avvakumov and Cote, 2007, Sapountzi and Cote, 2011, Roth et al., 2001). All of the MYST family members possess a common MYST domain, which is also a primary catalytic region (Marmorstein and Zhou, 2014). The MYST family can acetylate Histone H2B, H3 and H4 (Roth et al., 2001, Avvakumov and Cote, 2007). The p300/CBP HAT consists of p300 and CBP HATs which are present in higher eukaryotes (Chen et al., 1997). The p300 and CBP have high sequence homology and are often collectively called p300/CBP (Kalkhoven, 2004). The p300 and CBP acetylate histone H2A, H2B, H3 and H4 (Roth et al., 2001, Lee and Workman, 2007). Rtt109 is named for its initial identification as a regulator of Ty1 transposition gene product 109. Like the p300/CBP family, Rtt109 has limited sequence homology to other known HATs (Marmorstein and Zhou, 2014). The Rtt109 acetylates histone H3 (Han et al., 2007a, Han et al., 2007b). Rtt109 on its own has very low acetyltransferase activity and

requires one of two histone chaperone proteins, Vps75 or Asf1, for stimulated acetylation (Han et al., 2007a, Han et al., 2007b).

1.8.2 Histone deacetylases (HDACs)

HDACs remove the acetyl group from the histones which result in a compacted, transcriptionally silent closed chromatin structure. They are often associated with gene repression. In mammals, based on the structure, function and their localisation within the cells, HDACs are divided into four major classes (Class I, IIa, IIb, IV) (De Ruijter et al., 2003, Delcuve et al., 2012, Seto and Yoshida, 2014). The Class I HDACs are homologues of yeast Rpd3, a founding member of yeast HDACs (Taunton et al., 1996). Class I contains HDACI, HDAC2, HDAC3, and HDAC8 and are expressed ubiquitously (Seto and Yoshida, 2014). These HDACs are predominantly localised in the nucleus. The HDAC1 and HDAC2 are 85% identical in their structure (Gregoretti et al., 2004). Even though they share a high degree of homology, they have some non-redundant function. For example, deletion of the HDAC1 homologue in mice results in embryonic lethality but mice lacking the HDAC2 homologue survive until the prenatal period (Lagger et al., 2002, Zhang et al., 2014, Zupkovitz et al., 2006). HDAC1 and HDAC2 can form homo- and hetero-dimers which are required for HDAC activity (Taplick et al., 2001, Luo et al., 2009, Brunmeir et al., 2009).

HDAC1 and HDAC2 are generally found as a part of a multi-protein complex, including Sin3, nucleosome remodelling and histone deacetylase (NuRD) and Corepressor of REST (CoResT) complexes (Yang and Seto, 2003, De Ruijter et al., 2003, Yang and Seto, 2008). As shown in **Figure 1.6**, the Sin3 complex contains HDAC1/HDAC2, Sin3A/B, SAP18, SAP30, retinoblastoma-associated protein 46 and 48 (RbAp46 and RbAp48) (Silverstein and Ekwall, 2005, Hayakawa and Nakayama, 2011, Kelly and Cowley, 2013). The Sin3 complex plays an important role in nucleosome remodelling, histone methylation, DNA methylation and N-acetylglucosamine transferase activity (Silverstein and Ekwall, 2005, Hayakawa and Nakayama, 2011, Kelly and Cowley, 2013). The composition of the NuRD complex

varies depending upon external stimuli and cell types. The NuRD complex contains HDAC activity and ATP-dependent chromatin remodelling activity (**Figure 1.6**) (Delcuve et al., 2012). The HDAC activity is carried out by HDAC1 and HDAC2, whereas the ATP-dependent chromatin remodelling is carried out by the Mi-2α and or Mi-2β (Delcuve et al., 2012). The other important subunits of the NuRD complex are p66α or p66β, metastasis-associated protein family (MTA) MTA1, MTA2 or MTA3, RbAp46/RbAp48 and the methyl-CpG-binding domain-containing proteins (MBD2 or MBD3) (Spensberger et al., 2008, Denslow and Wade, 2007, Hayakawa and Nakayama, 2011). Similar to the SIN3 and NuRD complex, the CoREST complex contains HDAC1 and HDAC2, and other subunits such as KDM1/LSD1 demethylase, CoREST and BHC80, a PHD domain-containing protein (**Figure 1.6**) (Delcuve et al., 2012).

HDAC3 is found in two highly related complexes, namely the silencing mediator of retinoic acid and thyroid hormone receptor (SMRT or NCOR2), and nuclear receptor corepressor (NCoR or NCOR1) complexes (Hayakawa and Nakayama, 2011, Kelly and Cowley, 2013, Perissi et al., 2010). These complexes consist of HDAC3, SMRT, NCOR1, G-protein pathway suppressor 2 (GPS2), TBL-related 1 (TBLR1) and transducin β-like 1 (TBL1). HDAC8, unlike other class I HDACS, functions independently and is not associated with any known complexes (Delcuve et al., 2012).

The Class II HDACs includes HDAC4, HDAC5, HDAC6, HDAC7, HDAC9 and HDAC10 and are expressed in both the cytoplasm and nucleus (Delcuve et al., 2012). In response to the signal, they exhibit cell type-specific expression and shuttle between the cytoplasm and nucleus (Majdzadeh et al., 2008, Lu et al., 2000). The HDAC6 and HDAC10 contain two catalytic sites and are sometimes kept in a separate subclass (Class IIb). Other remaining HDACs (HDAC4, HDAC5, HDAC7 and HDAC9) are kept in Class IIa (Verdin et al., 2003, De Ruijter et al., 2003). In contrast to other HDACs, expression of the class IIa HDACs is restricted. For example, HDAC5 and HDAC9 are highly enriched in heart, muscle and brain, the HDAC4 is highly expressed in skeleton and brain, and the HDAC7 is expressed in thymocytes 34 (T-cell precursors derived from the thymus) and endothelial cells

(Zhang et al., 2002, Chang et al., 2004, Chang et al., 2006). HDAC6 is the main cytoplasmic deacetylase, and targets proteins such as cortactin, transmembrane proteins such as the interferon receptor IFNαR, and chaperones such as HSP90 (Haberland et al., 2009).

Class III HDACs (Sirtuins) are homologues of the yeast silent information regulator 2 (Sir2) and are structurally different from other classes of HDACs. Class III HDACs depend on the NAD^+ for deacetylation, and thus the deacetylase activity is controlled by the [NAD^+]/[NADH] ratio (Dai and Faller, 2008). There are seven members in this class *viz*., SIRT1—SIRT7. Among the seven members, as for the function/structure studies, SIRT1 which is also called as Sir2α, the closest orthologue of yeast Sir2, has been the most extensively studied (Seto and Yoshida, 2014, Schwer and Verdin, 2008, Dai and Faller, 2008, Frye, 2000). Sirtuins are ubiquitously present and are known to deacetylate histones H1 lysine 26 (H1K26Ac), H3 lysine 9 (H3K9Ac) and H4 lysine 16 (H4K16Ac) (Liu et al., 2009). They also deacetylate non-histone substrates such as forkhead box type O transcription factors (FOXO), p53, peroxisome proliferator activated receptor-γ (PPARγ), and nuclear factor-κB (NFκB) (Dai and Faller, 2008, Liu et al., 2009).

In the class IV family, there is only one member, HDAC11. HDAC11 shares homology with the catalytic domain of Class I and II HDACs (Gao et al., 2002). HDAC11 is enriched in the heart, kidney, brain, muscle and testis. The localisation of HDAC11 varies according to cell types and/or environmental cues (Gao et al., 2002, Yanginlar and Logie, 2018).

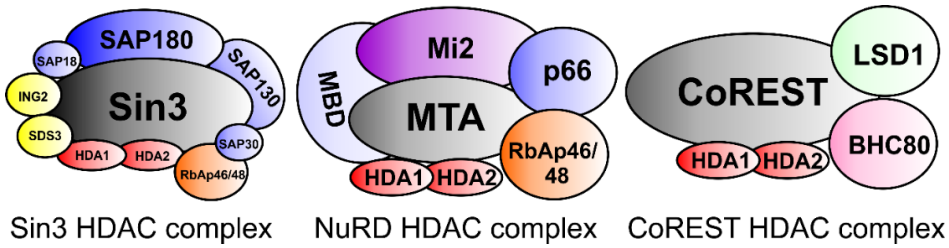


Figure 1.6: HDAC1/HDAC2 multi-protein complexes

The Sin3 complex contains HDAC1/HDAC2, Sin3A/B, SAP18, SAP30, retinoblastoma-associated protein 46 and 48 (RbAp46 and RbAp48). The NuRD complex contains HDAC1/HDAC2 p66α or p66β, MTA, RbAp46/RbAp48, MBD and Mi2. The CoREST complex contains HDAC1 /HDAC2, KDM1/LSD1 demethylase, CoREST, BHC80, a PHD domain-containing protein.

1.9 Crosstalk between H3K4me3 and acetylation

The interplay between the highly dynamic histone acetylation and H3K4me3 modifications plays an essential role in shaping the structure and function of chromatin (Zhang et al., 2015, Clouaire et al., 2014). Previous studies suggest that H3K4me3 present at the promoter regions helps in the recruitment of the HATs to facilitate the transcription process. For example, SGF29, a subunit of the SAGA HAT complex contains a tudor domain that binds to H3K4me3 and helps to recruit the SAGA HAT complex (Bian et al., 2011). Yng1, a component of the NuA3 HAT complex, contains a PHD domain that binds H3K4me3 and promotes the NuA3 HAT activity (Taverna et al., 2006, Doyon et al., 2004). Similarly, mammalian HBO1 histone acetyltransferase complex interacts with the H3K4me3 via its native subunit ING4 inhibitor of growth 4 (ING4) to acetylate H3 (Hung et al., 2009). The p300 (HAT) promotes the association of COMPASS complex at the p21/waf-1 (p53 regulated gene) promoter (Tang et al., 2013). In ESCs, a lack of H3K4me3 results in a reduced recruitment of the Gcn5 at *CDKN1A* promoters leading to a reduction in histone 3 lysine 9 (H3K9) acetylation and H3K4 methylation (Clouaire et al., 2014). MOZ and MLL coordinate with each other for the expression of the HOX genes (*HOXA5*, *HOXA7*, and *HOXA9)* in human cord blood CD34+ cells (Paggetti et al., 2010).

Different studies have suggested that similar to HATs, HDAC complexes also contain H3K4me3 binding domains (Doyon et al., 2006, Margaritis et al., 2012). A recent finding has indicated that pre-existing H3K4me3 marks on histones, promotes dynamic turnover of H3 acetylation, which is mediated by the combined action of HAT p300/CBP and HDACs (Crump et al., 2011). These findings suggest that the H3K4me3 could act as an epigenetic mark at the promoter regions to facilitate the dynamic turnover of acetylation modification, which plays a key role in transcriptional regulation.

1.10 *Caenorhabditis elegans:* a powerful genetic model

C. elegans is a multicellular genetic model used to understand various biological events. *C. elegans* is a free-living non-parasitic nematode and is small in size (1 millimetre long). *C. elegans* is easy to grow, they feed on bacteria and have a very short lifecycle (**Figure 1.7**) (Corsi, 2006, Brenner, 1974). *C. elegans* has skin, guts, neurons, muscles, gonads, and other tissues which are similar in function to those of other animals (Corsi et al., 2015). Thus, it can be used as a model to understand the genetic regulatory circuit present in these tissues. It is transparent in nature, with the help of Nomarski (differential interference contrast, DIC) optics, individual cells and subcellular details are easily visible. This allowed for the visualisation of development and differentiation process at a single cell resolution. Additionally, with the use of different fluorescence markers, numerous biological events can be easily studied (Corsi, 2006). *C. elegans* are hermaphrodites which can self-fertilise, and males, which only comprise ~0.02% of the total natural population (Sulston and Horvitz, 1977, Kimble and Hirsh, 1979, Fay, 2006). Therefore, it is easy to maintain the mutants over generations and also to create desired double-, triple- or more mutants by crossing the mutants with each other (Corsi, 2006). Another key feature of *C. elegans* is that it is amenable to RNA interference (RNAi). Thus any gene can be easily knockdown by RNAi feeding (Fire et al., 1998). This makes *C. elegans* an ideal model to examine the function of genes and knockdown phenotypes associated with the genes. Additionally, genetic screens are used in *C. elegans* to identify the gene function. It can be used to discover the genes responsible for a specific phenotype (forward genetics) or conversely, the gene can be modified to analyse the phenotypic effects (reverse genetics) (Sin et al., 2014, Kutscher and Shaham, 2014). With the help of forward genetics and reverse genetics in *C. elegans*, many genes involved in the development and biological process have been identified (Corsi, 2006).

Importantly, the function of the COMPASS complex is conserved in *C. elegans.* It contains two COMPASS complexes, (i) the SET-2/COMPASS, which is the orthologue of yeast Set1/COMPASS, and (ii) the SET-16/COMPASS, which is the orthologue of mammalian MLL3/4/COMPASS (Fisher et al., 2010, Simonet et al.,

2007). The SET-2/COMPASS is responsible for the majority of H3K4me3 deposition. CFP-1 is a key subunit of the Set-2/COMPASS complex and plays an important role in the development of an organism. However, how it contributes to gene regulation and development is not yet clear. I herein used *C. elegans* as a model to investigate the contribution of CFP-1 towards H3K4me3 deposition and its effect on gene regulation, and development of an organism. Furthermore, I used vulval development in *C. elegans* as a system to investigate the role of CFP-1 in cell fate specification and to identify its novel regulators.

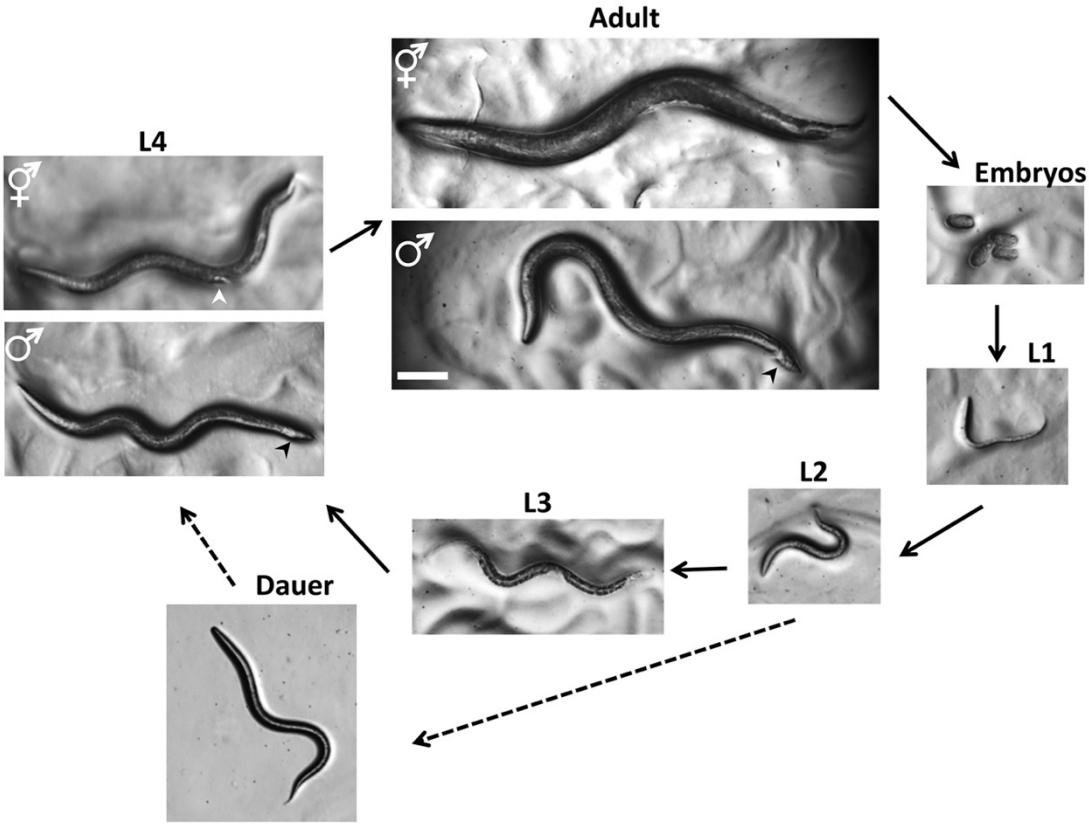


Figure 1.7: Life Cycle of *C. elegans*

C. elegans life cycle is comprised of the embryonic stage, four larval stages (L1, L2, L3 and L4) and adult stage. If the environmental conditions are not favourable, worms may enter the dauer stage at the end of the L2 stage. *C. elegans* has two sexes, male and hermaphrodite. Adult hermaphrodites have wider girth and a tapered tail whereas the adult male has slimmer girth and a fan-shaped tail. Males and hermaphrodites can be easily distinguished at the L4 stage. The picture is taken from Corsi et al. (2015).

1.11 Vulval development in *C. elegans*

The vulval development in *C. elegans* is an excellent system to study cell-fate specification, organogenesis events and signal transduction. *C. elegans* has six vulval precursor cells (VPCs) P(3-8).p, also called Pn.p cells. These cells have an ability to give rise to the vulva (Sternberg and Horvitz, 1986, Sulston and Horvitz, 1977). However, only three (P5.p, P6.p and P7.p) out of six adopt the vulval fate naturally and the other three (P3.p, P4.p and P8.p) fuse with the hypodermis (**Figure 1.8**). At the beginning of the third larval stage, Ras-mediated inducing signal (LIN-3; a ligand of Ras signalling pathway) from the anchor cell (AC) causes P5.p, P6.p and P7.p cells to adopt the vulval cell fate. P(5-7).p cells divide further and gives rise to 22 descendants which form the vulva (Sulston and Horvitz, 1977). The P(3,4,8).p cells which are far from the AC do not receive sufficient inductive signal to adopt the vulval cell fate (**Figure 1.8**). These cells divide once and fuse with the hypodermis. Even though the P(3,4,8).p cells fuse with the hypodermis, different studies have suggested that they can adopt the vulval cell fate (Thomas et al., 1990, Kimble, 1981). The ability of Pn.p cells to adopt the vulva fate is depended upon the inductive signal from the AC. In wild-type worms, P(5-7).p cells are near to the AC cells and receive strong inductive signals to adopt the vulva cell fate, but if the P(5-7).p cells are removed by laser ablation then the P(3,4,8).p cells will adopt the vulval fate (Sulston and White, 1980). Conversely when the AC is removed by laser ablation, then all the VPC cells adopt the non-vulval fates (Kimble, 1981, Sulston and White, 1980).

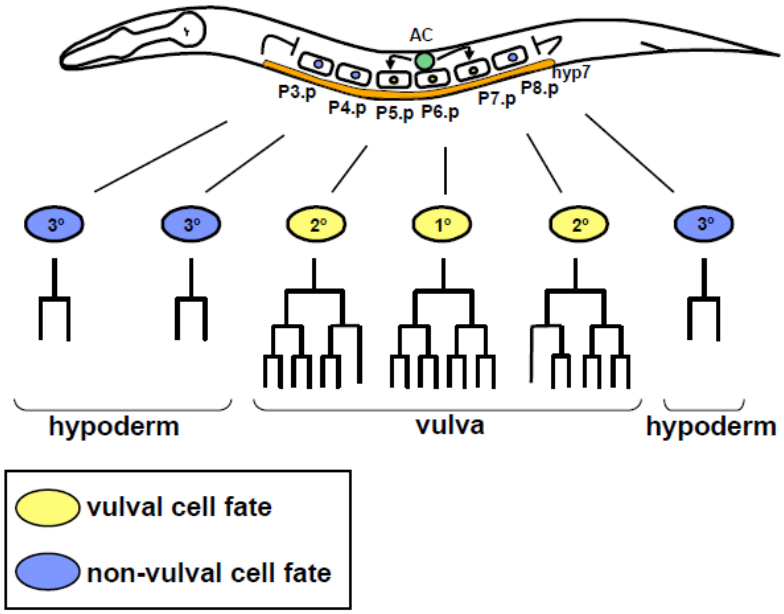


Figure 1.8: Vulva development at the third larval stage

Inductive signals from the anchor cell (green) induced the vulval cell fate (yellow) in the P6.p, P7.p and P8.p cells. Cells not receiving the inductive signal adopt the non-vulval cell fates (blue) and fuse with the hyp-7 (orange). The figure is taken from Andersen (2007).

1.11.1 Receptor Tyrosine Kinase (RTK)/Ras GTPase/MAP kinase (MAPK), Notch and Wnt signalling promotes vulval development

Mutants defective in vulval cell fate specification result in either a vulvaless (Vul) phenotype or a multivulva (Muv) phenotype (Horvitz and Sulston, 1980, Cui et al., 2006b, Harrison et al., 2006). If none of the VPC cells adopts a vulval fate, then the worms display Vul phenotype, whereas if all six VPC cells express vulval fates, then the worms display Muv phenotype. Three signalling pathways (Notch, receptor tyrosine kinase (RTK)/Ras/Mitogen-activated protein kinase (MAPK) and Wnt) interact with each other to direct vulval cell-fate specification (Cui et al., 2006b, Harrison et al., 2006).

A mutation that causes the loss of function in components of the RTK/Ras/MAPK pathway causes none of the six VPCs to adopt the vulval cell fate and display a vulvaless (Vul) phenotype (Aroian et al., 1990, Beitel et al., 1990, Han et al., 1990, Hill and Sternberg, 1992). Mutations that cause a gain of function in components of the RTK/Ras/MAPK pathway cause more than three cells to adopt the vulval cell fate resulting in the formation of ectopic pseudovulvae, the Muv phenotype (Aroian et al., 1990, Beitel et al., 1990, Han et al., 1990, Hill and Sternberg, 1992, Katz et al., 1996). RTK/Ras/MAPK signals from AC specify the fate of P5.p, P6.p and P7.p cells. The ACs express LIN-3, an epidermal growth factor (EGF)-like growth factor, which acts as an inductive signal to regulate the vulval development (Hill and Sternberg, 1992). LIN-3 is found in either membrane-bound or secreted forms. The AC mostly uses secreted LIN-3 to control vulval development. However, it can also use membrane-bound LIN-3 for vulval induction (Thomas et al., 1990, Aroian et al., 1990). As shown in **Figure 1.9**, LIN-3 binds to the LET-23 EGFR in the VPCs to specify the vulval fate (Koga and Ohshima, 1995, Simske and Kim, 1995).

The LET-23 epidermal growth factor receptor (EGFR) forms a complex with LIN-2 (membrane-associated guanylate kinase (MAGUK) protein), LIN-7 (PDZ domain-containing proteins) and LIN-10 (PDZ domain-containing proteins) for proper localisation at the basolateral side of the cell junction (**Figure 1.9**) (Hoskins et al., 1996, Simske et al., 1996, Whitfield et al., 1999). Mutation in *lin-2*, *lin-7*, or *lin-10*

results in mislocalisation of the LET-23 (Hoskins et al., 1996, Simske et al., 1996, Whitfield et al., 1999).

The LET-23 EGFR, LIN-2, LIN-7 and LIN-10 complex is required to receive the LIN-3 inductive signal from the anchor cell (Kaech et al., 1998). Once LIN-3 binds to the LET-23, LET-23 dimerises and autophosphorylates its C-terminal regions. These phosphorylated tyrosines act as a docking site for the adapter protein SEM-5(Grb2) (Clark et al., 1992). Docked SEM-5(Grb2) recruits SOS-1/LET-341, a Guanine Nucleotide Exchange Factor, and subsequently activates the LET-60/Ras (**Figure 1.9**) (Clark et al., 1992, Chang et al., 2000). Activated Ras activates the serine/threonine kinase LIN-45/RAF, which in turn activates the MAPK cascade by phosphorylating the serine/threonine kinase MEK. MEK phosphorylates and activates the MAP kinase homolog MPK-1/SUR-1 (Han et al., 1993, Kornfeld et al., 1995, Lackner et al., 1994, Wu and Han, 1994, Wu et al., 1995). Two TFs LIN-1 and LIN-31 regulate vulval development function directly downstream of the RTK/Ras/MAPK signalling pathway. LIN-1, an E26 transformation-specific TF, negatively regulates the vulval fate and loss of function mutation in the *lin-1* gene results in the Muv phenotype (Beitel et al., 1995). While LIN-31, a winged helix TF, acts as both a positive and negative regulator and loss of function of the *lin-31* gene leads to a mix of the Muv and Vul phenotypes (Miller et al., 1993). In addition to the key genes of the RTK/Ras/MAPK signalling pathway, there are many other positive and negative regulators of Ras signalling in *C. elegans* (Sundaram, 2006).

Similar to RTK/Ras/MAPK signalling pathway, Notch signalling also plays a crucial role in vulval cell fate specification. Delta/Serrate/LAG-2 (DSL) are ligands that bind to LIN-12*,* Notch receptor, to activate the Notch signalling pathway (Henderson et al., 1994, Chen and Greenwald, 2004). The APX-1 and LAG-2 are transmembrane ligands for the Notch receptor LIN-12, whereas the DSL-1 is a secreted ligand for the LIN-12 (Henderson et al., 1994, Tax and Thomas, 1994, Mello et al., 1994). Notch signalling is required for the specification of the AC (Lai, 2004, Greenwald, 1998). Cell-cell interactions mediated by the Notch signalling determine the AC and ventral uterine (VU) cell fate. Gain of function mutations of *lin-12* cause two VU to be specified whereas loss of function mutations of *lin-12*

cause two AC formation (Lai, 2004, Greenwald, 1998). In the VPC cells, Notch signalling promotes secondary (2º) vulval fate (Sternberg, 2005). Gain in the function of Notch signaling causes ectopic 2º cells, and the reduction in Notch signaling causes no 2º cells to be specified (Sternberg, 2005). The primary (1º) vulval cell expresses the Notch signaling ligands to promote the 2º vulval fate in P5.p and P7.p cells (Sternberg, 2005, Yoo et al., 2004). Once the Notch ligand binds to LIN-12, LIN-12 gets activated. After activation, LIN-12 is cleaved in the membrane, and the cytoplasmic domain of LIN-12 binds to CBF1/Su(H)/LAG-1 (CSL) transcription cofactor LAG-1, which promotes the expression of 2º vulval fates (Yoo et al., 2004, Sternberg, 2005, Christensen et al., 1996).

In addition to RTK/Ras/MAPK signalling and Notch signalling, Wnt signalling is also required for vulval cell fate specification. The Wnt signalling promotes the adoption of 1º and 2º vulval fate by preventing the fusion with the hypodermal syncytium hyp7 (Sternberg, 2005). Wnt signalling is also required for the asymmetric cell divisions of 2º cell descendants (Sternberg, 2005, He, 2004, Inoue et al., 2004).

1.11.2 Determination of Vulval fate

During vulval development, among the three VPC cells, P6.p adopts 1° fate, and P5.p and P7.p adopt 2° fate. The determination of 1° and 2° fate is regulated by crosstalk between the RTK/Ras/MAPK and Notch signalling pathways (Chen and Greenwald, 2004, Shaye and Greenwald, 2002, Yoo et al., 2004). Ras signalling is mostly responsible for the 1° fate adopted by P6.p, whereas a combination of Ras signalling and Notch signalling plays an important role in the adoption of 2° fate by P5.p and P7.p (**Figure 1.9**) (Katz et al., 1995).

Among the three VPCs, P6.p is closest to the AC, it receives the strongest LIN-3 signal and adopts the 1° fate (Katz et al., 1995). P5.p and P7.p cells are further from the AC and receive a lesser LIN-3 signal. The P6.p cell subsequently signals to P5.p and P7.p cells to adopt the 2° fate (**Figure 1.9**). The P6.p cell signals to adjacent

VPCs via APX-1, LAG-2 or DSL-1 (Simske and Kim, 1995, Yoo et al., 2004). LAG-2 is present in all VPCs before induction and its expression increases in the P6.p cell. On the other hand, expression of the *apx-1* and *dsl-1* is initiated by the P6.p cell to inhibit neighbouring cells to adopt the 1° fate, thereby inducing them to adopt the 2° fate (Chen and Greenwald, 2004, Wilkinson et al., 1994). The Notch signalling in neighbouring cells results in the activation of a number of negative regulators of Ras signalling such as the MAP kinase phosphatase LIP-1, the tyrosine kinase ARK-1 and the *lst* genes (Yoo et al., 2004, Chen and Greenwald, 2004, Wilkinson et al., 1994). In the P6.p cell, Ras signalling downregulates the Notch signalling pathway to maintain the 1° fate. In mosaic *C. elegans* in which both P5.p and P7.p cells have lost *let-23* but P6.p has wild-type *let-23*, P5.p and P7.p can still adopt the 2° cell fate (Koga and Ohshima, 1995, Simske and Kim, 1995). These studies suggest that Notch signalling alone, without Ras signalling, can specify the 2° fate of VPCs.

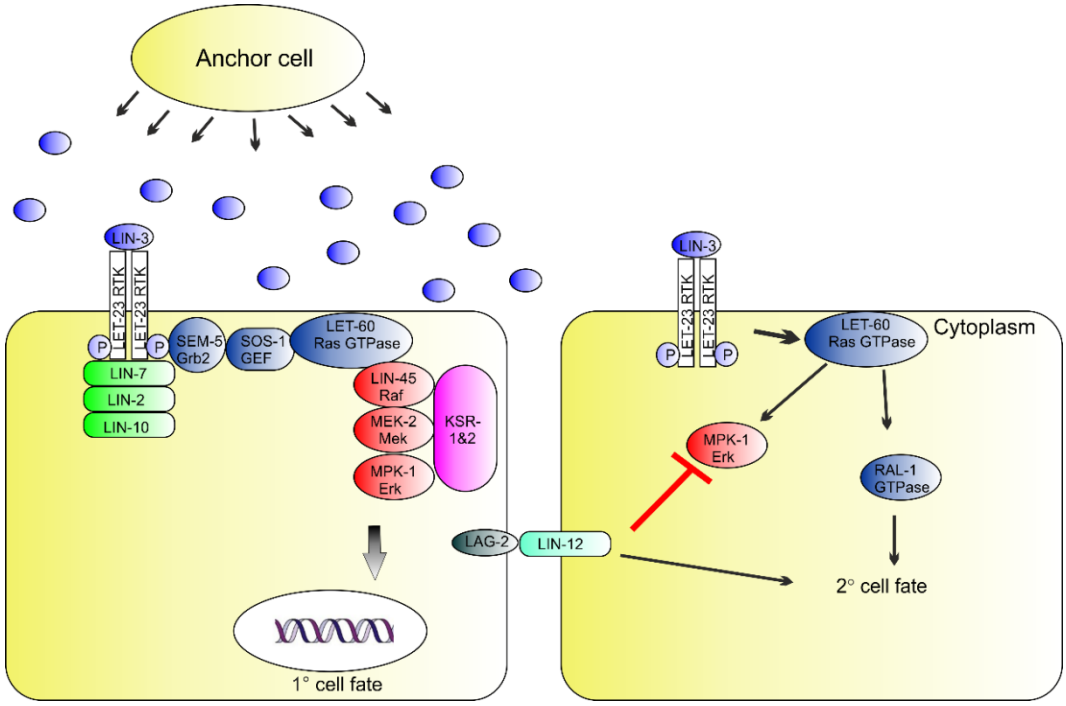


Figure 1.9: Determination of Vulval fate

The anchor cells release LIN-3 EGF-like ligands, which bind to the LET-23 and activates the RTK/Ras/MAPK signalling pathway to specify the vulval cell fates. The LIN-2/LIN-7/LIN-10 complex mediates basolateral localisation of the LET-23 EGFR and helps in the activation of the pathway. P6.p cell which is near to the anchor cell receives strong *lin-3* signal and adopts the 1°fate. The two adjacent cells (P5.p and P7.p) receive less *lin-3* signals and an inhibitory Notch signalling from the P6.p cell. Activation of the Notch signalling in (P5.p and P7.p) cells downregulates MPK-1 MAPK thereby, inducing them to adopt the 2° fate. The figure is adapted from Gauthier and Rocheleau (2017)

1.11.3 Synthetic Multivulva (SynMuv) genes

Excessive RTK/Ras/MAPK signalling can cause a Muv phenotype in which all VPCs adopt vulval fates. To prevent the aberrant development of ectopic vulva at the ventral surface of the worms, a set of genes called SynMuv genes antagonise the activity of the RTK/Ras/MAPK signalling (Ferguson and Horvitz, 1989, Lu and Horvitz, 1998). The SynMuv genes are classified into three main classes *viz*., A, B and C. Worms with a single mutation in any class of SynMuv genes have normal vulval development, but a combination of mutation of SynMuv genes from two different classes (AB, AC, and BC) exhibit the SynMuv phenotype (Ferguson and Horvitz, 1989, Fay and Yochem, 2007). Worms with two class A SynMuv mutations or two class B SynMuv mutations or two class C SynMuv mutations produce normal vulva (Fay and Yochem, 2007). This suggests that the SynMuv genes within each class function together, either in the same pathway or in a molecular complex, whereas the SynMuv genes in different classes are functionally redundant and act in parallel pathways. There are four class A SynMuv genes, 24 class B SynMuv genes and four class C SynMuv genes (Fay and Yochem, 2007). Additionally, three new genes, *smo-1*, *uba-2* and *ubc-9* have been identified, which, when mutated, can cause the Muv phenotype in class A, B, or C SynMuv genes (Fay and Yochem, 2007, Poulin et al., 2005). *smo-1* encodes for SUMO enzyme which is responsible for SUMOylation, *uba-2* encodes for SUMO activating enzyme and *ubc-9* encodes for a SUMO conjugating enzyme (Fay and Yochem, 2007).

In SynMuv class A, there are four genes *lin-8, lin-15a, lin-38* and *lin56,* which code for putative TFs. *lin-8* encodes for a novel acidic nuclear protein that interacts with the *lin-35* SynMuv gene (Davison et al., 2005). *lin-38* encodes a C2H2-type zinc-finger protein (Cayrol et al., 2007, Dejosez et al., 2008, Clark et al., 1994, Huang et al., 1994). *lin-56* and *lin-15A* encode THAP domain proteins which repress the transcription of many genes potentially by aiding in the recruitment of corepressor proteins (Cayrol et al., 2007, Dejosez et al., 2008, Clark et al., 1994, Huang et al., 1994). LIN-56 and LIN-15A interact with each other and are likely to function in a complex together (Davison et al., 2011).

The SynMuv Class B genes encode for TFs and chromatin-modifying complexes that play an important role in transcriptional regulation. The Class B SynMuv genes encode for the components of chromatin modulators such as *met-2* (histone H3 lysine 9 histone methyltransferases) and *hpl-2 (*heterochromatin protein 2) which are responsible for gene repression (Lu and Horvitz, 1998, Couteau et al., 2002, Coustham et al., 2006, Poulin et al., 2005). It also contains genes such as *let-418* (homologue of Mi-2) which are key components of the NuRD complex (Tong et al., 1998, Zhang et al., 1998).

Class B genes also include TFs which bind to specific DNA sequences to regulate gene expression (Korenjak et al., 2004, Lewis et al., 2004, Fay and Yochem, 2007). *lin-35* (orthologue of mammalian retinoblastoma protein Rb), *dpl-1* (homologue of DP), *efl-1*, *lin-9*, *lin-37*, *lin-52*, *lin-53* and *lin-54* encode for the components of the DREAM (DP, Retinoblastoma [Rb]-like, E2F, and MuvB) complex (Lu and Horvitz, 1998, Lipsick, 2004, Ceol and Horvitz, 2001). The DREAM complex is conserved in mammals, *D. melanogaster* and *C. elegans*, and plays a repressive role in gene expression (Korenjak et al., 2004, Lewis et al., 2004, Fay and Yochem, 2007). Some of the class B genes such as *lin-13*, *lin-15B*, and *lin-36* do not have obvious homologs in mammals (Huang et al., 1994, Clark et al., 1994, Melendez and Greenwald, 2000, Reddy and Villeneuve, 2004, Thomas et al., 1990).

Similar to class A and B, class C SynMuv genes are also involved in chromatin regulation. There are four genes in class C, *epc-1*, *mys-1*, *ssl-1* and *trr-1,* and they encode for the component of the TIP60 HAT chromatin remodelling complex (Fay and Yochem, 2007, Ceol and Horvitz, 2004, Ceol et al., 2006, Thomas et al., 2003). Sometimes the class C genes are considered to be a subset of the class B genes as they form only a weak Muv phenotype with the class B genes and a strong Muv phenotype with class A genes (Ceol and Horvitz, 2004).

1.12 Outline of the thesis

Previous studies have indicated the importance of Cfp1 in ESC differentiation and cell fate specification. However, early embryonic death due to loss of Cfp1 function in mammals prevented an assessment of the Cfp1 role during development. To investigate the function of Cfp1, a *cfp-1(tm6369)* loss-of-function *C. elegans* mutant was used in this study.

Chapter 3 of this thesis describes the role of CFP-1 in *C. elegans* development as investigated by using the *cfp-1(tm6369)* allele. The H3K4me3 levels were measured in *cfp-1(tm6369)* mutants and compared to those reported for a *set-2(bn129)* loss of function mutant. It was observed that in *cfp-1(tm6369)* mutants, H3K4me3 levels were significantly reduced and the levels of H3K4me3 were similar to those in a *set-2(bn129)* mutant. To investigate the functional consequences of loss of function of *cfp-1* in *C. elegans* development, fertility was assayed at 15 °C, 20 °C and 25 °C, and growth was assayed at 60 h. It was found that the fertility of both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants was significantly reduced at 15 °C, 20 °C and 25 °C. Furthermore, it was observed that in both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants development was delayed compared to the wild-type. These findings suggested that both the CFP-1 and the SET-2 are required for the fertility and development of *C. elegans*.

Chapter 4 of the thesis describes the role of CFP-1 in stress-induced gene regulation as investigated by using a salt stress reporter assay and measurement of stress-inducible gene expression by quantitative RT-PCR. It was observed that the induction of the stress-inducible reporter gene was significantly higher in both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants. Furthermore, the role of CFP-1 during stress-dependent gene induction was confirmed by measuring the endogenous gene expression of the salt-inducible and heat shock genes. It was observed that after induction the expression of the salt-inducible gene and heat shock genes was significantly higher in *cfp-1(tm6369)* and *set-2(bn129)* mutants. These findings suggested a potential role of CFP-1 and SET-2 in stress-dependent gene induction.

Chapter 5 of this thesis describes the function of CFP-1 in organogenesis as investigated by using vulval development in *C. elegans* as a model. It was observed that only loss of function of CFP-1, but not SET-2 could suppress the SynMuv phenotype. A mini RNAi screen was conducted to identify genes that either enhance or suppress the function of CFP-1 in the SynMuv phenotype. *sin-3*, *hda-1* and *gtbp-1* genes were found to enhance the function of *cfp-1* whereas *pbrm-1* was found to suppress the function of *cfp-1* in SynMuv development. The data presented in this Chapter suggests *cfp-1* interacts genetically with other chromatin modulators such as *hda-1* and *sin-3* to regulate the vulval development.

Chapter 6 of the thesis describes the genetic interaction between CFP-1 and HDACs as assayed using a fertility assay. RNAi knockdown of *hda-1, hda-2, hda-3* and key subunits of HDAC1/2 complexes significantly reduced the brood size of *cfp-1(tm6369)* mutants. In contrast, the brood size of *set-2(bn129)* mutants was not significantly affected. Furthermore, it was observed that fertility of *cfp-1(tm6369)* mutants was significantly reduced when treated with Trichostatin A, a class I/II HDAC inhibitor. Taken together, findings in this Chapter suggest that that CFP-1, but not SET-2 interacts with HDAC1/2 complexes.

Chapter 7 summarises the findings of this project. It also discusses the contribution of this work to the current understanding of the CFP-1 role in gene regulation and development, as well as potential further study.

Chapter 2

Materials and Methods

2.1 Strains and their Maintenance

C. elegans strains were maintained at 20 °C under standard growth condition unless otherwise stated. They were grown on *Escherichia coli* OP50 seeded nematode growth medium (NGM) petri plates (Brenner, 1974). Strains used in this study are listed in Table 2.1.

Table 2.1 List of strains used in this study

Table showing the strain name, genotype of strains and number of times they have been outcrossed. There is no strain name yet associated with the strains indicated with a star *.

Strain name	Genotype	Outcrossed
KC565	*sin-3(tm1276) I; him-5(e1490)* V	4
VC1815	*spr-1(ok2144)* V	4
PS1839	*let-23(sa62)* II	Not outcrossed
VP198	*kbIS5[gpdh-1p::GFP + rol-6(su1006)]*	4
MT2124	*let-60(n1046)* IV	Not outcrossed
MT1035	*lin-12(n137n460)* III	Not outcrossed
MT8189	*lin-15B&lin-15A (n765)* X	5
MH2430	*cbp-1 (ku258) III*	Not outcrossed
***cfp-1* mutant ***	*cfp-1(tm6369)* IV	5
***set-2* mutant ***	*set-2(bn129)* III	5
***mys-4* mutant ***	*mys-4(tm3161)* I	5
N2 Bristol	N2 (wild-type hermaphrodite)	
N2 Bristol	N2 (wild-type male)	

2.1.1 Preparation of NGM

To prepare NGM, agar, peptone, NaCl and water were mixed (see section 2.21.1) and autoclaved. The mixture was cooled down to 55 °C, and sterile cholesterol, $MgSO_4$, $CaCl_2$ and KPO_4 were added. The mixture was poured into a petri dish using a sterile peristaltic pump under a sterile hood. Solidified plates were turned upside down and dried for a week at room temperature. OP50 bacteria culture was spotted on the plates which were then incubated at room temperature for 3 days before using.

2.2 Genotyping by single-worm PCR

Single-worm PCR was done to confirm the genotype of wild-type, single and double mutants used in this experiment. For single-worm PCR, worm gDNA was extracted by lysing the worm in 10 µl of PK worm-lysis buffer using the "PK digestion program" on a thermocycler, as shown in **Table 2.2**. Once the digestion cycle was completed, 2.5 µl of the extracted gDNA template was mixed with PCR mixture (5 µl of Phusion HF buffer, 0.5 µl 20 mM dNTP (Applied Biosystems), 1.25 µl each of 10 mM reverse and forward primer (Integrated DNA Technologies) (**Table 2.3**) and 14.75 µl of nuclease-free H_2O). The PCR program used is shown in **Table 2.4**.

For analysis, 1.2% agarose gel, containing Ethidium bromide (Invitrogen), was prepared by dissolving 1.2 g agarose (SERVA) in 100 ml of 0.5 X Tris acetate and Ethylenediamine tetraacetic acid (EDTA) buffer (20 mM Tris acetate and 0.5 mM EDTA) (Severn Biotech, UK). The solution was heated in a microwave until the agarose was completely dissolved. Molten agar was then allowed to cool down to 50°C and 3 µl of 20 mg/ml Ethidium bromide was added. The gel was poured into gel trays and left at room temperature for 1 h to polymerise.

The PCR product was mixed with Orange G loading dye (50% glycerol, 0.05% orange G) and loaded into 1.2% agarose gel. The gel was then run on an electrophoretic tank containing 0.5 X Tris Borate EDTA buffer at 100 V for 20 min.

The gel was removed and visualized under the FastGene® FAS Digi imaging system, equipped with a Pentax MX 1 camera.

Table 2.2: PK digestion program

Step	**Temperature (°C)**	**Time (min)**	**Cycle**
Worm lysis	55	40	1
Worm lysis	65	30	1
Enzyme inactivation	95	20	1

Table 2.3: Primers used for genotyping

Gene name	Allele	Primer sequence (F-forward primer) (R-reverse primer)
cfp-1	*tm6369*	F:5′-TGG AAG AGT TAG TGG AGA ATT TGG-3′ R: 5′-TGT GCG AAA AAT GCA GTG C-3′
set-2	*bn129*	F: 5′-AAC ACC AAG AGC ACC ATC ATC-3′ R: 5′-CGT AGA AAG CAT CTG GCA GTC-3′
mys-4	*tm3161*	F: 5′-TCT CAA CTT CCC GGA ATT CTT-3′ R: 5′-TCT CCT TCG AAT TCC AGT TCA-3′
lin-15AB	*n765*	F: 5′-TTC AGG CTA GAC AGC AAC A-3′ R: 5′-TAT CGA TCC TGT GAT GGT C-3′
spr-1	*ok2144*	F: 5′-GGT ACT CCC GTT TGG TTG AA-3′ R: 5′-AGG CTT CAT GCA GCT TGT TT-3′
sin-3	*tm1276*	F: 5′-AAG GGC CAA CAG TCA ACA AC-3′ R: 5′-CCA AGA ACA AGT TCC GGT GA-3′

Table 2.4: PCR program

Step	Temperature (°C)	Time	Number of cycles
Initial denaturation	98	5 min	1
Denaturation	98	10 sec	
Annealing	56-62 (depending upon primers)	20 sec	31
Extension	72	40 sec	
Final extension	72	5 min	1

2.3 Genetic crosses

To generate double mutants, wild-type (N2 Bristol) males were crossed with L4 mutant hermaphrodites in a 3:1 ratio (**Figure 2.1**). From the crossed plate, heterozygous F1 males were picked and then crossed with hermaphrodites containing another mutation (**Figure 2.1**). From the crossed plate, eight L4 hermaphrodites (F1) were singled onto individual plates. Once the hermaphrodites reached the adult stage and started laying eggs, the hermaphrodites were picked for single-worm PCR to check for heterozygosity for both mutations. After the confirmation of heterozygosity, sixteen F2 hermaphrodites from the heterozygous plate were singled out. Once the F2 hermaphrodites started laying eggs, they were picked and analysed by single-worm PCR for homozygosity for both mutations. If F2 animals were homozygous for both mutations, they were further confirmed using single-worm PCR. For further confirmation, four *C. elegans* from a homozygous plate were separated in individual plates, and once they started laying eggs, they were picked for single-worm PCR. If isolated F2 progeny was homozygous for one mutation but heterozygous for another mutation, another 8 hermaphrodites (F3) were singled out and analysed by single-worm PCR for homozygosity for both mutations. For animals expressing a fluorescently tagged reporter gene, hetero or homozygosity was assayed by visualising fluorescence under a fluorescent dissecting microscope.

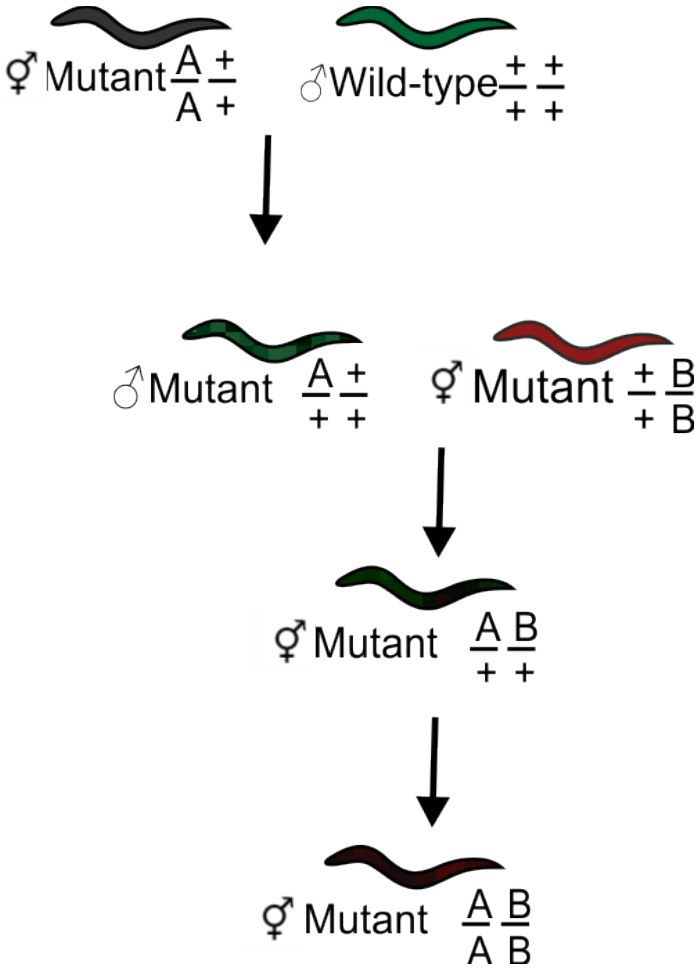


Figure 2.1: Genetic cross

Schematic representation of a double mutant generation. Hermaphrodites carrying the mutation A was crossed with a wild-type male. Heterozygous males from the crossed plate were picked and crossed with hermaphrodites carrying the mutation B. L4 (F1) hermaphrodites from the cross plate were singled, allowed to lay eggs, and single-worm PCR was performed on the F1 animal to confirm the presence of both mutations (heterozygosity). L4 (F2) animals of the heterozygous plate were singled out, allowed to lay eggs and single-worm PCR was perfomed to determine homozygosity of both mutations.

2.3.1 Double mutants generated in this study

set-2(bn129);mys-4(tm3161), cfp-1(tm6369);mys-4(tm3161), cfp-1(tm6369);kbIs5, set-2(bn129);kbIs5, set-2(bn129);spr-1(ok2144), cfp-1(tm6369);spr-1(ok2144), set-2(bn129);sin-3(tm1276), cfp-1(tm6369);sin-3(tm1276), cfp-1(tm6369);lin-15AB(n765), set-2(bn129);lin-15AB(n765), set-2(bn129);cfp-1(tm369).

2.4 Outcrossing

To remove or dilute the background mutations, *set-2(bn129), cfp-1(tm6369), mys-4(tm3161), kbIs5, spr-1(ok2144), sin-3(tm1276)* and *lin-15AB(n765)* mutant animals were outcrossed four-five times with wild-type males. For outcrossing, wild-type males were crossed with mutant hermaphrodites (L4 stage) in a 3:1 ratio (**Figure 2.2**). After 3-4 days, eight F1 hermaphrodites were picked from the crossed plate and transferred individually to eight separate seeded NGM plates. Once the F1 hermaphrodites reached the adult stage and started laying eggs, they were picked for single-worm PCR to check for heterozygosity. After the confirmation of heterozygosity, eight L4 hermaphrodites from the heterozygous plate were singled out and analysed by single-worm PCR for homozygosity. After obtaining the homozygous mutant, they were further confirmed using single-worm PCR. For further confirmation, four *C. elegans* from a homozygous plate were separated in individual plates, and once they started laying eggs, they were picked for single-worm PCR.

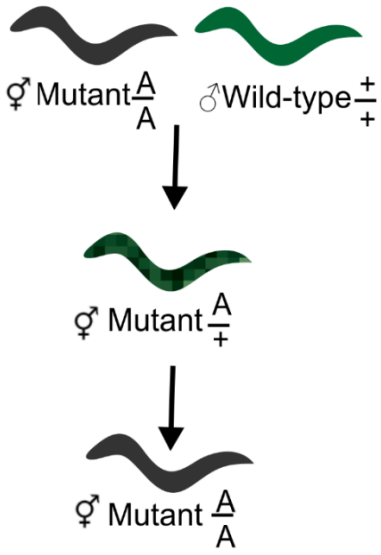


Figure 2.2: Outcrossing of single mutants

Schematic representation of outcrossing of single mutants. A hermaphrodite carrying mutation A was crossed with a wild-type male. F1 from the crossed plates were separated, and single-worm PCR was performed to check for heterozygosity. L4 hermaphrodites from the heterozygous plate were singled out and single-worm PCR was performed to check for homozygosity of mutation A.

2.5 Synchronization of *C. elegans* by bleaching

Animals were washed from agar plates with M9 buffer when more than 60% of the population reached the adulthood. An equal volume of bleach solution was added to tubes and left for 5 min. Tubes were centrifuged at 2000 rpm for 2 min at 20 °C, and the supernatant was aspirated. Pellets were washed with sterile water twice and once with M9 buffer. After the final wash, pellets were re-suspended in 10 ml of M9 buffer and left overnight on a shaker at 23 rpm at room temperature. Age synchronised L1 hatchlings were plated onto NGM agar plates seeded with OP50 bacteria.

2.6 Western Blot

To measure the levels of H3K4me3 in wild-type, *cfp-1(tm6369*) and *set-2(bn129)* mutants, a western blot analysis was performed. Embryos obtained from bleached adult worms were transferred in 15 mL Falcon tubes containing 10 mL of M9 buffer. Tubes were left on a shaker overnight at 20 °C to obtain starved L1 worms. Starved L1 (3.2-3.5 x10^3) larvae were pelleted in M9 buffer and snap-frozen in methanol and dry ice, and stored at −80 °C. Pellets were resuspended in 200 µl lysis buffer (50 mM Tris-Cl (pH 8), 300 mM NaCl, 1 mM PMSF, 1 mM EDTA, 0.5% Triton X-100) and protease inhibitor cocktail (Sigma Aldrich®) and sonicated at 20% amplitude (4 Watt) for 5-10 seconds and this step was repeated two times. Lysed samples were centrifuged at 13,523 g for 15 minutes at 4 °C, and supernatants were collected. Protein concentration in the supernatant was measured by the Bradford method of protein estimation (Kruger, 1994), and 50 µg of total protein was loaded onto a sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE, electric current condition: 45 mAmp, 45 minutes) in each well. To visualise H3K4me3 which is of similar molecular weight as H3, two sets of the same samples were loaded onto the same gel. One set of samples was used for H3 detection and another for H3K4me3 detection. Following electrophoresis, proteins were transferred from the gel to a nitrocellulose membrane, pre-soaked in transfer buffer (15 mM Tris.HCl, 150 mM glycine, 0.02% (w/v) SDS, 20% (v/v) methanol, pH 8.3), using Biorad TransBlot Turbo semidry Transfer apparatus at 25V, 1 A for 1 h.

After the transfer, the membrane was cut horizontally into two parts to be able to probe for tubulin (top) and, both Histone 3 and H3K4me3 (bottom). Furthermore, the bottom membrane was cut vertically into two parts to separate the two sets of samples and to be able to probe for H3K4me3 and Histone 3, which have approximately the same molecular weight (~15 kDa). Membranes were incubated with 5% non-fat milk in TBST buffer (20 mM Tris.HCl 150 mM NaCl, pH 7.4, 0.1% tween-20) for 1 h for blocking purpose. The membrane was washed with TBST and subsequently incubated overnight at 4 °C with 1:5,000 mouse monoclonal anti-H3K4me3 (Wako chemicals), 1:5,000 polyclonal rabbit anti-total H3 (Abcam) or 1:5,000 mouse monoclonal anti-tubulin (Sigma) antibodies diluted in blocking buffer (5% non-fat milk in TBST buffer). The membrane was washed twice with TBST for 10 minutes and incubated with 1:5,000 dilutions of HRP-linked secondary antibodies (goat anti rabbit (Abcam) or horse anti mouse Ig (New England Biolabs)). After incubation with the secondary antibody, the membrane was washed thrice with TBST for 10 min and once with TBS (20 mM Tris.HCl, 150 mM NaCl, pH 7.4). After the wash step, the western blot was developed using chemiluminescence (Pico Plus Kit, Thermo Scientific) and imaged using Alliance Q9 advanced gel imager (Uvitec, Cambridge).

2.7 Brood size assay

2.7.1 Brood size at 15 °C and 20 °C

*C. elega*ns was grown at respective temperatures for four generations. Either 10 L4 larvae were picked and transferred to an individual plate, or 3-5 L4 larvae per plate were picked onto two to three plates. After the animals reached the adult stage, they were left for ~24 h on the plate to lay eggs. After ~24 h, animals were transferred to new plates every day or every other day until egg laying stopped (**Figure 2.3**). Old plates were counted for a total number of eggs on plates. These plates were stored at respective temperatures for ~24-48 h and were scored for the number of live progeny. Eggs that were not hatched after ~24 h incubation were counted as dead. Statistical analysis was performed using a one-sided t-test. A p-value of 5% or lower

was considered to be statistically significant. Animals that crawled out of the plates and were lost were not included in the study.

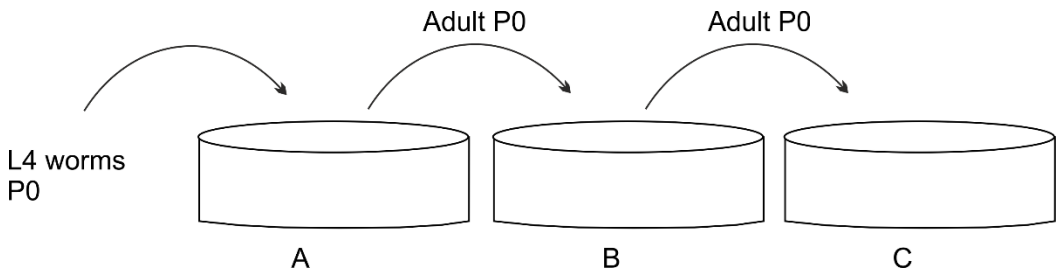


Figure 2.3: Schematic representation of worm transfer for fertility assessment at 20 °C and 15 °C

L4 larvae maintained at respective temperatures were transferred into plate A. After animals reached the adulthood, they were left for ~24 h on the plate A for egg laying. After ~24 h animals were transferred to a fresh plate B, and after ~24-48 h the adult animals were transferred to new plate C and left for ~24-48 h.

2.7.2 Brood size at 25 °C

For the brood size assay at 25 °C, twenty L4 (P0) larvae were picked from 20 °C and transferred to new OP50 seeded plates. They were allowed to lay eggs overnight at 25 °C. The next day, all mother (P0) hermaphrodites were picked and transferred to new plates and left for another 5-6 h (**Figure 2.4**). Mothers (P0) from new plates were removed, and eggs were allowed to reach L4 (F1) at 25 °C. Ten L4 stage F1 animals were picked and transferred to an individual plate. After the animals reached the adult stage, they were left for ~24 h on the plate to lay eggs. After ~24 h, animals were transferred to new plates every day or every other day until egg laying stopped (**Figure 2.4**). Old plates were counted for a total number of eggs on plates. These plates were stored at 25 °C for ~24-48 h and were scored for the number of live progeny. Unhatched eggs after ~24 of incubation were counted as dead. A one-sided t-test was used to measure the level of significance. A p-value of 5% or lower was considered to be statistically significant.

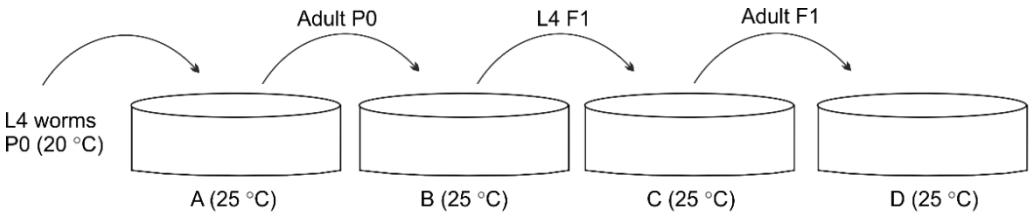


Figure 2.4: Schematic representation of worm transfer for fertility assessment at 25 °C

L4 larvae maintained at 20 °C were transferred into plate A at 25 °C. After ~24 h the animals were transferred to a fresh plate B and left for 5-6 h. L4 (F1) animals from plate B were transferred to new plate C. After the animals reached the adulthood, they were left for ~24 h on the plate C to lay eggs. After ~24-48 the adult animals were transferred to new plate D and left for ~24-48 h.

2.8 Development assay

Twenty to forty synchronised L4 (P0) larvae were picked from 20 °C and then transferred onto fresh OP50 seeded plates. They were allowed to lay eggs overnight at 20 °C. The next day, all mother hermaphrodites (P0) were picked and transferred to new plates to lay eggs and left for 5-6 h. Mothers (P0) from the new plates were removed, and eggs were allowed to grow for the indicated number of hours. F1 animals were transferred to Eppendorf tubes, washed twice with M9 buffer, fixed with 95% methanol, frozen for 1 h at -20 °C, washed twice with M9 buffer and stained with 1 ng/mL 4′,6-diamidino-2-phenylindole (DAPI) (Sigma) for 10 min. After staining, animals were washed three times with M9. *C. elegans* were transferred onto microscope slides and visualized by fluorescence microscopy. I scored the development stage of the animals using gonad structure (**Figure 2.5**). Animals were scored as L3 larvae if gonads were extended outwards but not turned. I scored animals as L4 larvae if gonad arms were extended outwards and turned. I also checked for the presence of sperm and immature vulva in L4 larvae. Animals with completely turned gonad and had oocytes or/and embryos along with matured vulva were scored as an adult.

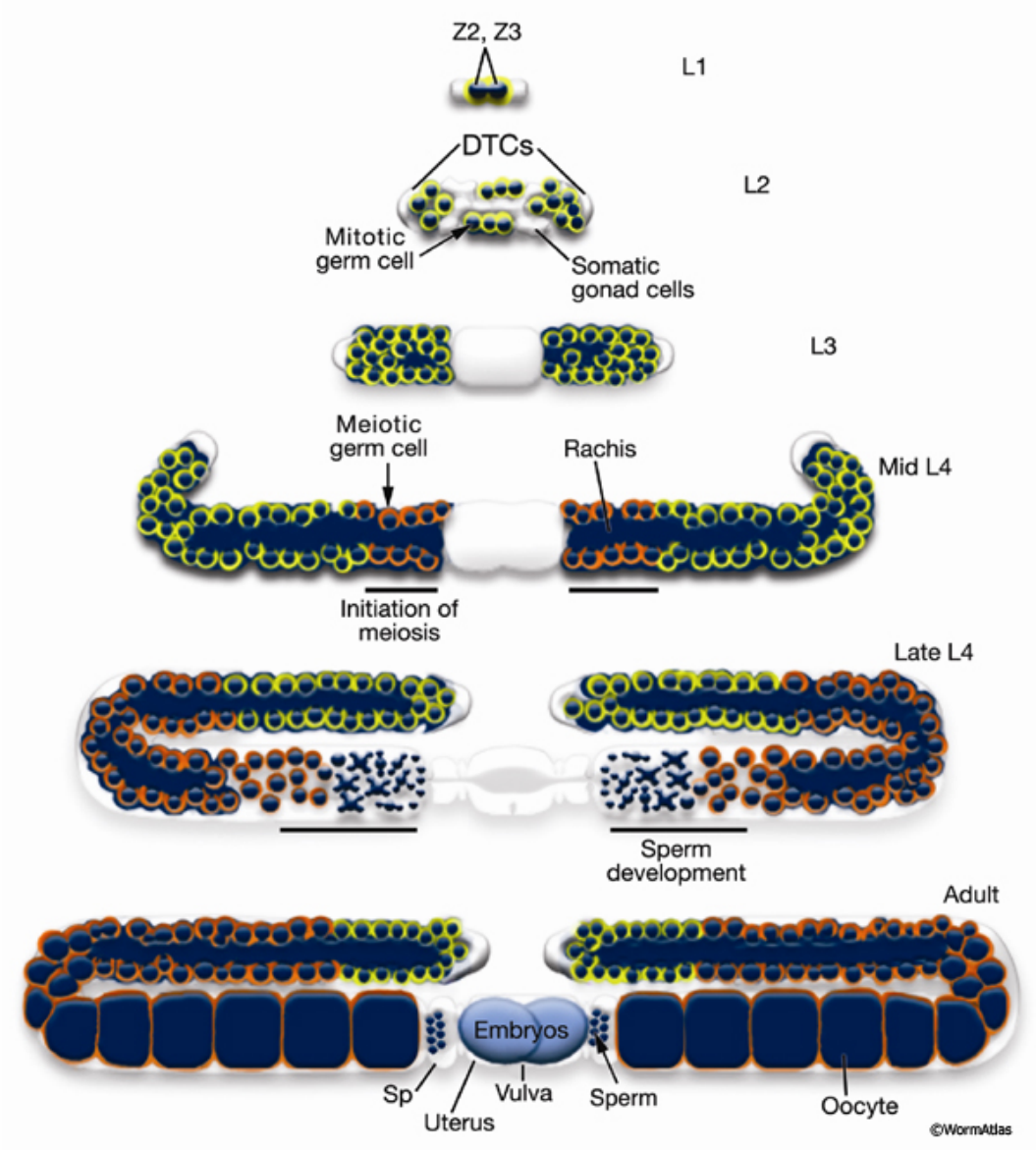


Figure 2.5: Scoring scheme for development assay

L3: gonad arms are extending outwards but not turned. L4: gonad arms are extending outwards and turned, sperm is present and vulva is immature. Adult: gonad arms are completely turned and positioned at midline, oocytes or/and embryos are present, and the vulva is matured. The figure is taken from (Lints, 2009).

2.9 Analysis of chromosomal abnormalities

Synchronised young adult animals were DAPI stained as described in section 2.8. Animals were imaged using a Zeiss LSM880 confocal microscope through a 63x/1.4 or 40x/1.4 oil DIC M27. The excitation and emission wavelengths used for DAPI was 405 nm and 459 nm, respectively. DAPI-stained bodies present within diakinesis (DI) nuclei were imaged as collapsed Z-stacks. The total number of DAPI stained bodies present within DI nuclei were counted. In the wild-type DI nucleus, 6 DAPI stained bodies were present. Chromosomal abnormalities were identified by the presence of more than six or less than six DAPI stained bodies in DI nucleus. For the presentation, images were processed using CorelDraw software.

2.10 Agar pad preparation

For agar pads, 2-5% agarose (Life Science, UK) was mixed in distilled water and kept on a heat block at 65 °C. Two glass slides, each with tape attached to increase the elevation, were placed on either side of a clean glass slide. 150 µl of the molten agarose was added at the centre of the middle slide, and a clean glass slide was placed on the top (perpendicular to the centre slide). After the agar settled, the top glass slide was gently removed. For imaging, *C. elegans* was anaesthetised in 10-25 mM sodium azide (Acros Organics) and positioned on the agar pad with the help of an eyelash.

2.11 SynMuv suppressor screen

For the SynMuv suppression screen, twenty L4 larvae were picked from 15 °C and transferred to OP50 seeded plates at 20 °C. They were allowed to lay eggs overnight. Next day, all mothers were picked and transferred to new plates. Eggs on the new plates were incubated until 80-90% reached the adult stage. *C. elegans* were observed under a Nomarski microscope for protrusions on the ventral side, and the number of protrusions was scored (Cui et al., 2006b).

2.12 RNAi plates

RNAi clones were streaked on LB agar plates containing ampicillin (100 µg/mL) and tetracycline (12.5 µg/mL) and incubated overnight at 37 °C. RNAi clones were obtained from the Ahringer RNAi library. The overnight culture was inoculated in 2 ml LB broth with ampicillin (100 µg/mL) and incubated for 6-8 h at 37 °C in a shaking incubator. The grown bacterial culture was seeded on dried NGM plates containing 1 mM IPTG (Generon, UK) and ampicillin (100 µg/mL). Seeded plates were dried at room temperature for 3 days. For the RNAi negative control, HT115(DE3) containing the empty L4440 RNAi feeding vector was used.

2.12.1 Fertility assay at 20 °C

For all fertility experiments using RNAi bacteria, except for *hda-1*, L1 larvae were spotted on RNAi plates, and their progeny (F1) were used for the experiments (**Figure 2.6**). For *hda-1* RNAi, spotted L1 (P0) were used for the experiments. Fertility was assayed as described in section 2.7.1 at 20 °C.

2.12.2 Fertility assay at 15 ° C.

For fertility assay at 15 °C, synchronised adult *C. elegans* maintained at 15 °C for at least 4 generations were bleached and hatched L1 larvae were seeded onto RNAi plates. Plates were incubated at 15 °C and fertility of their progeny (F1) were assayed as described in section 2.7.1 at 15 °C.

2.12.3 Fertility assay at 25 ° C

Synchronised adult animals maintained at 20 °C were bleached and hatched L1 larvae were seeded onto RNAi plates. When animals reached the L4 (P0) stage, they were picked from 20 °C and transferred to fresh RNAi plates and incubated overnight at 25 °C. The next day, animals were picked and transferred to fresh RNAi plates and left for another 5-6 h for laying. After 5-6 h, mothers (P0) were removed and eggs were allowed to reach L4 (F1) at 25 °C. Fertility of F1 was assayed as described in section 2.7.2.

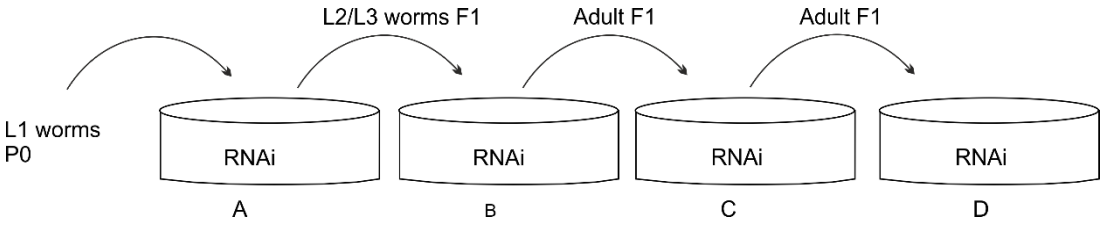


Figure 2.6: Schematic representation of worm transfer for fertility screen
L1 larvae maintained at 20 °C were transferred into RNAi plate A. L2/L3 (F1) larvae were transferred to a fresh RNAi plate B. After animals reached the adult stage, they were left for ~24 h on the plate B for laying. After ~24 h the larvae were transferred to a fresh RNAi plate C. After ~24-48 h the adult larvae were transferred to new RNAi plate D and left for ~24-48 h.

2.12.4 SynMuv suppressor screens

Synchronised L1 larvae were grown at 15 °C on RNAi plates until they reached young adulthood. Young adults were bleached, and embryos were re-suspended in 10 ml of M9 buffer and left on a shaker at 23 rpm overnight at 15 °C. Hatched L1 larvae were spotted on an RNAi plate, and the plates were incubated for 3 days at 20 °C. After 3 days, animals were scored for the Muv phenotype under a stereomicroscope and/or Nomarski microscope. For *lin-12(n137n460),* strain plates were incubated for 3 days at 15 °C.

2.12.5 Egg laying defective assay

C. elegans grown at 20 °C were bleached, and embryos were re-suspended in 10 ml of M9 buffer and left on a shaker at 23 rpm overnight at 15 °C. Hatched L1 larvae were spotted on desired RNAi plates, and the plates were incubated for 3-4 days at 20 °C. After 3-4 days, animals were scored for the egg-laying defective (Egl) phenotype. An animal was scored as *egl* if it retained its eggs (bagged) or its progeny had hatched internally.

2.12.6 Development assay

L1 larvae were spotted onto an RNAi plate, and when they reached the L4 stage, twenty to forty synchronised L4 larvae were picked and transferred to a fresh RNAi plate. They were allowed to lay eggs overnight at 20 °C. Next day, all mother hermaphrodites were picked and transferred to new RNAi plates and left to lay eggs for 5-6 h. Mothers from the new plates were removed, and eggs were allowed to grow for an indicated number of hours. After a respective time, animals were transferred to Eppendorf or falcon tubes, washed twice with M9, frozen for 1 h at -20 °C, DAPI stained and observed under a Nomarski microscope.

2.13 Fertility assay of TSA treated *C. elegans*

NGM plates containing 4 µM Trichostatin A (TSA) (Sigma) or Dimethyl sulphoxide (DMSO) (Fisher Bioreagents) were prepared. TSA was dissolved in DMSO, therefore DMSO was used as the control. OP50 containing 4 µM TSA or DMSO was seeded onto the respective plates. L1 (P0) larvae were transferred into TSA or DMSO plates and incubated at 20 °C. Either ten L4 (F1) larvae were singled onto individual plates, or 3 L4 (F1) larvae per plate were picked onto three TSA or DMSO plates. Fertility was assayed as described in section 2.7.1, at 20 °C.

2.14 T-test for fertility assays

A student t-test was performed to investigate the potential interaction between the two genes. Under the null hypothesis, where no genetic interaction between two genes is assumed, the expected brood size of the double mutants (or RNAi knockdown of a gene in a single mutant) is the product of the brood size of the single mutants (or single mutant and the RNAi knockdown of the gene in a wild-type) divided by the average brood size of wild-type. A one-sided t-test is done to compare the expected (under null-hypothesis) brood size with the observed brood size of double mutants (or RNAi knockdown of a gene in a single mutant).

$$Brood_{H_0} = \frac{Brood_{gene\ 1} \times Brood_{gene\ 2}}{Brood_{WT}}$$

$Brood_{H_0}$ = Expected Brood size of double mutant (or RNAi) under the null hypothesis

$Brood_{gene\ 1}$ = Actual Brood size of first mutant (or RNAi)

$Brood_{gene\ 2}$ = Actual Brood size of second mutant (or RNAi)

$Brood_{WT}$ =Actual Brood size of Wild-type

2.15 Fisher's exact-test for multivulva assays

Fisher's exact-test was performed to investigate the potential interaction between two genes. Under the null hypothesis, where no genetic interaction between two genes is assumed, the expected percentage of multivulva in double mutants (or RNAi knockdown of a gene in a single mutant) is the product of the percentage of multivulva in the single mutants (or single mutant and the RNAi knockdown of the gene in a *lin-15AB(n765)* mutants) divided by the average percentage of multivulva in *lin-15AB(n765)* mutants. Fisher's exact-test was done to compare the expected (under null-hypothesis) percentage of multivulva with the observed percentage of multivulva in double mutants (or RNAi knockdown of a gene in a single mutant).

$$Muv_{H_0} = \frac{Muv_{gene\ 1} \times Muv_{gene\ 2}}{Muv_{lin-15AB}}$$

Muv_{H_0} = Expected percentage of multivulva of double mutant (or RNAi) under the null hypothesis

$Muv_{gene\ 1}$ = Actual percentage of multivulva in first mutant (or RNAi)

$Muv_{gene\ 2}$ = Actual percentage of multivulva in second mutant (or RNAi)

$Muv_{lin-15AB}$ =Actual percentage of multivulva in *lin-15AB*

2.16 Heat Shock experiments

Synchronised L1 larvae were grown at 20 °C until they reached young adulthood. For heat shock, plates containing desired strains were tightly parafilmed and placed into a pre-heated water bath at 33 °C for 1 h. After heat shock, animals were washed off from the plate with M9 buffer, collected into Eppendorf tubes, centrifuged (2000 rpm for 2 min) and washed three times in M9 buffer. Pellets were then snap frozen in methanol and dry ice, and stored at -80 °C.

2.17 Salt induction experiments

2.17.1 Salt induction reporter assay

For the salt induction assay, NGM plates containing 52 (control) and 200 mM NaCl were prepared. Starved L1-stage larvae expressing a salt-stress inducible reporter gene (*gpdh-1p::GFP*) were placed on NGM plates with 52 or 200 mM NaCl. Plates were incubated at 20 °C until animals reached the adult stage (~72 h). Animals were visualized using a Leica fluorescence microscope and scored for the expression of GFP.

2.17.2 Fluorescence Image Quantification

L1-stage larvae were seeded onto NGM-Agar plates supplemented with 52 mM (control) or 200 mM NaCl concentration. Plates were incubated at 20 °C. Once the animals reached the adult stage, they were anesthetised in 5 mM levamisole solution and mounted on 2% agar pads. *C. elegans* were positioned on the agar pad using a worm pick. *C. elegans* were imaged using a Zeiss LSM700 confocal microscope through a 10x/1.0 numerical aperture objective. The excitation and emission wavelengths used for GFP was 488 nm and 518 nm, respectively. Image settings were kept constant and images were analysed with ZEN 2.6 software. The expression of GFP during osmotic stress (200 mM NaCl) compared to control was quantified. For the quantification of *gpdh-1p::GFP* fluorescence intensity, ImageJ software was used. Mean fluorescence intensity per animal was calculated for three biological replicates ($n > 15$).

2.17.3 Salt induction experiment for qPCR

For qPCR, starved L1-stage larvae were placed on NGM plates containing 52 mM, and 150 mM NaCl. After ~72 h, animals were collected in Eppendorf tubes by washing off the animals with M9 buffer, and centrifuged (2000 rpm for 2 min). Collected animals were washed three times in M9 buffer. Pellets were then snap frozen in methanol and dry ice, and stored at -80 °C.

2.18 RNAi efficiency experiment for qPCR.

L1 larvae were seeded onto plates containing RNAi bacteria and incubated at 20 °C. Once they reached the adult stage, they were bleach-synchronised, and eggs were hatched overnight at room temperature. Hatched L1 larvae were spotted onto fresh RNAi plates and grown at 20 °C until they reached young adulthood (Day 1 adults). Animals were picked and transferred to unseeded NGM plates (transfer plate) to minimise bacterial contamination. Animals were washed off from the transfer plate with M9 buffer and collected into Eppendorf tubes on ice. To remove residual bacteria, animals were washed three times with M9 buffer. For washing, animals were left in tubes on ice for 3-5 minutes to settle down on the bottom of the tube. M9 buffer present in the tube was carefully aspirated, without disturbing the *C. elegans* pellet in the bottom. Fresh M9 was added and the above steps were repeated 3 times. Pellets were then snap frozen in liquid nitrogen, and stored at -80 °C.

2.19 RNA extraction and qPCR

2.19.1 RNA extraction

RNA was extracted using a Direct-zol RNA miniprep kit. Frozen animals were homogenized in 600 µl of TRIzol by vigorous shaking for 15 min. After homogenization, an equal volume of 100% ethanol was added, mixed thoroughly and centrifuged at 10,000 g for 30 sec (all following centrifuge steps were done at this speed for 30 sec). The supernatant was transferred into a Zymo-Spin™ III CG column in a collection tube and centrifuged. Flow-through was discarded, and the column was transferred into a new collection tube. 400 µl of RNA wash buffer was added to the column and centrifuged. In an RNase-free tube, 5 µl DNase I (6 U/µl) and 75 µl of DNA digestion buffer were added. This mixture was directly loaded into the top of the column and left to incubate at room temperature for 15 min. 400 µl of Direct-zol™ RNA prewash was added to the column and centrifuged, and the flow-through was discarded. This step was repeated once. 700 µl of RNA wash buffer was added to the column and centrifuged for 2 min 10,000 g. The column was transferred into an RNase-free tube. To elute RNA, 50 µl of DNase/RNase-free

water was added directly to the column matrix and centrifuged. The extracted RNA was stored at -80 °C.

2.19.2 cDNA synthesis

Before cDNA synthesis, RNA sample quality was assessed by running the sample on an agarose gel. I checked whether 18S and 28S rRNA appeared as sharp bands after electrophoresis. A clear and sharp band indicates the presence of intact RNA and smeared bands indicate degraded mRNA. I also checked whether the intensity of 28S rRNA was approximately twice as intense as the 18S rRNA. The concentration of RNA was measured using a nanodrop. Extracted RNA was reverse transcribed to obtain cDNA using Thermo Scientific™ RevertAid first strand cDNA synthesis kit. For cDNA synthesis, an equal concentration of RNA was used. 5 µl of the RNA with 1 µl of a random primer, 4 µl 5X reaction buffer, 6 µl nuclease-free water, 2 µl dNTP mix (10 mM), 1 µl RiboLock RNase inhibitor (20 U µl-1) and 1 µl RevertAid M-MuLV RT (200U µl-1). The mixture was incubated for 5 min at 25 °C, 60 min at 42 °C and for 5 min at 70 °C in the thermocycler.

2.19.3 qPCR

qPCR was performed with SYBR® Green (Biorad). For qPCR, a reaction mix containing 10 µl SYBR Green supermix, 1.5 µl (2 µM) forward primer, 1.5 µl (2 µM) reverse primer (**Table 2.5)**, 3 µl cDNA and 4 µl nuclease-free water was prepared. The reaction mix was loaded as technical duplicates or triplicates onto a 96-well plate, and the plate was then placed into a thermal cycler (Stratagene Mx3000P) to perform the qPCR (see **Table 2.6** for the program). For higher qPCR efficiency (90-100%), the reaction was optimized by changing the concentration of primers and the concentration of templates (see “efficiency of primers” below). mRNA abundance was measured using $1/2^{-\Delta Ct}$ formula.

ΔCt (cycle threshold) = Ct (a target gene) – Ct (average of Ct value of two reference genes).

mRNA abundance = $1/2^{-\Delta Ct}$

A t-test was performed in ΔCt values.

Table 2.5: Primers used for qPCR

Gene name	Primers F-forward and R-reverse
C12C8.1	F:5′-TCA ATG GGA AGG ACC TCA ACT-3′ R:5′-AGT GGG ACA ACA TCA ACG AGT-3′
F44E5.4	F:5′-ATC TAT CAG AAT GGA AAG GTT GAG A-3′ R:5′-CTG GAT TAC GAG CTG CTT GA-3′
hsp-16.11	F:5′-CTC CAT CTG AAT CTT CTG AGA TTG T-3′ R:5′-GAA TTG ATA ATG TAT GTC CAT CCA AA-3′
tba-1	F:5′-TGA TCT CTG CTG ACA AGG CTT AC-3′ R:5′-GCA CTT CAC CAT TTG GTT GG-3′
pmp-3	F:5′-GCT TGA TAA TCC AGA TCA ACG TCT-3′ R:5′-GGA CCA ATC CAA CCG GAA C-3′
gpdh-1	F:5′-GTG ATA AGC CGC TCC GAG T-3′ R:5′-AAC ATC CTG AGC TGT TGG CG-3′
cfp-1	F:5′-CTCTGGGAATGGAGGAGCTG-3′ F:5′-CGGGTTCTGCATTCAACTGT-3′

Table 2.6: qPCR program

Step	Temperature (°C)	Time	Cycle
Initial denaturation	95	30 sec	1
Denaturation	95	30 sec	
Annealing	60	30 sec	40
Extension	72	40 sec	
Final extension	72	5 min	1

2.20 Efficiency of primers

Template cDNA was serially diluted in a 1:10 ratio (10, 100, 1000, 10000, 10000). 3 µL of appropriate template dilution was added to the wells of 96-well plate containing 10 µl SYBR Green Supermix, 1.5 µl (2 µM) forward primer, 1.5 µl (2 µM) reverse primer and 4 µl nuclease-free water. The reaction mix was loaded in duplicates/triplicates in a 96-well plate, and the plate was then placed into a thermal cycler (qPCR) program (**Table 2.6**). The Ct values were used to obtain a calibration curve. Ct (Y-axis) values were plotted against template concentration (X-axis) and a straight linear curve with correlation coefficient >0.98 was obtained. Efficiency was calculated by using the formula below:

Efficiency (E) = $(10^{\wedge(-1/slope)}-1)*100$

The concentration of template and/or primers and/or the annealing temperature was changed to obtain 90-100% efficiency.

2.21 The composition of media and buffers

2.21.1 The composition of NGM media

- 17 g agar (Sigma Aldrich®)
- 3 g NaCl (Acros Organics)
- 2.5 g peptone (Sigma Aldrich®)
- 1 mL of 5 mg/mL cholesterol (Sigma) dissolved in ethanol
- 1 mL of 1 M $MgSO_4$ (Sigma Aldrich®)
- 1 mL of 1 M $CaCl_2$ (Sigma Aldrich®)
- 25 mL of 1 M KPO_4 (Sigma Aldrich®) buffer pH 6.0.
- H_2O to 1 litre

2.21.2 The composition of Luria-Bertani (LB) medium

- 10 g Bacto-tryptone (Sigma Aldrich®)
- 5 g Bacto-yeast (Sigma Aldrich®)
- 5 g NaCl (Acros Organics)

- 15 g agar (Sigma Aldrich®)
- H_2O to 1 litre, pH 7.5

Note: For LB broth, agar was not added.

2.21.3 The composition of M9 buffer

- 3 g KH_2PO_4 (Sigma Aldrich®)
- 6 g Na_2HPO_4 (Sigma Aldrich®)
- 5 g NaCl (Acros Organics)
- 1 ml 1 M $MgSO_4$ (Sigma Aldrich®)
- H_2O to 1 litre

2.21.4 The composition of the bleach solution

- 3 ml of 4.99% NaClO (Sigma Aldrich®)
- 1 ml of 10 M NaOH (Sigma Aldrich®)
- 6 ml of sterile water

2.21.5 The composition of TBST buffer solution

- 20 mM Tris HCl
- 150 mM NaCl
- 0.1% Tween 20

2.21.6 The composition of PK lysis buffer

- 7 µl of nuclease-free H_2O (Ambion, UK)
- 2 µl of Phusion High-Fidelity Buffer (New England Biolabs)
- 1 µl of 20 mg/ml proteinase K (Thermo Scientific, UK)

Chapter 3

Investigating the role of CFP-1 in *C. elegans* development

3.1 Introduction

Chromatin regulation controls almost all biological processes in the nucleus. Chromatin can be modified by enzymatic complexes that coordinately regulate chromatin composition and structure. One of the major forms of chromatin modifications is histone modification, which plays a central role in cellular gene expression, underpinning cellular physiology and developmental processes (Kouzarides, 2007, Portela and Esteller, 2010, Shilatifard, 2012, Shilatifard, 2008, Tessarz and Kouzarides, 2014, Salz et al., 2014, Bertero et al., 2015). Perturbation of histone modifications is associated with different developmental defects and diseases, including cancer (Shilatifard, 2012, Allis and Jenuwein, 2016, Salz et al., 2014, Bertero et al., 2015). However, it is unknown how histone modifications contribute to these events.

Histone 3 lysine 4 tri-methylation (H3K4me3) is a histone lysine methylation modification often found at active promoters (Barski et al., 2007, Pokholok et al., 2005). H3K4me3 is mainly deposited by the evolutionarily conserved Complex Proteins Associated with the Set1(COMPASS) complex. The COMPASS complex, first identified in yeast, is conserved from yeast to mammals (Miller et al., 2001, Lee et al., 2007a, Ardehali et al., 2011, Xiao et al., 2011). In addition to yeast Set1, other subunits of the yeast Set1/COMPASS complex are also evolutionarily conserved and play an important role in H3K4 methylation (Miller et al., 2001, Lee et al., 2007a, Ardehali et al., 2011, Xiao et al., 2011, Hughes et al., 2004, Simonet et al., 2007). One of the major subunits of the COMPASS complex is Cfp1 (Chun et al., 2014, Chen et al., 2014, Thomson et al., 2010, Shilatifard, 2012).

Cfp1 is an evolutionarily conserved epigenetic regulator that recognises non-methylated CG dinucleotides (CpGs) (Chun et al., 2014, Chen et al., 2014, Thomson

et al., 2010). Cfp1 binds to 81% of unmethylated CpG islands *in vivo* (Thomson et al., 2010). Cfp1 is required for H3K4me3 deposition in mammalian promoter regions (Lee and Skalnik, 2005, Ardehali et al., 2011, Shilatifard, 2012, Yu et al., 2017, Clouaire et al., 2014, Clouaire et al., 2012, Thomson et al., 2010). Additionally, Cfp1 is important during the early stage of mammalian development (Carlone and Skalnik, 2001, Young and Skalnik, 2007, Chun et al., 2014, Young et al., 2006). However, as the deletion of *Cfp1* causes early embryonic lethality, it is not clear how exactly Cfp1 contributes to the development of an organism.

To understand the role of Cfp1 in development, in this study, the *cfp-1(tm6369) C. elegans* mutant was phenotypically characterised and compared with the reported *set-2(bn129)* loss-of-function mutant allele, which has low levels of H3K4me3. It was observed that in the *cfp-1(tm6369)* mutant, H3K4me3 levels were significantly reduced and the level of H3K4me3 was similar to that reported in the *set-2(bn129)* mutant. Furthermore, *cfp-1(tm6369)* mutants or wild-type animals during *cfp-1* RNAi have a significant reduction in fertility at 15 °C or 20 °C and were nearly sterile at 25 °C. I also found that the development of the *cfp-1(tm6369)* mutant was delayed compared to wild-type animals. Additionally, I observed chromosomal abnormalities in the *cfp-1(tm6369)* mutant oocytes. Taken together, these results suggest that CFP-1 is required for global H3K4me3, and plays an important role in maintaining the fertility and the development of *C. elegans*.

3.2 Results

3.2.1 The global level of H3K4me3 is drastically reduced in the *cfp-1(tm6369)* mutant

Reports show that the deletion of Cfp1 homologues in yeast and mammals results in lower levels of H3K4me3 at 5′ sites of active genes (Voo et al., 2000, Lee and Skalnik, 2005, Brown et al., 2017, Clouaire et al., 2012). The function of SET-2 and CFP-1 in H3K4me3 deposition is also conserved in *C. elegans*. It has been reported that both *set-2(bn129)* loss of function mutant and RNAi knockdown of *cfp-1* in *C. elegans* lead to a significant reduction in global H3K4me3 levels (Robert et al., 2014, Simonet et al., 2007, Xiao et al., 2011). To further investigate the role of CFP-1 in development, I used *cfp-1(tm6369)* mutants. *cfp-1(tm6369)* is a deletion allele, in which 254 bp encompassing part of intron 4, exon 5 and part of intron 5 is deleted (**Figure 3.1 A**). This deletion is predicted to produce a truncated CFP-1 protein. The *cfp-1(tm6369)* mutant was confirmed by genotyping before conducting experiments (**Appendix 9.1 A**) and outcrossed 5 times.

To investigate whether *cfp-1(tm6369)* is a loss of function allele, the global expression levels of H3K4me3 in *cfp-1(tm6369)* mutants was measured by western blot analysis. Western blot analysis was carried out in collaboration with Dr Amit Kumar (Brockwell Laboratory, University of Leeds). The *cfp-1(tm6369)* mutant repeatedly showed a drastic reduction of the H3K4me3 levels compared to wild-type animals (**Figure 3.1 B**)(Pokhrel et al., 2019). The *set-2(bn129)* loss of function mutant was used to compare H3K4me3 levels with the *cfp-1(tm6369)* mutant. The *set-2(bn129)* mutant was genotyped (**Appendix 9.1 B**) and outcrossed 5 times before further use. In the *set-2(bn129)* mutant, global levels of H3K4me3 were similar to those in the *cfp-1(tm6369)* mutant (**Figure 3.1 B and Appendix 9.2**) (Xiao et al., 2011, Robert et al., 2014, Herbette et al., 2017). This suggests that *cfp-1(tm6369)* is a loss of function allele and has a similar impact on H3K4me3 deposition as the loss-of-function *set-2(bn129)* mutant.

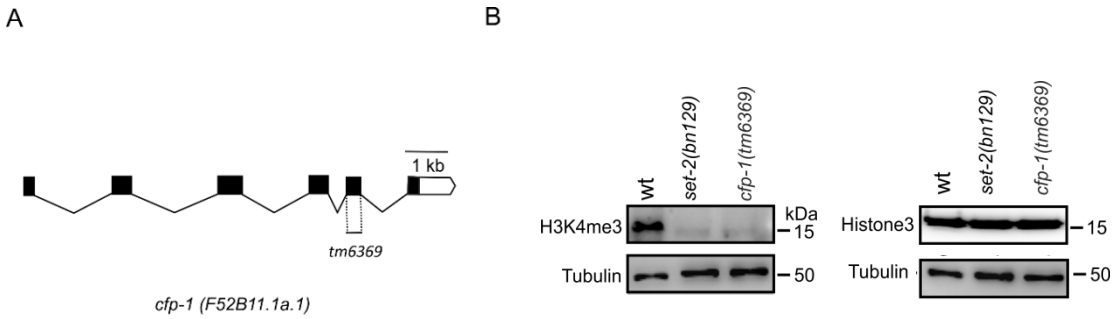


Figure 3.1: Molecular characterization of the *cfp-1(tm6369)* allele

(A) Diagrammatic representation of the *cfp-1(tm6369)* allele. The sequence of exon 5 of F52B11.1a.1 is conserved in all other transcripts of the *cfp-1* gene. 254 bp encompassing exon 5 of the F52B11.1a.1 and part of the intron upstream and downstream region is deleted. Black colour denotes the exons. (B) Western blot analysis showing that the H3K4me3 levels were significantly reduced in *cfp-1(tm6369)* and *set-2(bn129)* mutants compared to wild-type (wt). Histone 3 (H3) and tubulin were used as loading controls. The molecular weight of H3K4me3 and H3 is approximately similar. Therefore, to probe H3K4me3 and H3, I loaded three samples (wt, *set-2(bn129)* and *cfp-1(tm6369)*) twice on the same SDS gel (Appendix 9.2). The image on the left shows the levels of H3K4me3 and corresponding tubulin levels (loading control) in wt, *cfp-1(tm6369)* and *set-2(bn129)* mutants. The image on the right shows the levels of H3 and corresponding tubulin levels in wt, *cfp-1(tm6369)* and *set-2(bn129)* mutants.

3.2.2 CFP-1 promotes fertility

Loss of H3K4me3 in *set-2(bn129)* mutants is reported to be accompanied by a reduction in the brood size at 15 °C and 20 °C (Xiao et al., 2011, Herbette et al., 2017). Since CFP-1 and SET-2 are the major subunits of the SET-2/COMPASS complex and are required for the majority of H3K4me3 (**Figure 3.1 B**), they might have a similar phenotype. To explore the role of CFP-1 in fertility, wild-type (N2) animals were treated with *cfp-1* RNAi or RNAi negative control (empty vector) and the brood size was measured, distinguishing between live and dead animals, at 15 °C and 20 °C. The average live brood size of wild-type animals treated with *cfp-1* RNAi was significantly reduced compared to RNAi negative control, at both temperatures (**Figure 3.2 A and B**). However, no significant changes in the number of dead progeny was observed (**Figure 3.2 A and B**). To confirm that the observed reduction of brood size of *cfp-1* RNAi-treated animals is due to reduced *cfp-1* mRNA levels, mRNA transcript levels were measured by qRT-PCR. Upon *cfp-1* RNAi, levels of the *cfp-1* mRNA were significantly reduced in wild-type animals, and the level of reduction was 74% compared to RNAi negative control (**Figure 3.2 C**). This result supports that the loss of function of CFP-1 results in a significant reduction of brood size.

To further confirm that the observed reduction of brood size upon *cfp-1* RNAi was due to the loss of CFP-1 function, brood size of the *cfp-1(tm6369)* mutant was measured at 15 °C and 20 °C. Consistent with the RNAi results, the average live brood size of *cfp-1(tm6369)* animals was significantly reduced compared to the wild-type by 59% and 57% at 15 °C and 20 °C, respectively (**Figure 3.2 D and E**). Additionally, the level of reduction was similar to published data for the *set-2(bn129)* loss of function allele (**Figure 3.2 D and E**) (Xiao et al., 2011). The average live brood size of *set-2(bn129)* animals was reduced by 49% and 54% at 15 °C and 20 °C, respectively (**Figure 3.2 D and E**). Similar to RNAi results, I did not observe any significant change in the percentage of dead eggs in *cfp-1(tm6369)* and *set-2(bn129)* mutants. Taken together, these results suggest that both SET-2 and CFP-1 play a vital role to promote *C. elegans* fertility.

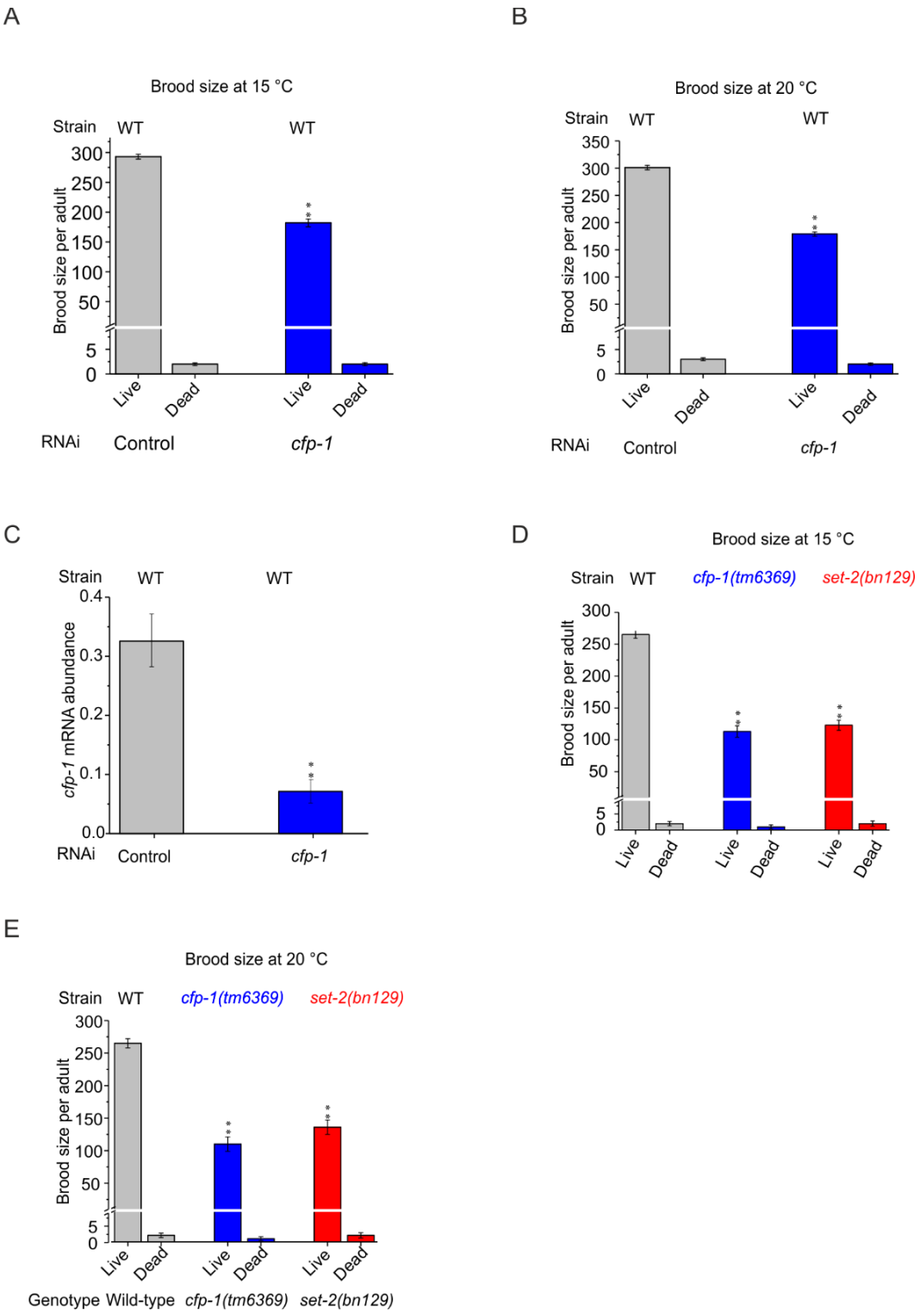


Figure 3.2: CFP-1 is required for fertility

(A and B) The average brood size of *C. elegans* treated with *cfp-1* RNAi compared to RNAi negative control (control) at 15 and 20 °C. The average live brood size of wild-type animals during *cfp-1* RNAi was significantly reduced compared to

control. The data represent three biological replicates (n=9 in each replicate). (C) *cfp-1* mRNA transcript levels in *C. elegans* treated with control RNAi (grey) and *cfp-1* RNAi (blue). *pmp-3* and *tba-1* genes were used for normalisation. The figure represents the average of three biological replicates (n = 50 in each replicate). ** = P<0.01. (D and E) Total brood size (dead and live) assay for wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants at 15 and 20 °C. The average live brood size of *cfp-1(tm6369)* and *set-2(bn129*) mutants was significantly reduced compared to wild type at both temperatures. Two biological replicates were combined for the figure (n=10 in each experiment). P-values were calculated using a student t-test: ** = P<0.01. Error bars represent ± standard error of the mean (SEM).

3.2.3 Loss of CFP-1 function results in temperature-sensitive sterility

C. elegans are maintained at 20 °C, and their fertility decreases with an increase in temperature (Mcmullen et al., 2012). For example, when *C. elegans* are shifted to 25 °C, they become less fertile (Xiao et al., 2011, Mcmullen et al., 2012). The previous study has reported that loss of SET-2 function results in a further reduction of fertility at 25 °C (Xiao et al., 2011). To investigate the impact of the loss of CFP-1 function in the fertilty at 25 °C, the brood size of *C. elegans* treated with *cfp-1* RNAi was measured. RNAi-mediated knockdown of *cfp-1* resulted in a drastically reduced live brood size by 83% compared to wild-type animals upon RNAi negative control (**Figure 3.3 A**). This suggests that CFP-1 is required for fertility at 25 °C. The role of CFP-1 in maintaining *C. elegans* fertility was further confirmed by using the *cfp-1(tm6369)* mutant. The average brood size of *cfp-1(tm6369)* mutants was reduced by 96% at 25 °C. The level of reduction was similar to the reported *set-2(bn129)* mutants (**Figure 3.3 B**). Additionally, 70% (n = 20) of both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants were sterile. These results suggest that both CFP-1 and SET-2 are required to maintain *C. elegans* fertility at 25 °C.

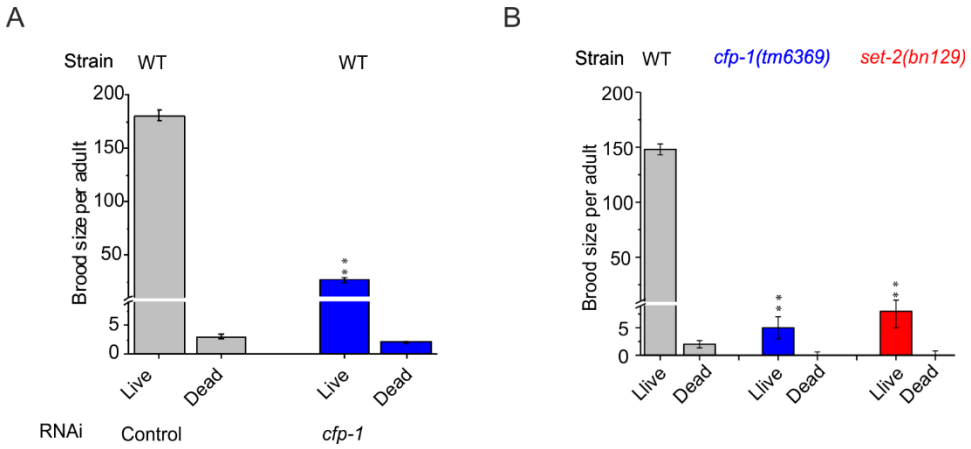


Figure 3.3: *cfp-1(tm6369)* and *set-2(bn129)* display temperature-sensitive sterility

(A) The average brood size (dead and live) of wild-type animals upon *cfp-1* RNAi (blue) and RNAi negative control (control) (grey) at 25 °C. Fertility of *C. elegans* treated with *cfp-1* RNAi (blue) was severely compromised at 25 °C. The figure represents the average of three biological replicates (n = 9 in each experiment). (B) The average brood size of wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants at 25 °C. The brood sizes of *cfp-1(tm6369)* and *set-2(bn129)* mutants was severely compromised with 70% of both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants being sterile. The figure represents the average of two independent experiments (n = 10 in each experiment). P-values were calculated using a student t-test: ** = P<0.01. Error bars represent ± SEM.

3.2.4 CFP-1 promotes *C. elegans* development

To further explore the impact of the loss of CFP-1 function in *C. elegans* development, the *cfp-1* levels were depleted by RNAi and a developmental assay was performed. *C. elegans* embryos pass through four larval development stages (L1, L2, L3 and L4) to reach adulthood (**Figure 3.4 A**). I scored the development stage of the animals using gonad structure (see methods).

RNAi-mediated knockdown of *cfp-1* resulted in a developmental delay: after 60 h, 92% of animals treated with RNAi negative control reached the adult stage, whilst only 45% of *cfp-1* RNAi treated animals reached the adult stage (**Figure 3.4 B**). Similar to the RNAi result, it was observed that the development of *cfp-1*(*tm6369*) mutant animals was delayed compared to the wild-type (**Figure 3.4 C**). Additionally, the development of the *set-2(bn129)* mutant was also delayed (**Figure 3.4 C**). The observed delays in development from embryo to adult upon *cfp-1* RNAi and in both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants suggest that CFP-1 and SET-2 play an important role in the development of the organism.

Together, the present findings confirm that *cfp-1(tm6369)* is a loss of function allele and is phenotypically similar to the *set-2(bn129)* mutant.

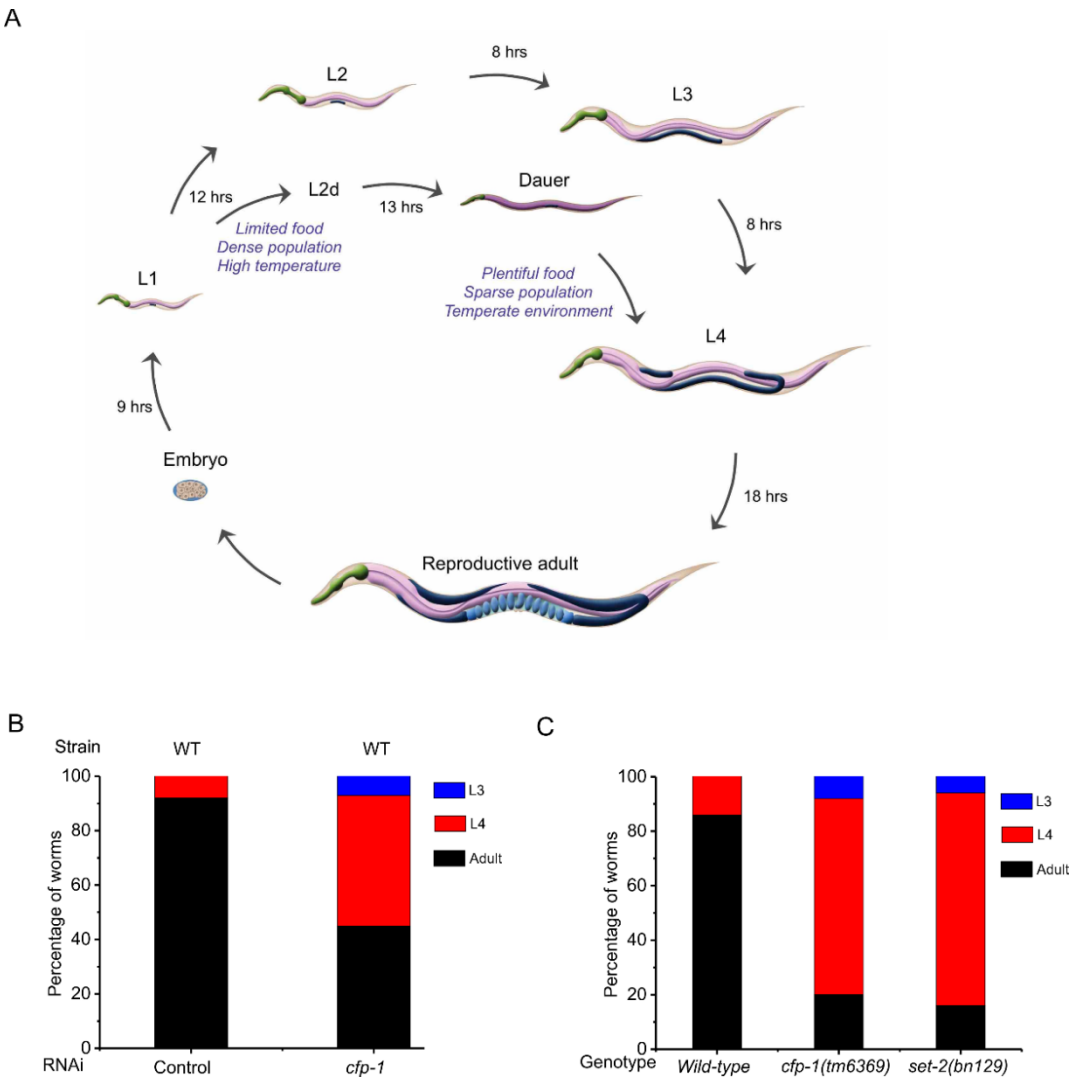


Figure 3.4: *cfp-1(tm6369)* and *set-2(bn129)* mutants development is delayed
(A) The life cycle of *C. elegans.* The generation time of *C. elegans* is only three days. *C. elegans* embryos pass through four larval development stages (L1, L2, L3 and L4) to reach adulthood. The picture was taken from Wolkow (2015). (B) Developmental progress of *C. elegans* grown on *cfp-1* RNAi and RNAi negative control (control) for 60 h after hatching at 20 °C. *cfp-1* RNAi delayed development into adulthood. The figure represents the average of three biological replicates (n > 30 per strain in each experiment. Combined number of animals from three replicates: WT during control RNAi (149) and *cfp-1* RNAi (113)). (C) Developmental assay of *cfp-1(tm6369)* and *set-2(bn129)* mutants compared to wild-type animals after 60 h at 20 °C. After 60 h most of the control animals reached adulthood, whereas both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants displayed

delays in development from an embryo into a young adult. The figure is the average of two independent experiments (n > 30 per strain in each experiment. Combined number of animals from two replicates: WT (172), *cfp-1(tm6369)* (101), *set-2(bn129)* (137)).

3.2.5 Loss of CFP-1 function results in gross chromosomal changes in germ cells.

Loss of function of CFP-1 or SET-2 resulted in a significant reduction of brood size. Previously in our lab, it was observed that 1-2% of *cfp-1(tm6369)* mutants developed spontaneously into males, which is significantly greater compared to ~0.02% males in wild-type animals (Hanouf et al. unpublished data) (Fay, 2006). This high incidence of males (Him phenotype) in the *cfp-1* mutant suggests an increase in chromosomal aberration in germ cells. Similarly, previous studies have suggested that the loss of function of SET-2 results in reduced fertility accompanied by increased chromosomal aberration (Xiao et al., 2011, Herbette et al., 2017, Robert et al., 2014). It is possible that the reduced fertility in the *cfp-1(tm6369)* loss of function mutant may be due to chromosomal abnormalities in germ cells. To investigate whether the poor fertility phenotype in the *cfp-1* mutant is associated with chromosomal aberration, the germ lines of *cfp-1(tm6369)* animals were analysed by DAPI staining. All unfertilised wild-type oocytes (n=136 oocytes) contain six DAPI-stained bodies **Figure 3.5 A**. In the *cfp-1(tm6369)* mutant, 6 % of oocytes (n= 229 oocytes) contained three or four or five DAPI-stained bodies (**Figure 3.5 B-D**), and 2 % of oocytes contained more than six DAPI-stained bodies at diakinesis (**Figure 3.5 E and F**). The presence of fewer or more than expected DAPI-stained bodies in the *cfp-1(tm6369)* mutant indicates chromosomal fusions and fragmentation, respectively. These defects are usually associated with errors in meiosis, suggesting a role of CFP-1 in maintaining meiotic fidelity (Hillers and Villeneuve, 2009, Xiao et al., 2011).

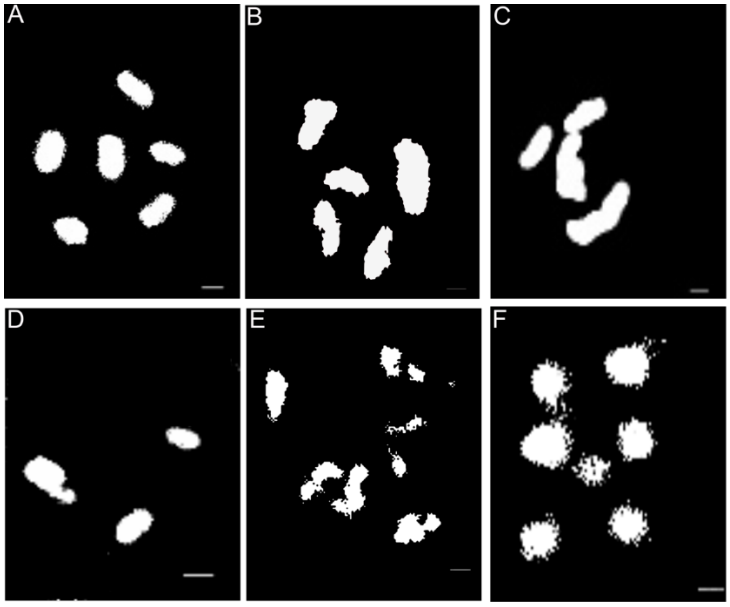


Figure 3.5: CFP-1 deficiency results in defects in chromosome segregation during meiosis

Representative images of DAPI-stained diakinetic chromosomes from oocytes (prior to spermatheca entry) of wild-type animals and the *cfp-1(tm6369)* mutants. (A) Diakinetic nuclei of wild-type animals (n=136 oocytes) showing six DAPI-stained bodies. (B-F) Diakinetic nuclei of *cfp-1(tm6369)* mutants (n=229 oocytes) showing fewer (B-D) or more (E and F) than six DAPI-stained bodies. A Fisher's exact test was used to assess the chromosomal abnormalities in *cfp-1(tm6369)* mutants compared to wild-type animals, and was found to be significant ($p<0.01$). Scale bar, 1 μm.

3.2.6 Reduced fertility in *cfp-1(tm6369)* and *set-2(bn129)* mutants is not due to increased apoptosis in germ cells

Loss of CFP-1 and SET-2 function resulted in gross chromosomal changes in germ cells **Figure 3.5** (Xiao et al., 2011). The increase in genomic instability in these mutants may lead to an increase in the number of apoptotic events in the germ cells and thereby contribute to fertility defects. Therefore, I hypothesised that if the observed drop in brood size of *cfp-1(tm6369)* and *set-2(bn129)* mutants is due to an increase in the apoptosis of germ cells, then the knockdown of apoptotic-inducing genes such as *cep-1* should rescue the brood size.

To explore this, expression of *cep-1,* an orthologue of the p53 gene which promotes DNA damage-induced apoptosis, was reduced by RNAi-mediated knockdown in both the *cfp-1(tm6369)* mutant and the *set-2(bn129)* mutant and brood size was subsequently measured at 20 °C (Hoffman et al., 2014, Schumacher et al., 2001). Interestingly, it was observed that the average live brood size of wild-type animals was significantly reduced upon *cep-1* RNAi (**Figure 3.6**). However, no significant change in the number of dead eggs was observed (**Figure 3.6**). On the other hand, RNAi knockdown of *cep-1* in the *cfp-1(tm6369)* and *set-2(bn129)* mutants did not result in any significant change on the total brood size (live and dead progeny) (**Figure 3.6**). Additionally, RNAi-mediated knock-down of *hus-1*, which acts upstream of *cep-1* and is required for DNA damage-induced apoptosis and cell cycle arrest (Hofmann et al., 2002, Gartner et al., 2008), did not result in a reduced brood size (**Figure 3.6**). Similarly, the average brood size of the *cfp-1(tm6369)* mutant did not change upon RNAi-mediated knockdown of *hus-1* (**Figure 3.6**). On the other hand, the average live brood size of *set-2(bn129)* animals grown on *hus-1* RNAi was increased compared to RNAi negative control (**Figure 3.6**). However, the increment was not statistically significant. Taken together, these finding suggests that a reduction in fertility in *cfp-1(tm6369)* and *set-2(bn129)* is not due to an increase in *cep-1* mediated apoptosis.

Previously, it has been observed that *kri-1* activates DNA damage-dependent cell death which is independent of *cep-1* (Ito et al., 2010). It is possible that there could

be an apoptosis event, independent of *cep-1*. To investigate this, *kri-1* was knocked down by RNAi in *cfp-1(tm6369)* and *set-2(bn129)* mutants, and fertility was assayed. The knockdown of *kri-1* in wild-type animals did not have a significant impact on the average brood size (**Figure 3.6**). Similarly, the average brood size of *cfp-1(tm6369)* and *set-2(bn129)* mutants treated with *kri-1* RNAi remained unchanged (**Figure 3.6**). In conjunction, these findings suggest that the loss of fertility in *cfp-1(tm6369)* and *set-2(bn129)* mutants was not due to an increase in apoptotic events in the germ cells. Consistent with these findings, Herbette et al. (2017) observed that a number of apoptotic cells in *set-2(bn129)* mutants animals was similar to wild-type animals.

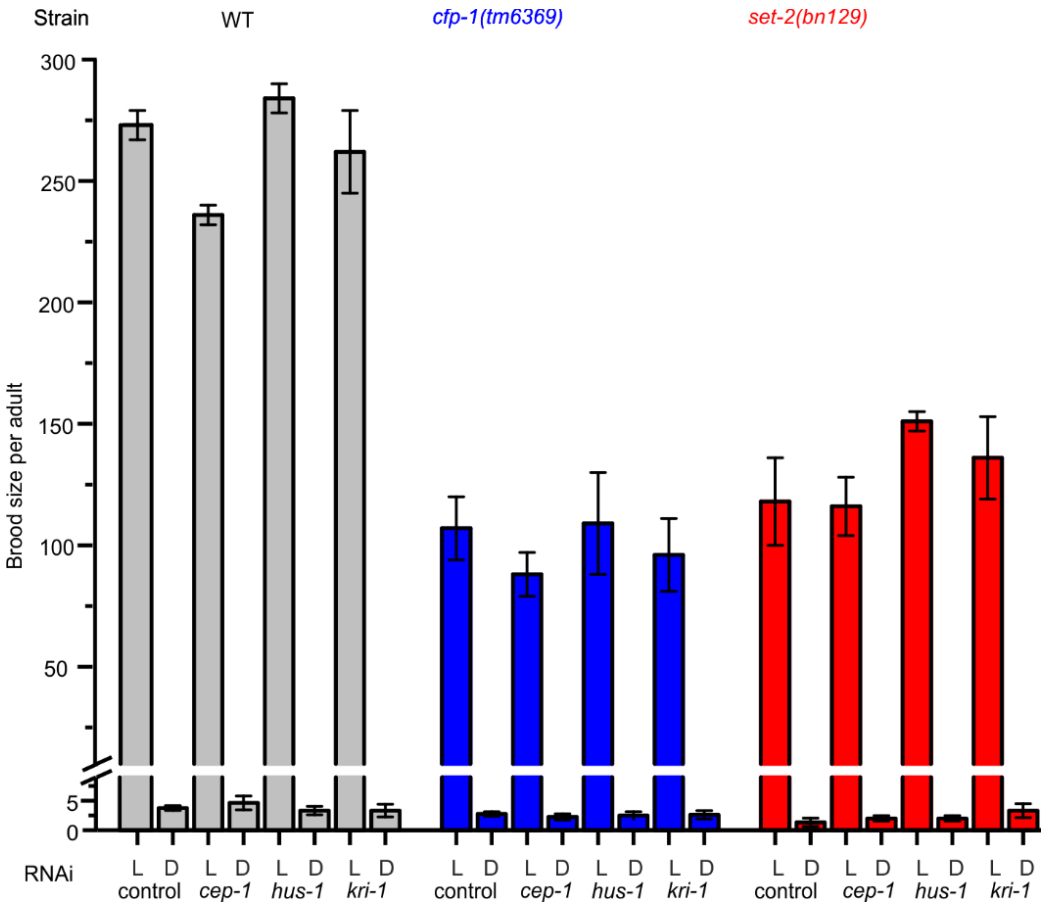


Figure 3.6: Knockdown of key apoptotic genes did not rescue the poor fertility of *cfp-1(tm6369)* and *set-2(bn129)* mutants

The average brood size of wild-type (WT) (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants grown on *cep-1*, *hus-1* or *kri-1* RNAi compared to RNAi negative control (control). The brood size of the *cfp-1(tm6369)* mutant was slightly reduced during RNAi mediated knockdown of *cep-1*. However, the reduction was not statistically significant. *cep-1* or *kri-1* RNAi had no significant impact on the average brood size of *cfp-1(tm6369)* mutants whereas *hus-1* RNAi slightly increased the brood size of *set-2(bn129)* mutants but the increment was not statistically significant. A null hypothesis t-test (refer to the methods section) was used to test the genetic interactions. The figure is the average of two biological repeats (n=10-15 in each repeat). P-values were calculated using a one-tailed student t-test. Error bars represent ± SEM. L= live brood and D=dead eggs.

3.3 Discussion

The importance of Cfp1 in different biological processes is emphasised by various previous research. Cfp1 is a major subunit of the Set1/COMPASS complex, and characterisation studies show that the role of Cfp1 in H3K4me3 deposition is conserved in mammals and yeast (Voo et al., 2000, Lee and Skalnik, 2005, Brown et al., 2017, Clouaire et al., 2012, Ardehali et al., 2011, Shilatifard, 2012, Yu et al., 2017, Clouaire et al., 2014). *Cfp1* depleted mice are not viable, and *Cfp1*$^{-/-}$ embryonic stem cells cannot differentiate (Young and Skalnik, 2007, Carlone and Skalnik, 2001, Carlone et al., 2005). However, the contribution of Cfp1 in development is not clearly elucidated. Here, to investigate the role of Cfp1 in development, a *cfp-1(tm6369) C. elegans* mutant was used. I found that *cfp-1(tm6369)* mutants have dramatically reduced H3K4me3 levels suggesting that the CFP-1 is responsible for the bulk of H3K4me3 deposition. Furthermore, phenotypic characterisation of the *cfp-1(tm6369)* mutant suggests that CFP-1 is required for the fertility and development of *C. elegans*.

In both the *set-2(bn129)* and the *cfp-1(tm6369)* mutants, the level of H3K4me3 was drastically reduced, suggesting that CFP-1 and SET-2 are both required for H3K4me3 modification. I also investigated the impact of the loss of CFP-1 function for fertility by depleting the levels of *cfp-1* by RNAi-mediated knockdown in wild-type animals. Upon *cfp-1* RNAi, the brood size was significantly reduced at 15 °C and 20 °C. Furthermore, fertility was severely compromised at 25 °C. The role of CFP-1 in fertility was further confirmed by using the *cfp-1(tm6369)* genetic mutant. Consistent with the RNAi result, the fertility was reduced dramatically, and the level of reduction was higher in the genetic mutant compared to animals treated with *cfp-1* RNAi. The less severe phenotype upon *cfp-1* RNAi compared to the *cfp-1(tm6369)* mutant could be due to the fact that *cfp-1* RNAi did not result in a complete knockdown. Indeed, *cfp-1* mRNA levels as measured by qRT-PCR were reduced by 74% during *cfp-1* RNAi, thus it is possible that the remaining 26% could partially contribute to CFP-1 function.

The observed reduction in the brood size of both the *set-2(bn129)* and the *cfp-1(tm6369)* mutants combined with the dramatically reduced levels of H3K4me3 suggests a role of H3K4me3 in promoting fertility. Furthermore, previous studies suggest that knockdown of *wdr-5, ash-2* and *rbbp-5*, subunits of the COMPASS complex, which are responsible for H3K4me3 modification, results in a significant reduction in the brood size at 20 °C and 25 °C (Wang et al., 2011b, Xiao et al., 2011). These findings strongly suggest that H3K4me3 is required to maintain fertility.

The observed reduced fertility in *cfp-1* or *set-2* loss of function mutants could be due to the accumulation of heritable cellular damage, which leads to genomic instability. This can be supported by the fact that the loss of CFP-1 or SET-2 function results in gross chromosomal changes (**Figure 3.5** and (Xiao et al., 2011)). Xiao et al. (2011) reported that the *set-2(bn129)* mutant exhibits more or fewer than six DAPI-stained chromosomal bodies. Similarly, in this study, less or more than six DAPI- stained structures were observed in *cfp-1(tm6369)* mutants. A reduced number of DAPI-stained bodies could be due to chromosomal fusions and more than six DAPI-stained bodies could be due to fragmentation (Hillers and Villeneuve, 2009, Xiao et al., 2011). Furthermore, a recent study reported that *set-2(bn129)* mutants are sensitive to DNA damage-causing agents and 90% of *set-2(bn129)* L1 larvae treated with ionising radiation developed into sterile adults (Herbette et al., 2017). This suggests that CFP-1 and SET-2 could function to regulate chromatin structure to maintain meiotic fidelity, and that loss of function of either could result in increased chromosomal aberrations leading to genomic instability in germ cells that could affect fertility.

The reduced genomic stability in *cfp-1* or *set-2* loss of function mutants could potentially increase the apoptotic events resulting in a rediced number of germ cells that are able to fertilise the oocyte. However, the brood size of both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants did not increase in RNAi knockdown of key genes responsible for apoptosis. Furthermore, a number of apoptotic cells in *set-2(bn129)* mutants were similar to wild-type animals (Herbette et al., 2017). These observations suggest that a reduction in brood size of *cfp-1(tm6369)* and *set-*

2(bn129) mutants is not due to an increase in the apoptotic event in germ cells. It is possible that the increase in genomic instability in germ cells results in failure to fertilise rather than apoptosis. As an increase in genomic instability does not always lead to apoptosis (Jaramillo-Lambert et al., 2010).

Alternatively, the reduced brood size could be due to a loss of pluripotency and premature differentiation of germ cells resulting in a reduction in the number of germ cells capable of fertilisation. This is alined with the fact that H3K4me3 maintains the pluripotency of germ cells by repressing the somatic gene expression in the *C. elegans* germline (Robert et al., 2014, Cui et al., 2006b, Wang et al., 2011b). Similarly, in mammals, H3K4me3 maintains the pluripotency of germ cells by maintaining the late developmental genes in a silenced state poised for activation (Lesch and Page, 2014, Sachs et al., 2013). Silenced state genes have a bivalent chromatin domain: chromatin that has both active (H3K4me3) and repressive (H3K27me3) marks. The loss of the bivalent domain results in a loss of pluripotency of the cells (Lesch and Page, 2014). Together, these findings suggest that a loss of H3K4me3 could change the germline transcriptional program and result in the precocious expression of late development genes or expression of somatic genes in *cfp-1(tm6369)* and *set-2(bn129)* germline. This is most likely to contribute to the observed loss of fertility.

I also observed that the depletion of *set-2* or *cfp-1* results in a development delay. After 60 h, most of the wild-type animals reached the adult stage whereas most of the *cfp-1(tm6369)* and *set-2(bn129)* mutants were in the L4 stage. The observed delay in development could be due to the reduced or late expression of genes required for stage transitions. Further research will shed light on the actual mechanism behind the observed development delay.

Chapter 4

Investigating the role of CFP-1 and SET-2 in gene induction

4.1 Introduction

The trimethylation of histone 3 lysine 4 (H3K4me3) is a well-recognised "active promoter" chromatin mark, frequently found in active promoter regions. The loss of H3K4me3 can lead to developmental defects in both mammals and *C. elegans*, indicating that H3K4me3 plays a key role in epigenetic control during metazoan development (Shilatifard, 2012, Shilatifard, 2008, Xiao et al., 2011). Therefore, it is important to understand how H3K4me3 is regulated, and how this controls gene expression in metazoan development. H3K4me3 is mainly deposited by the major H3K4me3 methyltransferase, called SET-2 in *C. elegans* and Set1A in mammals, which is a part of the pan-eukaryotic Set1/COMPASS complex that places H3K4me3 modifications at active promoters (Santos-Rosa et al., 2002, Barski et al., 2007, Shilatifard, 2012, Bernstein et al., 2002, Xiao et al., 2011, Ardehali et al., 2011). In addition to H3K4me3 methyltransferase, all the other subunits of the pan-eukaryotic Set1/COMPASS complex are also conserved in *C. elegans* (Shilatifard, 2012). Within the complex, CXXC1/CFP-1 is a key subunit essential for H3K4me3 modifications (Shilatifard, 2012).

CFP1 is a conserved epigenetic regulator that binds to unmethylated CpG-rich DNA sequences known as CpG islands (CGIs) (Chun et al., 2014, Chen et al., 2014, Thomson et al., 2010). CFP1 helps in the recruitment of the SET1/COMPASS complex at the promoter region of active genes (Clouaire et al., 2012, Tate et al., 2010). In mammals, CFP1 is reported to be important for cell differentiation and cell fate specification (Clouaire et al., 2014, Mahadevan and Skalnik, 2016). Using *Cxxc1* conditional knock-out mES, Brown *et al.* recently reported that Cxxc1 plays a role in shaping the context-dependent gene expression in mice (Brown et al., 2017). However, the molecular principle of CFP-1's role in gene expression regulation is yet to be better elucidated at the organismal level.

Several genome-wide analyses have been carried out in yeast, *Drosophila* and mammals to understand the interplay between H3K4me3 and gene transcription. These genome-wide studies have shown that the relative level of H3K4me3 is strongly correlated with the gene expression level, suggesting that H3K4me3 can promote gene expression (Howe et al., 2017, Guillemette et al., 2011, Dong and Weng, 2013, Araki et al., 2009). Despite H3K4me3 being associated with gene expression, loss of H3K4me3 however, results in very little change in global gene expression level (Howe et al., 2017, Clouaire et al., 2014). Most of these studies were conducted in steady-state conditions.

In contrast, single gene studies of chromatin regulators including COMPASS subunits suggest that chromatin regulators play important roles in a dynamic process that are masked in steady state conditions (Weiner et al., 2012). For example, H3K4me3 plays a repressive role in galactose-inducible *GAL* genes (Margaritis et al., 2012, Zhou and Zhou, 2011). As both *cfp-1(tm6369)* and *set-2(bn129)* mutants exhibit low H3K4me3 levels, I wanted to understand the transcriptional role of H3K4me3 during stress conditions in these mutants. Heat and osmotic stress were chosen, as these have been shown to play a role in dramatic and rapid change in gene expression in *C. elegans* (Lamitina et al., 2004, Lamitina et al., 2006, Choe and Strange, 2007, Kultz, 2007, Brunquell et al., 2016, Li et al., 2016b).

During heat stress, heat shock transcription factor 1 (HSF1) binds to the heat shock response elements and upregulates the transcription of heat shock chaperones such as heat shock protein (Hsp)90, Hsp70 and the small HSPs (Hentze et al., 2016, Gomez-Pastor et al., 2018, Labbadia and Morimoto, 2015). These chaperones help to maintain protein structure by maintaining unfolded proteins in non-native states, which can refold to the final native states (Hentze et al., 2016, Gomez-Pastor et al., 2018, Ritossa, 1962).

Similar to heat shock response, when the salt concentration is increased (hypertonic), the body starts to produce osmolytes to prevent water loss, and failure to do so results in cell shrinkage and protein aggregations (Lamitina et al., 2004, Choe and Strange, 2007, Kultz, 2007). When there is low salt concentration, the

body releases the osmolyte and also reduces its production to prevent excessive water uptake. For example, when *C. elegans* are treated under hypertonic conditions, the expression level of the *gpdh-1* gene is significantly upregulated (Lamitina et al., 2004, Lamitina et al., 2006). The *gpdh-1* gene encodes a glycerol- 3- phosphate dehydrogenase (GPDH-1), which induces the synthesis of the glycerol (Lamitina et al., 2004, Choe and Strange, 2007). Accumulation of the glycerol in cells increases the osmolyte concentration and prevents fluid loss (Lamitina et al., 2004, Lamitina et al., 2006).

In Chapter 3, I observed that both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants have a drastic reduction in the H3K4me3 levels. Furthermore, it was observed that CFP-1 and SET-2-dependent H3K4me3 plays an essential role in maintaining the fertility and the development of the whole organism. Here, in this Chapter, I observed that the induction of the salt-inducible reporter gene was significantly higher in both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants compared to wild-type animals. Furthermore, the role of H3K4me3 in gene induction was confirmed by measuring the endogenous gene expression of the salt-inducible gene and heat shock chaperones. Collectively, findings from this Chapter suggest that H3K4me3 may play a repressive role in gene induction.

4.2 Results

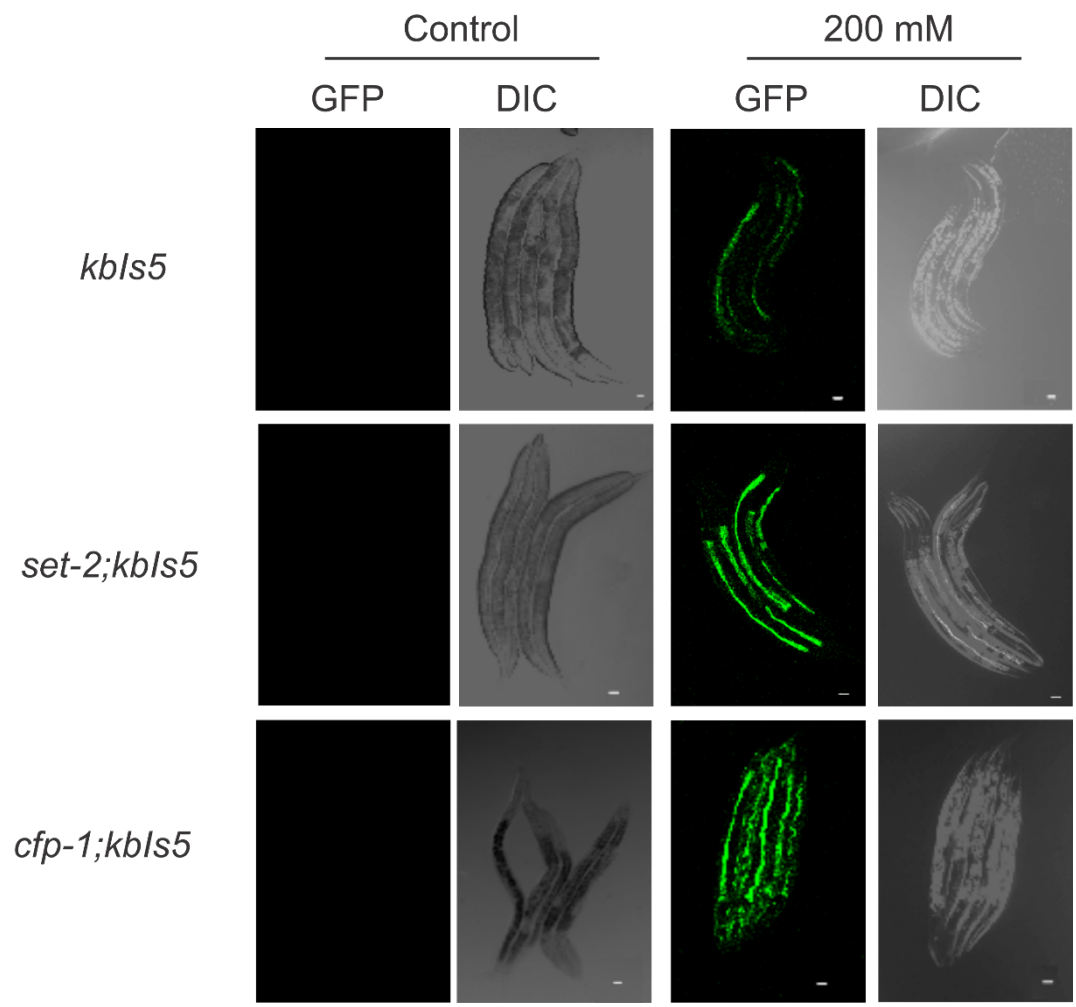
4.2.1 A salt-inducible reporter gene is strongly expressed in *cfp-1(tm6369)* and *set-2(bn129)* mutants

To investigate the role of H3K4me3 in gene regulation, a salt-inducible reporter strain (VP198 (*kbIs5* [*gpdh-1p::GFP + rol-6(su1006)*]), which contains the green fluorescent protein (GFP) reporter gene downstream of the salt-inducible gene *gpdh-1* promoter was used (Lamitina et al., 2004, Lamitina et al., 2006). This reporter strain was genetically outcrossed four times into the wild-type *C. elegans* (N2 Bristol) to remove any background mutations. After outcrossing, the reporter strain was genetically crossed into the background of *cfp-1(tm6369)* and *set-2(bn129)* mutant strains to generate *cfp-1(tm6369);kbIs5* and *set-2(bn129);kbIs5* mutants. *C. elegans* is normally grown on a salt (NaCl) concentration of 52 mM. Thus 52 mM was used as a control throughout this study.

gpdh-1p::GFP reporter strains harbouring either the *cfp-1(tm6369)* or the *set-2(bn129)* mutant allele were treated with salt (NaCl) concentrations of either 52 mM (control) or 200 mM (stress condition) and GFP expression of the reporter was monitored under a fluorescence microscope. At salt concentrations of 52 mM, the salt-inducible GFP reporter was not expressed in either mutant strain background tested, *cfp-1(tm6369)* or *set-2(bn129)*. Interestingly, at 200 mM salt concentration, expression of *gpdh-1p::GFP* in both *cfp-1(tm6369);kbIs5* and *set-2(bn129);kbIs5* mutants was induced compared to *kbIs5* control animals (**Figure 4.1**). In both mutants, GFP was strongly visible in the intestine and hypodermis compared to *kbIs5* animals (**Figure 4.1 A**). As shown in the **Figure 4.1 B,** the average intensity of GFP expression was increased 3-fold in *cfp-1(tm6369);kbIs5* and 2.3-fold in *set-2(bn129);kbIs5* mutants when compared to the *kbIs5* animals.

Thus, the level of stress-inducible *gpdh-1* gene expression is increased in the *cfp-1* and *set-2* mutants.

A



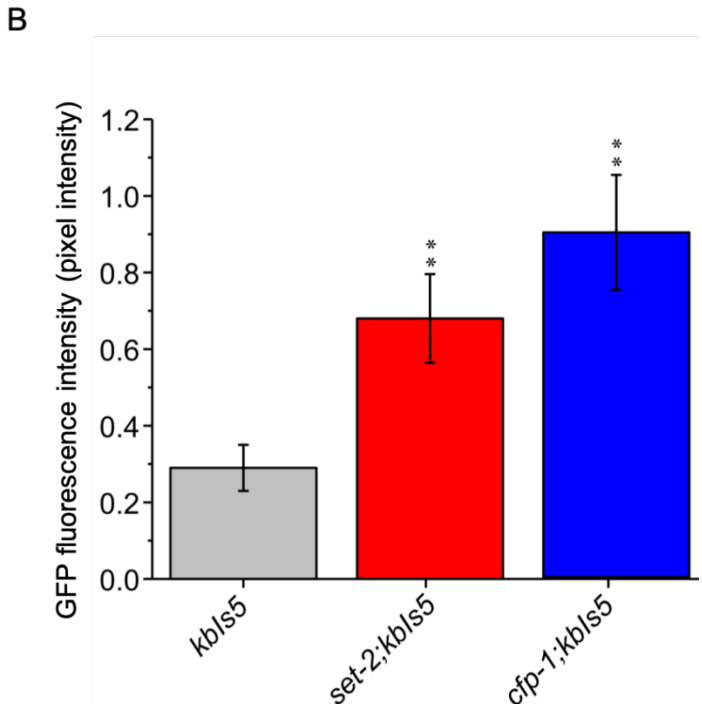


Figure 4.1: Loss of CFP-1 or SET-2 function results in stronger expression of GFP protein

(A) Representative images of *kbIs5, cfp-1;kbIs5, set-2;kbIs5* animals expressing GFP under the control of the *gpdh-1* promoter, during control conditions (52 mM salt) and osmotic stress (200 mM salt concentration). Differential interference contrast (DIC) Nomarski and GFP images. Expression of GFP was higher in *cfp-1;kbIs5* and *set-2;kbIs5* animals compared to *kbIs5* animals. Scale bar 20 µm. (B) Quantification of average GFP fluorescence intensity in *kbIs5, cfp-1;kbIs5, set-2;kbIs5* animals expressing *gpdh-1p::GFP* during osmotic stress compared to the control. For the image quantification, I performed a basic pixel intensity analysis by ImageJ, which quantifies grey levels (corresponding to GFP intensity levels) across a selected area of the image relative to background. Average GFP fluorescence intensity per animal in *cfp-1;kbIs5* and *set-2;kbIs5* mutants is significantly higher compared to the *kbIs5.* The figure shows an average of three biological replicates (n > 15). P-values were calculated using student t-test: **= P<0.01. Error bars represent ± standard error of the mean (SEM).

4.2.2 The *gpdh-1* gene is strongly induced in *cfp-1(tm6369)* and *set-2(bn129)* mutants

To further confirm the observation that salt-inducible expression of *gpdh-1* was increased in the *set-2* and *cfp-1* mutants, endogenous transcript levels of *gpdh-1* were measured using qPCR. For efficient qPCR, and to avoid genomic DNA contamination, primers were designed that bind to the exon-exon junction, i.e. one half of the primer hybridises to the 3′ end of one exon and another half of the primer binds to the 5′ site of an adjacent exon. Primers were optimised by changing the concentration of primer, template and reaction mixture to achieve more than 90% efficiency (see methods).

Wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants were grown on control plates (containing 52 mM salt) or plates containing a higher concentration of salt (150 mM). Animals treated with 52 mM salt were used as a control, and *tba-1* and *pmp-3* genes were used as reference genes as the expression levels of *tba-1* and *pmp-3* are stable compared to other housekeeping genes in *C. elegans* (Zhang et al., 2012). Interestingly, mRNA transcript levels of *gpdh-1* in *cfp-1(tm6369)* and *set-2(bn129)* mutants were already induced during the control condition and significantly higher than measured in control wild-type animals (**Figure 4.2 and Appendix 9.3**). During salt stress, the levels of *gpdh-1* transcripts were further induced: 7-fold in wild-type animals, 10-fold in the *set-2* mutant and 9-fold in *cfp-1* mutants compared to control animals (**Figure 4.2**). Thus, the expression level of *gpdh-1* transcripts in the mutants remained higher than compared to wild-type animals, even after exposure to salt stress.

This result suggests that the expression of the salt-inducible gene *gpdh-1* is increased in the *cfp-1* and *set-2* mutants.

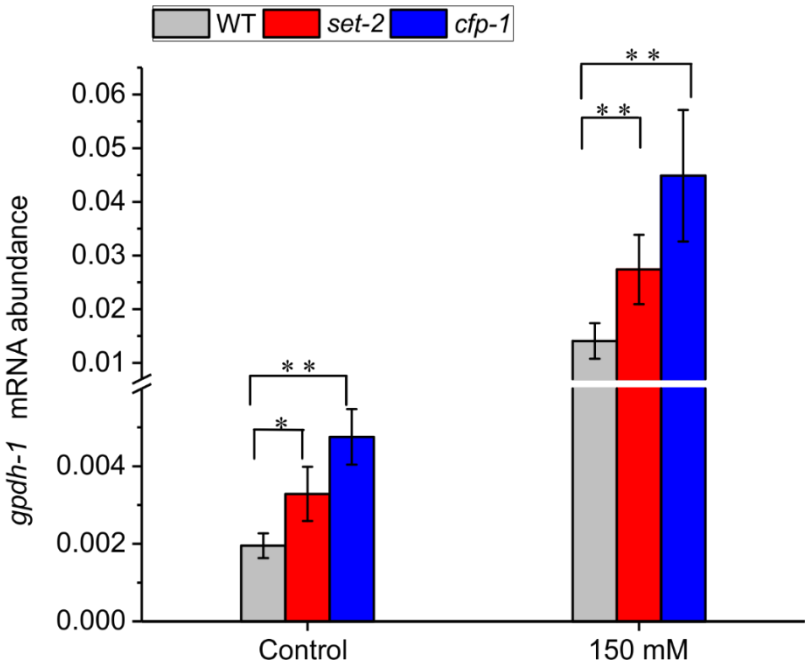


Figure 4.2: The expression of *gpdh-1* is higher in *cfp-1(tm6369)* and *set-2(bn129)* mutants

gpdh-1 transcript levels as measured by qPCR during control (52 mM salt) conditions and osmotic stress (150 mM) in wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants. *pmp-3* and *tba-1* genes were used for normalisation. During osmotic stress (150 mM), *gpdh-1* expression levels remained higher in *cfp-1(tm6369)* and *set-2(bn129)* mutants than wild-type animals. The bar graph is the average of four biological replicates (n=130-150 in each experiment). P-values were calculated using a student t-test: * = P<0.05, ** = P<0.01. Error bars represent ± standard error of the mean (SEM).

4.2.3 Loss of CFP-1 or SET-2 function results in increased transcription of heat inducible genes

I observed that CFP-1 and SET-2 regulate the induction of a salt-inducible gene. It is possible that they may also regulate other stress responses. To further investigate the role of CFP-1 and SET-2 in stress gene induction, the mRNA expression levels of heat-inducible genes, *hsp-70* (*C12C8.1 and F44E5.4*) and *hsp-16.2*, were measured in wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants. *C12C8.1, hsp-16.2* and *F44E5.4* are heat inducible chaperones and are expressed during heat stress (Prahlad et al., 2008, Snutch et al., 1988, Jones et al., 1986, Brunquell et al., 2016). Similar to the *gpdh-1* gene, primers that bind to the exon-exon junction were designed and optimised by changing the concentration of primer, template and reaction mixture to achieve greater than 90% efficiency.

Wild-type, *cfp-1(tm6369)* and *set-2(bn129)* animals were heat shocked at 33 °C for 1 h, and the endogenous mRNA transcript levels of *C12C8.1, hsp-16.2* and *F44E5.4* genes were measured. Non-heat shocked animals were used as a control, and *tba-1* and *pmp-3* were used as reference genes. Interestingly, expression levels of all three chaperons were significantly upregulated in *cfp-1(tm6369)* and *set-2(bn129)* animals before heat shock (**Figure 4.3 and Appendix 9.4**). After heat shock, transcript levels of *C12C8.1, hsp-16.2* and *F44E5.4* were strongly induced in *cfp-1(tm6369)* and *set-2(bn129)* mutants compared to wild-type animals (**Figure 4.3 and Appendix 9.4**). The transcript levels of *C12C8.1* were induced 167-fold in wild-type animals, 260-fold in the *set-2(bn129)* mutant and 454-fold in *cfp-1(tm6369)* mutants, compared to non-heat shock controls. Similarly, the mRNA expression of *F44E5.4* was induced 798-fold in wild-type animals, 953-fold in the *set-2(bn129)* mutant and 2100-fold in the *cfp-1(tm6369)* mutant, compared to non-heat shock controls. The expression of *hsp-16.2* was induced 61-fold in wild-type animals, 128-fold in the *set-2(bn129)* mutant and 110-fold in the *cfp-1(tm6369)* mutant, compared to non-heat shock controls (**Figure 4.3**).

These findings illustrate that similar to a salt inducible gene, expression of heat shock chaperones is strongly induced in *cfp-1(tm6369)* and *set-2(bn129)* mutants compared to wild-type animals.

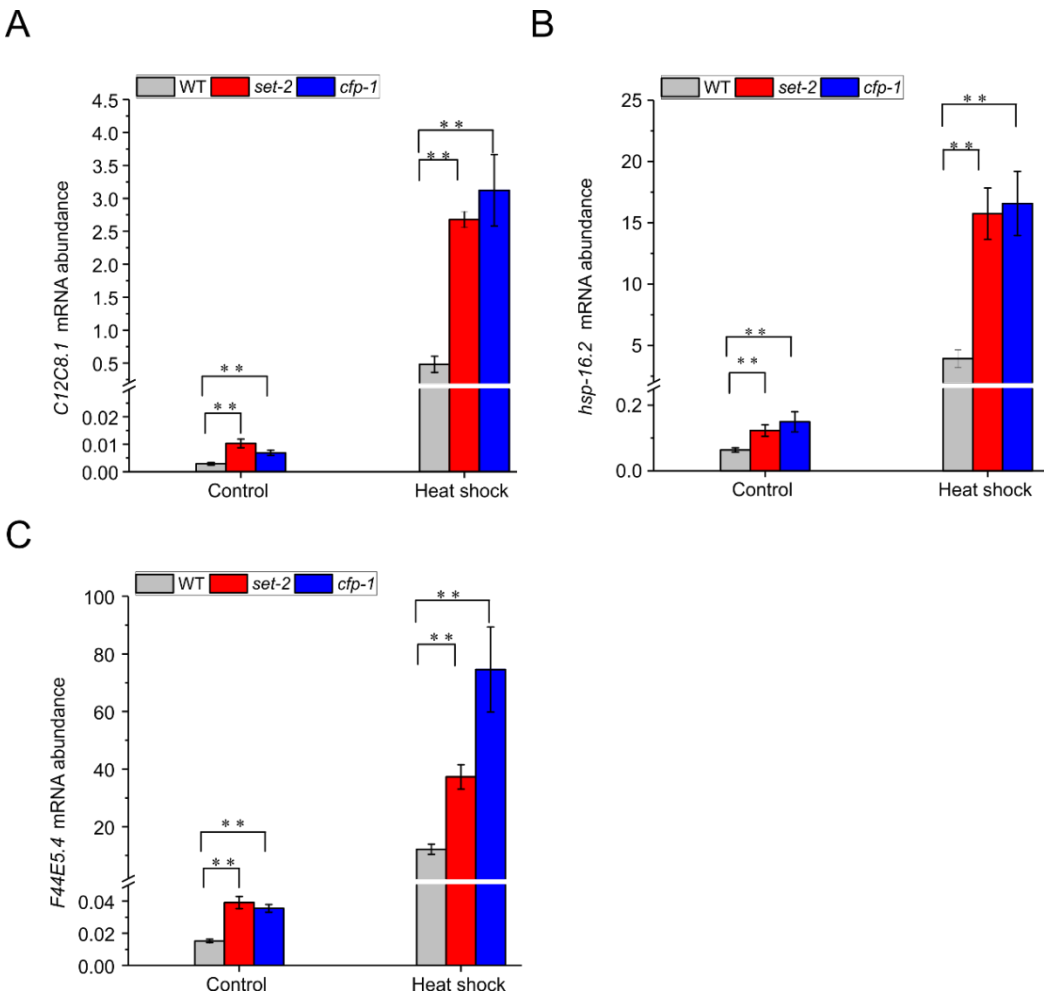


Figure 4.3: Heat shock response genes are strongly expressed in *cfp-1(tm6369)* and *set-2(bn129)* mutants.

qPCR of transcripts of heat shock genes, (A) *C12C8.1*, (B) *hsp-16.2* and (C) *F44E5.4* in wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants after heat shock at 33 °C for 1 h. After heat shock, the expression levels of *C12C8.1, hsp-16.2* and *F44E5.4* were significantly higher in both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants compared to wild-type animals. The bar graph is a combined average of three biological replicates (n=130-150 in each experiment). P-values were calculated using a student t-test: **= P<0.01. Error bars represent ± SEM.

4.3 Discussion

H3K4me3 is commonly known as an active promoter mark and is associated with active genes. The role of H3K4me3 in gene expression is still debatable. Previous studies have shown that the level of H3K4me3 is correlated with an expression of a subset of genes (Howe et al., 2017, Guillemette et al., 2011, Dong and Weng, 2013, Araki et al., 2009). However, recent findings in yeast have pointed towards the repressive role of H3K4me3 in gene expression (Margaritis et al., 2012, Zhou and Zhou, 2011, Lorenz et al., 2014, Lorenz et al., 2012). Additionally, in *C. elegans*, H3K4me3 represses somatic gene expression in the germline to maintain the pluripotency of germ cells (Robert et al., 2014, Cui et al., 2006b). Consistent with the repressive role of H3K4me3 in the germline, in this study, it was observed that in both the *cfp-1(tm6369)* and the *set-2(bn129)* mutants the induction of heat and salt stress response genes was significantly higher than in wild-type animals. This suggests that H3K4me3 could play a repressive role in gene induction (**Figure 4.4**).

Data presented here show that endogenous gene expression levels of the *gpdh-1* gene were induced in the *cfp-1(tm6369)* and *set-2(bn129)* mutants during control and osmotic stress conditions. While no induction of the *gpdh-1p::GFP* reporter in *cfp-1;kbIs5* and *set-2;kbIs5* mutants was detected during control conditions, GFP expression was clearly induced during osmotic stress and was also higher in *cfp-1;kbIs5* and *set-2;kbIs5* animals than in *kbIs5* animals. The observed difference in reporter assay and qPCR could be credited to the higher sensitivity of measuring endogenous mRNA transcript levels by qPCR as opposed to GFP fluorescence intensity.

The expression of stress inducible genes (*gpdh-1*, *C12C8.1, hsp-16.2* and *F44E5.4*) were already induced during control conditions in the *set-2(bn129)* mutants and *cfp-1(tm6369)* mutants. This suggests that CFP-1 and SET-2 repress transcription of stress inducible genes under control conditions. The expression levels during stress conditions (osmotic or heat stress) were also strongly induced in both mutants compared to wild-type animals. It is possible that during stress, the transcription of stress inducible genes results in an increase in deposition of H3K4me3 at the

promoter region which could contribute to the repression of gene induction in the presence of functional CFP-1 and SET-2.

The observed hyper-induction of salt and heat-inducible genes in the *cfp-1(tm6369)* and *set-2(bn129)* mutants could be due to the pausing of RNA polymerase II (Pol II). Promoter pausing is one of the regulatory mechanisms to control gene transcription (Teves and Henikoff, 2013). Poised promoters are primed for transcription activation in response to a stimulus (Teves and Henikoff, 2013). The presence of paused Poll II helps to maintain the accessible promoter region that can be bound by transcription factors and activators resulting in efficient activation of genes (Teves and Henikoff, 2013, Shopland et al., 1995, Samarakkody et al., 2015). Studies in *D. melanogaster* show that heat shock response genes also contain the paused Pol II downstream of the transcription start sites (TSS) even in uninduced conditions (Teves and Henikoff, 2013, Shopland et al., 1995, Samarakkody et al., 2015). Furthermore, Shopland et al. (1995) showed that Pol II pausing is required for gene activation during heat stress, and loss of Pol II pausing results in decreased binding of HSF-1 at heat shock protein promoter regions. Paused regions have an open chromatin structure, and H3K4me3 is often found at the promoter of active genes. H3K4me3 could act as a regulator to maintain paused Pol II and prevent the burst of transcription before induction. This could be a reason for the increased expression of stress-inducible genes (*gpdh-1*, *C12C8.1, hsp-16.2* and *F44E5.4*) in *cfp-1(tm6369)* and *set-2(bn129)* mutants.

Alternatively, H3K4me3 could act as a binding site for repressor complexes to block transcription and the loss of H3K4me3 could result in a reduced recruitment of repressor complexes, thus facilitating transcription. This could be supported by the fact that in yeast, H3K4me2/3 represses subsets of gene expression by recruiting repressor complexes (Margaritis et al., 2012, Zhou and Zhou, 2011, Lorenz et al., 2014, Lorenz et al., 2012). For example in *Saccharomyces cerevisiae*, H3K4me2/3 represses the *GAL1* gene induction by recruiting histone deacetylase complex called RPD3S, and also act as a memory to repress the *GAL1* reactivation by recruiting Isw1 ATPase, which limits the Pol II activity (Margaritis et al., 2012, Zhou and Zhou, 2011). Similar to *S. cerevisiae*, in *S. pombe* Set1C interacts with Clr3; class II

histone deacetylase to repress a subset of genes, including stress response genes (Lorenz et al., 2014). Since HDACs and Isw1 ATPase are also conserved in *C. elegans,* their function could also be conserved.

CFP-1 and SET-2 are required for bulk H3K4me3 modification which is associated with various biological events including gene expression (Howe et al., 2017, Guillemette et al., 2011, Dong and Weng, 2013, Araki et al., 2009). It is also possible that the observed hyper-induction of stress inducible genes in both the *cfp-1(tm6369)* mutant and the *set-2(bn129)* mutant could be due to an indirect effect of loss of H3K4me3. The loss of H3K4me3 could result in a decrease in expression of genes which play a repressive role in gene induction or increase in expression of genes which contribute to the systematic enhancement of resistance to stress (Ermolaeva et al., 2013). Alternatively, the observed stronger induction of heat and salt inducible genes in *cfp-1(tm6369)* and *set-2(bn129)* mutants could be due to activation of other biological pathways which can contribute to the hyper-induction phenotype. For example, Ermolaeva et al. (2013) have reported that the DNA-damage and innate immunity pathways mediated by MPK-1/MAPK can contribute to the activation of stress response pathways. Chapter 3 shows that the loss of *cfp-1* results in increased chromosomal aberrations in germ cells. It is likely that the DNA-damage pathways may be activated in the loss of function of CFP-1, which then contribute to the increased expression of stress-inducible genes. Another hypothesis for the observed hyper-induction phenotype could be that CFP-1 and SET-2 may play an important role in maintaining the structure of chromatin by interacting with other chromatin regulators, and the loss of function of CFP-1 or SET-2 results in an increase in chromatin accessibility. Therefore, the loss of CFP-1 or SET-2 function could increase the rate of recruitment of HSF-1 or other transcription factors, thereby contributing to the hyper-induction.

Recently, Labbadia and Morimoto (2015) found that *C. elegans* germline stem cells (GSCs) play a repressive role in the stress response by providing a signal for the H3K27 demethylase *jmjd-3.1*. The signal from GSCs reduces the activity of *jmjd-3.1* at heat shock response genes, resulting in an increase in the H3K27me3 repressive mark. Their study suggests that the heat shock response starts to decrease

at day 1 of adulthood when egg laying starts. GSC number continues to increase after the onset of egg laying. They observed that reducing GSC number upholds a heat shock response, preserves the expression of *jmjd-3.1* and enhances the stress resistance during adulthood. They also found that this repressive role of GSCs is not confined to the heat shock response, as they observe similar repressive role in other stress responses including osmotic stress (Labbadia and Morimoto, 2015). In Chapter 3, it was reported that both the *cfp-1(tm6369)* and *set-2(bn129)* have a fertility defect at 20 °C and the defect is more severe at 25 °C. The hyper-induction phenotype observed in *cfp-1(tm6369)* and *set-2(bn129)* mutants could be due to deformed germline which might have reduced number of GSCs.

The increased expression of heat and salt inducible genes observed in this study suggests that H3K4me3 may play a repressive role in gene expression. However, it is possible that the main role of H3K4me3 is not in gene regulation, rather it may play a role in fine tuning the gene expression. Stress inducible genes are either not expressed or expressed at low levels before induction, but are strongly induced during stress. H3K4me3 may play a role in repressing the expression of stress inducible genes during normal growth conditions, therefore loss of H3K4me3 results in increased expression of stress inducible genes under control condition. Similarly, during stress, H3K4me3 may repress the transcription of stress inducible genes to some degree, so as to avoid an “over-induction” of stress response pathways.

In summary, in this study, I observed that the expression of heat- and salt-inducible genes were upregulated in *cfp-1* and *set-2* mutants. This suggests that CFP-1 and SET-2 may play a repressive role in gene regulation. However, findings from this study cannot distinguish whether the CFP-1 and SET-2 play a direct or indirect role in the transcription of stress inducible genes. Further research is required to understand the mechanistic action of CFP-1 and SET-2 in gene induction.

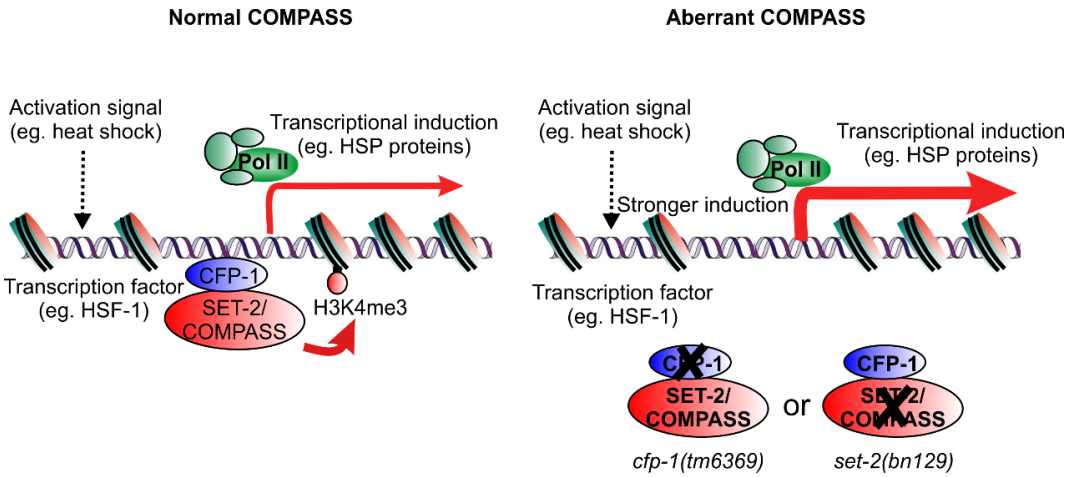


Figure 4.4: Model showing regulation of gene induction by H3K4me3

The SET-2/COMPASS complex deposits H3K4me3 modifications near the promoter of stress-inducible genes (e.g. heat shock response genes). When there is normal SET-2/COMPASS complex, the expression of stress-inducible genes is normal. Loss of H3K4me3 due to deletion of *cfp-1* or *set-2* cause stronger induction of stress-inducible genes.

Chapter 5

SET-2/COMPASS independent function of CFP-1 in vulval development

5.1 Introduction

Cfp1 is an evolutionarily conserved key subunit of the Set1/COMPASS complex which catalyses the H3K4me3 modification at promoter regions (Lee and Skalnik, 2005, Ardehali et al., 2011, Shilatifard, 2012, Yu et al., 2017, Clouaire et al., 2014, Clouaire et al., 2012, Thomson et al., 2010). Cfp1 plays an essential role in vertebrate development; for example, mouse embryos lacking *Cfp1* (*Cfp1* $^{-/-}$) exhibit a pre-implantation death, and they fail to gastrulate (Carlone and Skalnik, 2001, Carlone et al., 2005). Furthermore, injection of *cfp1* transcript-specific antisense oligonucleotides in zebrafish resulted in shorter embryos, failure of primitive haematopoiesis, and cardiac oedema (Young et al., 2006). Additionally, knockdown of the *Cfp1* gene in adult mice leads to haematopoiesis failure, dramatic loss of lineage-committed progenitors, and death within two weeks (Chun et al., 2014). *Cfp1* $^{-/-}$ mice embryonic stem cells (ESCs) are viable, but they are unable to differentiate and exhibit a longer doubling time compared to wild-type ESCs (Carlone and Skalnik, 2001, Carlone et al., 2005). Upon removal of leukaemia inhibitory factor for the induction of differentiation, *Cfp1* $^{-/-}$ ESCs fail to induce lineage-restricted markers and fail to downregulate pluripotency markers of stem cells, such as alkaline phosphatase and Oct4 (Carlone and Skalnik, 2001, Carlone et al., 2005). Taken together, previous data suggest that Cfp1 is required for cell fate specification and pre- and post-embryonic development. Importantly, the early death of *Cfp1* knocked down mouse embryos prevents an assessment of *Cfp1* during development and cell fate specification. *Cfp1* is also conserved in *C. elegans*. Furthermore, *cfp-1* depleted *C. elegans* are viable. Previous studies have reported that *cfp-1* plays a role in vulval development (Cui et al., 2006b, Simonet et al., 2007).

In this study, vulval development in *C. elegans* was used to investigate the role of CFP-1 in organogenesis and to identify its genetic regulators (please see section 1.11 for detailed information regarding vulval induction). Three signalling pathways, Notch, Receptor Tyrosine Kinase (RTK)/Ras/Mitogen-activated Protein Kinase (MAPK) and Wnt, interact with each other for the proper development of vulva (Cui et al., 2006b, Harrison et al., 2006). Mutations that eliminate or reduce the function of RTK/Ras/MAPK signalling results in vulvaless (Vul) animals. In contrast, a mutation that increases the activity of RTK/Ras/MAPK signalling results in the aberrant development of ectopic vulva at the ventral surface of the animals called the multivulva (Muv) phenotype (**Figure 5.1**). A large set of genes, collectively known as Synthetic Multivulva (SynMuv) genes, antagonise the activity of RTK/Ras/MAPK signalling and prevent the aberrant development of ectopic vulva (Cui et al., 2006b, Fay and Yochem, 2007). Animals with a mutation in SynMuv genes from any two classes have a Muv phenotype called synthetic multivulva (please see section 1.11.3 for detailed information regarding SynMuv genes and classes). In a SynMuv suppressor screen, it was found that a number of genes, including *cfp-1,* suppress the SynMuv phenotype (Cui et al., 2006b).

Chapter 3 shows that both CFP-1 and SET-2 are required for H3K4me3, and both play an essential role in fertility and development. Here, results show that CFP-1 but not SET-2 promotes the multivulva phenotype in SynMuv mutants, suggesting the non-overlapping function of CFP-1 and SET-2 for vulval development. Furthermore, a small RNAi screen was conducted to identify genes that could either enhance or repress the role of *cfp-1* in the SynMuv phenotype. I found that *cfp-1* interacts genetically with other chromatin regulators to regulate the multivulva phenotype in SynMuv mutants.

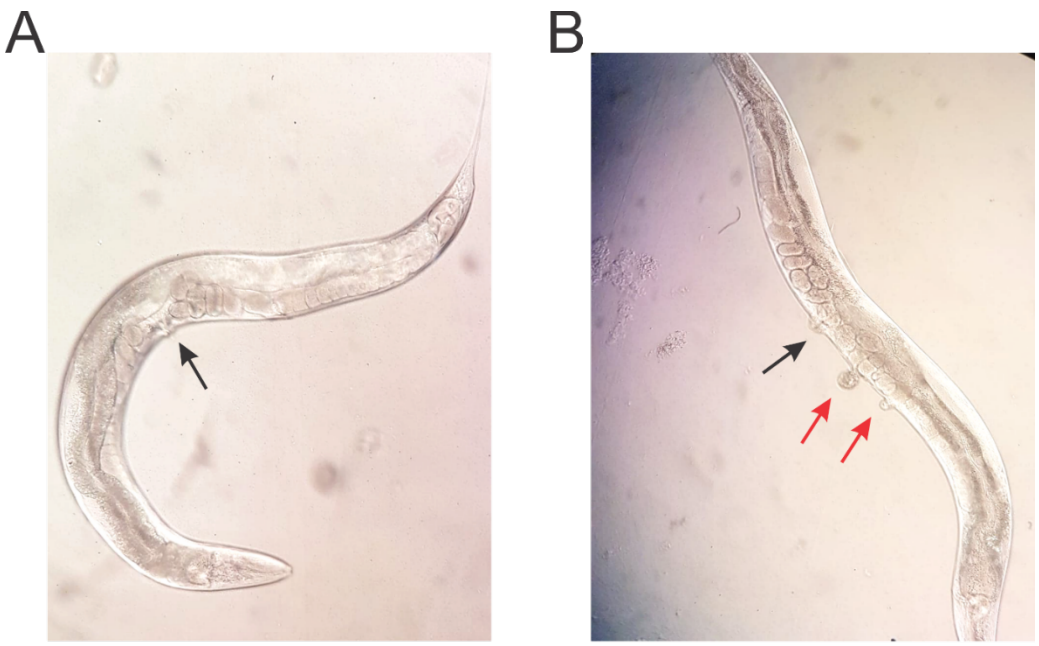


Figure 5.1: Synthetic multivulva phenotype

(A) Wild-type animals have one vulva (black arrow). (B) SynMuv mutants have one normal vulva (black arrow) and one or more vulva like protrusions (red arrows).

5.2 Results

5.2.1 CFP-1 promotes vulval induction in a SynMuv mutant

To investigate the role of CFP-1 in VPC fate decision, a *lin-15AB*(*n765*) strain was used. It contains an insertion that destroys the open reading frame of genes *lin-15A* and *lin-15B* (Cui et al., 2006b, Fay and Yochem, 2007). *lin-15A* is a SynMuv class A gene and *lin-15B* is a SynMuv class B gene. Loss of function of both *lin-15A* and *lin-15B* genes result in the SynMuv phenotype at 20 °C (Cui et al., 2006b, Fay and Yochem, 2007). The *lin-15AB(n765)* strain was confirmed by genotyping (**Appendix 9.1 C**) and outcrossed five times into the wild-type *C. elegans* (Bristol N2) to remove any background mutations. After outcrossing, the *cfp-1* was knocked down by RNAi in the *lin-15AB(n765)* mutant and percentage of the SynMuv phenotype was scored.

As shown in **Table 5.1**, at 20 °C during RNAi negative control conditions, all *lin-15AB(n765)* animals had the SynMuv phenotype, which is exhibited as 2.35 ± 0.09 vulva-like protrusions (ectopic/synthetic vulva) on average (**Figure 5.1 B** and **Table 5.1**). Worms with normal vulva i.e. one vulva without extra protrusions were counted as zero protrusions. Wild-type (Bristol N2) animals had 0 vulva-like protrusions, i.e. no synthetic multivulva (**Figure 5.1 A** and **Table 5.1**). RNAi mediated knockdown of *cfp-1* partially suppressed the SynMuv phenotype (**Table 5.1**) (Cui et al., 2006b). Furthermore, the average number of protrusions in *lin-15AB(n765)* mutants treated with *cfp-1* RNAi was also reduced (**Table 5.1**). This suggests that depletion of *cfp-1* expression suppressed the SynMuv phenotype of *lin-15AB(n765)* mutants.

To confirm that the suppression of the SynMuv phenotype was indeed due to *cfp-1* depletion, I measured levels of the *cfp-1* mRNA in *lin-15AB(n765)* mutants treated with *cfp-1* RNAi. The levels of *cfp-1* mRNA were reduced by 89% (**Figure 5.2**), suggesting that the partial suppression of the SynMuv phenotype of *lin-15AB(n765)* mutants is due to depletion of *cfp-1*.

Table 5.1: Percentage of *lin-15AB(n765)* mutant animals with the SynMuv phenotype during *cfp-1* RNAi.

The percentage of the SynMuv phenotype was scored in *lin-15AB(n765)* mutants treated with RNAi negative control (control) or *cfp-1* RNAi. The average number of vulva-like protrusions were counted. Animals with wild-type vulva were counted as 0 protrusions. Experiments were conducted at 20 °C. The table represents the average of three biological replicates. n is total animals from three replicates. **P < 0.01. P-value was derived using a Fisher's exact test. SEM, standard error of the mean.

Genotype (RNAi)	**Percentage of SynMuv ± SEM (n)**	**Avg number of protrusions ± SEM**
***lin-15AB(n765)* (control)**	100 (142)	2.35 ± 0.09
***lin-15AB(n765)* (*cfp-1*)**	78 ± 1.15 (207) **	1.45 ± 0.03
Wild-type (control)	0 (66)	0

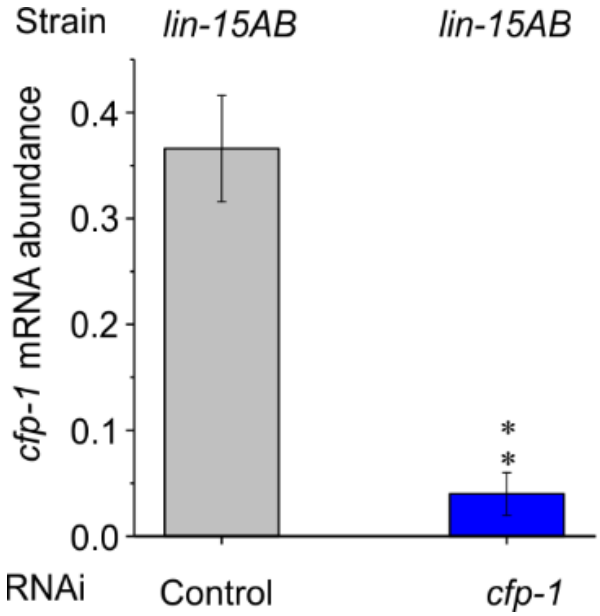


Figure 5.2 : Expression levels of the *cfp-1* mRNA upon RNAi negative control and *cfp-1* RNAi.

Levels of the *cfp-1* transcript in wild-type animals were significantly reduced during *cfp-1* RNAi (blue) compared to RNAi negative control (control) (grey). *pmp-3* and *tba-1* genes were used for normalisation. The figure represents the average of three biological replicates (n = 50 in each replicate). P-values were calculated using a student t-test: ** = P<0.01. Error bars represent ± standard error of the mean (SEM).

5.2.2 The *cfp-1(tm6369)* genetic mutant confirmed the role of CFP-1 in the SynMuv phenotype.

To further confirm the role of CFP-1 in the SynMuv phenotype, I crossed the *lin-15AB(n765)* strain with the *cfp-1(tm6369)* mutant to generate *cfp-1(tm6369)*;*lin-15AB(n765)* double mutants. Consistent with the RNAi result, I observed a significant reduction in the SynMuv phenotype in *cfp-1(tm6369)*;*lin-15AB(n765)* mutants. In addition, the average number of protrusions were also reduced (**Table 5.2**). This strongly suggests that the loss of *cfp-1* function suppressed the SynMuv phenotype of *lin-15AB(n765)* mutants.

In the previous studies, it has been observed that histone methyltransferases can suppress the SynMuv phenotype (Cui et al., 2006b, Fay and Yochem, 2007, Andersen, 2007). CFP-1 and SET-2 are the key subunits of the H3K4me3 methyltransferase and loss of CFP-1 or SET-2 function results in a drastic reduction of H3K4me3. The suppression of SynMuv in the *cfp-1(tm6369)*;*lin-15AB(n765)* mutant could be due to loss of H3K4me3. To investigate this hypothesis, double mutants of *set-2(bn129)* and *lin-15AB(n765)* were generated. In contrast to *cfp-1(tm6369);lin-15AB(n765)* mutants, all of *set-2(bn129);lin-15AB(n765)* mutants displayed the SynMuv phenotype (**Table 5.2**) suggesting that the loss of function of SET-2 does not suppress the SynMuv phenotype of *lin-15AB(n765)* mutants. Additionally, the average number of protrusions were similar to the average number of protrusions in the *lin-15AB(n765)* mutant (**Table 5.2**). This finding suggests that the suppression of SynMuv in the *cfp-1(tm6369)*;*lin-15AB(n765)* mutant was not due to loss of H3K4me3, indicating a non-overlapping role of CFP-1 and SET-2/H3K4me3 in vulva development.

Table 5.2: Percentage of animals with the SynMuv phenotype and the average number of protrusions

The percentage of SynMuv and the average (Avg) number of protrusions were scored in *lin-15AB(n765), cfp-1(tm6369);lin-15AB(n765)* and *set-2(bn129);lin-15AB(n765)* mutants. Experiments were conducted at 20 °C. The table combines the results of three biological replicates. n is total animals from three replicates. **P < 0.01. P-Value was derived using Fisher's exact test. SEM, standard error of the mean.

Genotype	**Percentage of SynMuv ± SEM (n)**	**Avg number of protrusions ± SEM**
lin-15AB(n765)	100 (264)	2.31 ± 0.09
cfp-1(tm6369);lin-15AB(n765)	76 ± 5 (341) **	1.39 ± 0.11
set-2(bn129);lin-15AB(n765)	100 (258)	2.34 ± 0.05
Wild-type	0 (80)	0

5.2.3 Role of CFP-1 in the RTK/Ras/MAPK pathway during vulval development

RTK/Ras/MAPK signalling plays an important role in vulval development as gain or loss of this signalling results in the Muv or vulvaless phenotype, respectively (Sundaram, 2006, Thomas et al., 2003). Previous studies suggest that SynMuv genes antagonise RTK/Ras/MAPK signalling to prevent ectopic vulva development (Cui et al., 2006a, Cui et al., 2006b). CFP-1 might act in the RTK/Ras/MAPK signalling in vulva development. Thus, loss of function of CFP-1 might have resulted in downregulation of RTK/Ras/MAPK signalling. To explore this possibility, *cfp-1* was knocked down in the Ras gain of function mutant strain MT2124 [*let-60(n1046)*]. LET-60 is the *C. elegans* orthologue of mammalian Ras; it stimulates a MAPK cascade consisting of the kinases MEK-2, MPK-1 and LIN-45 (Kornfeld et al., 1995, Lackner et al., 1994, Wu and Han, 1994, Wu et al., 1995). *let-60(n1046)* is a non-lethal gain of function *let-60* (Ras) allele and has 97% penetrance of the Muv phenotype (Singh and Han, 1995). If the loss of *cfp-1* function results in downregulation of RTK/Ras/MAPK signalling in vulval development, then RNAi-mediated knockdown of *cfp-1* in *let-60(n1046)* mutants would suppress the Muv phenotype of this mutant. Surprisingly, knockdown of *cfp-1* did not decrease the Muv phenotype of *let-60(n1046)* mutants (**Table 5.3**), which suggests that CFP-1 does not act on RTK/Ras/MAPK signalling. However, it is possible that CFP-1 might act upstream of LET-60 in RTK/Ras/MAPK signalling.

To explore this, the *let-23* (receptor of Ras signalling) gain of function mutant, PS1839 [*let-23(sa62)*] was used (Katz et al., 1996). The *let-23(sa62)* mutant has 97 ± 1.16% penetrance of the Muv phenotype (**Table 5.3**). Interestingly, RNAi knockdown of *cfp-1* had no significant impact on the Muv phenotype of *let-23(sa62)* mutants (**Table 5.3**). Since there is no suppression of the Muv phenotype, CFP-1 may not interact genetically with the RTK/Ras/MAPK pathway in vulval development.

Table 5.3: Percentage of *let-60(n1046), let-23(sa62)* and *lin-12(n137n460)* mutants exhibiting the Muv phenotype during RNAi-mediated knockdown of *cfp-1*

The percentage of Muv was scored in *let-60(n1046), let-23(sa62)* and *lin-12(n137n460)* mutants treated with *cfp-1* RNAi compared to RNAi negative control (control). Two biological replicates were combined for the table. n is total animals from two replicates. SEM, standard error of the mean.

Genotype (RNAi)	**Percentage of Muv ± SEM (n)**
***let-60(n1046)* (control)**	95 ± 0.6 (146)
let-60(n1046) (cfp-1)	95 ± 1.16 (190)
***let-23(sa62)* (control)**	97 ± 1.16 (185)
let-23(sa62) (cfp-1)	100 (217)
***lin-12(n137n460)* (control)**	96 ± 1.16 (198)
lin-12(n137n460) (cfp-1)	96.5 ± 0.3 (135)

5.2.4 Role of CFP-1 in the Notch pathway during vulval development

The Notch pathway also plays an essential role in vulval induction. CFP-1 might act in a Notch pathway to regulate vulval induction. To explore this possibility, *cfp-1* was knocked down in the MT1035 *[lin-12(n137n460)]* strain by RNAi. *lin-12(n137n460)* is a gain of function mutant, which shows a wild-type phenotype at 25 °C and displays Muv and egg-laying defective (Egl) phenotype at 15 °C (Lussi et al., 2016). RNAi-mediated knockdown of *cfp-1* in the *lin-12(n137n460)* mutant did not further reduce the percentage of animals exhibiting the Muv phenotype (**Table 5.3**). This suggests that CFP-1 may not act on the Notch pathway to regulate vulval development.

5.2.5 Suppressor/enhancer screen

Mutations causing the loss of *cfp-1* function (*cfp-1(tm6369)*) or depletion of *cfp-1* mRNA levels by RNAi suppresses the SynMuv phenotype of *lin-15AB(n765)* mutants. Interestingly, RNAi-mediated knockdown of *cfp-1* could not suppress the Muv phenotype of Ras and Notch gain of function mutants. It is possible that some other genes might regulate *cfp-1* in vulval induction. To identify gene(s) that could either enhance or suppress the SynMuv phenotype of *cfp-1(tm6369);lin-15AB(n765)* mutants, a small RNAi screen was carried out. For the RNAi screen, 15 genes that have been reported to suppress or promote the SynMuv/Muv phenotype or linked with H3K4me3 were selected (Margaritis et al., 2012, Zhou and Zhou, 2011, Lorenz et al., 2014, Lorenz et al., 2012, Cui et al., 2006b, Lehner et al., 2006). Loss or reduction of function of *gtbp-1*, *ham-3, mes-6*, *mys-4*, *phf-15, sin-3* or *sumv-1* genes is reported to suppress the SynMuv phenotype (Cui et al., 2006b, Lehner et al., 2006). RNAi of *athp-1, hda-1* or *rbp-11* enhances the Muv phenotype (Wannissorn, 2016, Poulin et al., 2005, Dufourcq et al., 2002, Slaughter et al., 2018). *egl-18* promotes vulval development (Sternberg, 2005). *ape-1*, *cec-1*, *pbrm-1* and *hat-1* orthologues are associated with H3K4me3 (Kim and Buratowski, 2009, Beilharz et al., 2017, Lorenz et al., 2014, Engelen et al., 2015, Slaughter et al., 2018). In addition to their role in the Muv/SynMuv phenotype, *athp-1, egl-18*, *hda-1*, and *sin-*

3 orthologues are also associated with H3K4me3 (Margaritis et al., 2012, Zhou and Zhou, 2011, Lorenz et al., 2014, Lorenz et al., 2012, Cui et al., 2006b, Lehner et al., 2006, Ang et al., 2016, He et al., 2013, Morishita et al., 2014, Wannissorn, 2016).

As shown in **Table 5.4**, more than 96% of *lin-15AB(n765)* mutants treated with *cec-1, egl-18, hat-1* or *rpb-11* RNAi exhibited the SynMuv phenotype. RNAi mediated knockdown of the same genes in *cfp-1(tm6369);lin-15AB(n765)* mutants had no significant effect (**Table 5.4**). RNAi-mediated knockdown of *mys-4* in *lin-15AB(n765)* mutants resulted in an 11% reduction. However, RNAi-mediated knockdown of *mys-4* in *cfp-1(tm6369);lin-15AB(n765)* mutants did not decrease the percentage of *cfp-1(tm6369);lin-15AB(n765)* mutants exhibiting the SynMuv phenotype (**Table 5.4**). This suggests that these genes do not suppress or enhance the function of *cfp-1* in the SynMuv phenotype.

On the other hand, individual RNAi knockdown of *ape-1, athp-1* or *ham-3* on *cfp-1(tm6369);lin-15AB(n765)* mutants further decreased the percentage of the SynMuv phenotype by 2-5% (**Table 5.4**). Similarly, individual RNAi of *mes-6*, *sumv-1* or *phf-15* further decreased the percentage of *cfp-1(tm6369);lin-15AB(n765)* mutants expressing the SynMuv phenotype by 10-18% (**Table 5.4**). These results suggest that individual RNAi of *athp-1*, *ham-3*, *ape-1, mes-6*, *sumv-1* or *phf-15* may partially suppress the SynMuv phenotype in *cfp-1(tm6369);lin-15AB(n765)* mutants. However, the suppression was not strong and statistically significant ($P>0.1$) to suggest a genetic interaction between *cfp-1* and *athp-1, ham-3, ape-1, mes-6, sumv-1* or *phf-15.*

Table 5.4: RNAi screen identified three genes that could enhance and one gene that could suppress the role of *cfp-1* in the vulval development

The percentage of SynMuv was scored in the *lin-15AB(n765)* and *cfp-1(tm6369);lin-15AB(n765)* background following RNAi of each candidate genes. 2-5 biological replicates were combined for the table. n is combined total animals from at least two replicates. **P < 0.01. P-value was derived using a Fisher's exact test. SEM, standard error of the mean. SEM, standard error of the mean.

	lin-15AB(n765)	***cfp-1(tm6369);lin-15AB(n765)***	**Reason for selection**	**References**
RNAi	**Percentage of SynMuv ± SEM (n)**	**Percentage of SynMuv ± SEM (n)**		
control	99.3 ± 1 (1191)	68 ± 7.8 (907)		
ape-1	95 ± 0.29 (153)	65 ± 0.6 (141)	Linked with H3K4me3	(Kim and Buratowski, 2009, Beilharz et al., 2017)
athp-1	98 ± 1.16 (184)	63.5 ± 3.8 (67)	Enhancer of Muv Linked with H3K4me3	(He et al., 2013, Morishita et al., 2014, Wannissorn, 2016)
cec-1	98 ± 1.16 (272)	72.5 ± 2.03 (64)	Linked with H3K4me3	(Lorenz et al., 2014)
egl-18	98.5 ± 1 (105)	67.5 ± 3.8 (81)	Role in vulval induction Linked with H3K4me3	(Sternberg, 2005) (Ang et al., 2016)
gtbp-1	95.75 ± 1.1 (180)	34 ± 9.6 (165) **	Muv suppression	(Cui et al., 2006b)

ham-3	93 ± 0.58 (249)	66.5 ± 1.45(44)	Muv suppression	(Lehner et al., 2006)
hat-1	100 (59)	80.5 ± 3.19 (118)	Linked with H3K4me3	(Engelen et al., 2015)
hda-1	100 (109)	45 ± 9.3 (136) **	Promotes SynMuv Linked with H3K4me3	(Dufourcq et al., 2002, Cui et al., 2006b, Lorenz et al., 2014).
mes-6	67.5 ± 0.29 (50)	52 ± 1.45 (138)	Muv suppression	(Cui et al., 2006b)
mys-4	89 ± 2.32 (69)	72 ± 0.3 (75)	Muv suppression	(Cui et al., 2006b)
pbrm-1	100 (277)	97 ± 1.3 (342) **	Linked with H3K4me3	(Slaughter et al., 2018)
phf-15	73 ± 7.55 (119)	50 ± 8.2 (262)	Muv suppression	(Cui et al., 2006b)
rpb-11	99 ± 0.58 (175)	66.5 ± 0.29 (101)	Promotes SynMuv	(Poulin et al., 2005)
sin-3	66 ± 13.37 (366)	20.6 ± 8.48 (332) **	Muv suppression Linked with H3K4me3	(Cui et al., 2006b, Lorenz et al., 2014)
sumv-1	65.5 ± 20 (464)	58.25 ± 17 (349)	Muv suppression	(Cui et al., 2006b)

5.2.6 PBRM-1 suppresses the function of CFP-1 in vulval induction

pbrm-1 is an orthologue of the mammalian *PBRM1* gene which encodes a BAF180 protein, a subunit of the switch/sucrose non-fermentable (SWI/SNF) complex (Shibata et al., 2012). I found that most of the *cfp-1(tm6369);lin-15AB(n765)* mutant animals treated with *pbrm-1* RNAi exhibited SynMuv compared to RNAi negative control (**Table 5.4**). Additionally, RNAi knockdown of *pbrm-1* in *cfp-1(tm6369);lin-15AB(n765)* also rescued the average number of protrusions to *lin-15AB(n765)* level (**Table 5.5**). This finding suggests that PBRM-1 suppresses the function of CFP-1 in the SynMuv phenotype.

5.2.7 GTBP-1, SIN-3 and HDA-1 enhance the function of CFP-1 in the SynMuv phenotype of the *lin-15AB(n765)* mutant

In the RNAi screen, it was observed that individual RNAi of *hda-1, sin-3* or *gtbp-1* in the *cfp-1(tm6369);lin-15AB(n765)* mutant remarkably reduced the number of animals with the SynMuv phenotype (**Table 5.4**). This suggests that *cfp-1* may interact genetically with *hda-1, sin-3* or *gtbp-1* to regulate vulva development in the SynMuv mutant. Next, I used a Fisher's exact test under the null hypothesis (see methods) to investigate whether the observed interaction is statistically significant. The Fisher's exact test results suggested that the reduction of the percentage of *cfp-1(tm6369);lin-15AB(n765)* animals with the SynMuv phenotype in individual RNAi knockdown of *hda-1, sin-3* or *gtbp-1* was statistically significant ($P<0.01$).

To further explore the impact of *hda-1, sin-3* or *gtbp-1* RNAi on the appearance of the SynMuv phenotype in *cfp-1(tm6369);lin-15AB(n765)* mutants, the average number of protrusions were counted. The average number of protrusions of the *lin-15AB(n765)* mutant during *hda-1, sin-3* or *gtbp-1* RNAi was significantly lower than the average number of protrusions in the *cfp-1(tm6369);lin-15AB(n765)* mutant grown on RNAi negative control (**Table 5.5**). Thus, the stronger and statistically significant reduction of the SynMuv phenotype suggests that *cfp-1* interacts genetically with *hda-1, sin-3* or *gtbp-1* to regulate vulval fate in the SynMuv mutant.

Table 5.5: Number of protrusions in *lin-AB(n765)* and *cfp-1(tm6369);lin-15AB(n765)* in RNAi knockdown of *pbrm-1, gtbp-1, sin-3* or *hda-1*

The average number of protrusions were counted under a Nomarski microscope. Animals with wild-type vulva were counted as 0 protrusion. The average number of protrusions were reduced in *cfp-1(tm6369);lin-15AB(n765)* mutants treated with *gtbp-1, sin-3* or *hda-1* RNAi compared to RNAi negative control (control). On the other hand, the average number of protrusions were increased in the *cfp-1(tm6369);lin-15AB(n765)* mutant when treated with *pbrm-1* RNAi. Two biological replicates were combined for the table. n is combined total animals from two replicates. SEM, standard error of the mean.

Genotype (RNAi)	**Avg protrusions per animal ± SEM (n)**
***lin-15AB(n765)* (control RNAi)**	2.42 ± 0.01 (140)
***lin-15AB(n765)* (*pbrm-1*)**	2.23 ± 0.01 (104)
***lin-15AB(n765)* (*hda-1*)**	2.3 ± 0.01 (92)
***lin-15AB(n765)* (*gtbp-1*)**	2.22 ± 0.01 (183)
***lin-15AB(n765)* (*sin-3*)**	1.45 ± 0.12 (311)
***cfp-1(tm6368);lin-15AB(n765)* (control RNAi)**	1.37 ± 0.02 (161)
***cfp-1(tm6368);lin-15AB(n765)* (*pbrm-1*)**	2.12 ± 0.01 (172)
***cfp-1(tm6368);lin-15AB(n765)* (*hda-1*)**	1.025 ± 0.03 (129)
***cfp-1(tm6368);lin-15AB(n765)* (*gtbp-1*)**	1.11 ± 0.03 (284)
***cfp-1(tm6368);lin-15AB(n765)* (*sin-3*)**	0.35 ± 0.04 (167)

5.2.8 GTBP-1 and RTK/Ras/MAPK signalling

The mammalian homologue of GTBP-1 is known to modulate the transduction of Ras-mediated signalling (Irvine et al., 2004). In this study, it was observed that RNAi of *gtbp-1* suppressed the SynMuv phenotype of the *lin-15AB(n765)* mutant and the *cfp-1(tm6369);lin-15AB(n765)* mutant. It is possible that GTBP-1 could act in the RTK/Ras/MAPK signalling pathway to modulate the SynMuv phenotype. To further explore this, *gtbp-1* was knocked down by RNAi in the *let-23(sa62)* mutant. Interestingly, knockdown of *gtbp-1* did not decrease the Muv phenotype of the *let-23(sa62)* strain (**Table 5.6**).

As RNAi knockdown of *gtbp-1* in the *cfp-1(tm6369);lin-15AB(n765)* mutant significantly reduced the percentage of animals exhibiting the SynMuv phenotype, they could also interact to regulate RTK/Ras/MAPK signalling. To investigate this hypothesis, *gtbp-1* and *cfp-1* were knocked down simultaneously by making a bacterial culture containing an equal mixture of *gtbp-1* and *cfp-1* RNAi bacteria. Interestingly, simultaneous RNAi knockdown of *gtbp-1* and *cfp-1* in the *let-23(sa62)* mutant did not reduce the percentage of the Muv phenotype (**Table 5.6**). This finding suggests that GTBP-1 does not interact genetically with Ras to regulate vulval induction. However, the observed results could be due to RNAi sensitivity of the *let-23(sa62)* strain and decreased effectiveness due to the combined RNAi knockdown approach.

Table 5.6: Percentage of *let-23(sa62)* mutants with the Muv phenotype in RNAi of *cfp-1*, *gtbp-1* and *cfp-1+gtbp-1* RNAi

The percentage of Muv was scored in the *let-23(sa62)* mutant background following RNAi of each candidate genes. The experiments were conducted at 20 °C. Two biological replicates were combined for the table. n is total animals from two replicates. SEM, standard error of the mean.

Genotype (RNAi)	**Percentage of Muv ± SEM (n)**
let-23(sa62) (*control*)	99 ± 0.6 (137)
let-23(sa62) (cfp-1)	100 (116)
let-23(sa62) (gtbp-1)	99.5 ± 0.3 (124)
let-23(sa62) (gtbp-1+cfp-1)	100 (125)

5.2.9 CFP-1 and GTBP-1 genetically interact with each other in *C. elegans* development

In the SynMuv assay, it was found that CFP-1 interacts genetically with GTBP-1 to regulate vulval development. It is possible that CFP-1 and GTBP-1 could genetically interact to regulate other phenotypes. To further investigate the genetic interaction between CFP-1 and GTBP-1 in *C. elegans* development, both a developmental assay and a fertility assay of the *cfp-1* mutant upon *gtbp-1* RNAi were conducted.

It was observed that after 68 h, 56% of the *lin-15AB(n765)* mutant reached the adult stage and 41% were in the L4 stage (**Figure 5.3 A**). RNAi knockdown of *gtbp-1* in *lin-15AB(n765)* mutants resulted in slight developmental delay, as 45% of *lin-15AB(n765)* mutants were in the adult stage during *gtbp-1* RNAi (**Figure 5.3 A**). I observed that the *cfp-1(tm6369);lin-15AB(n765)* mutant development was delayed compared to the *lin-15AB(n765)* mutant. As after 68 h, only 35% *cfp-1(tm6369);lin-15AB(n765)* mutants reached adulthood and most of the *cfp-1(tm6369);lin-15AB(n765)* mutant were still in the L4 stage (**Figure 5.3 A**). Interestingly, RNAi knockdown of *gtbp-1* in the *cfp-1(tm6369);lin-15AB(n765)* mutant enhanced the developmental delay phenotype of *cfp-1(tm6369);lin-15AB(n765)*, as only 2% of the *cfp-1(tm6369);lin-15AB(n765)* mutant during *gtbp-1* RNAi were in adult stage and most of the animals were in the L4 stage (**Figure 5.3 A**). This result suggests that *gtbp-1* interacts genetically with *cfp-1,* to promote the development of *C. elegans*.

In the fertility assay, *gtbp-1* was knocked down by RNAi in the *cfp-1(tm6369)* mutant, and the average brood size was measured. RNAi knockdown of *gtbp-1* in wild-type animals did not have a significant impact on fertility. Interestingly, the average brood size of the *cfp-1(tm6369)* mutant was significantly reduced during *gtbp-1* RNAi (**Figure 5.3 B**). This result indicates that *cfp-1* interacts genetically with *gtbp-1* to promote fertility.

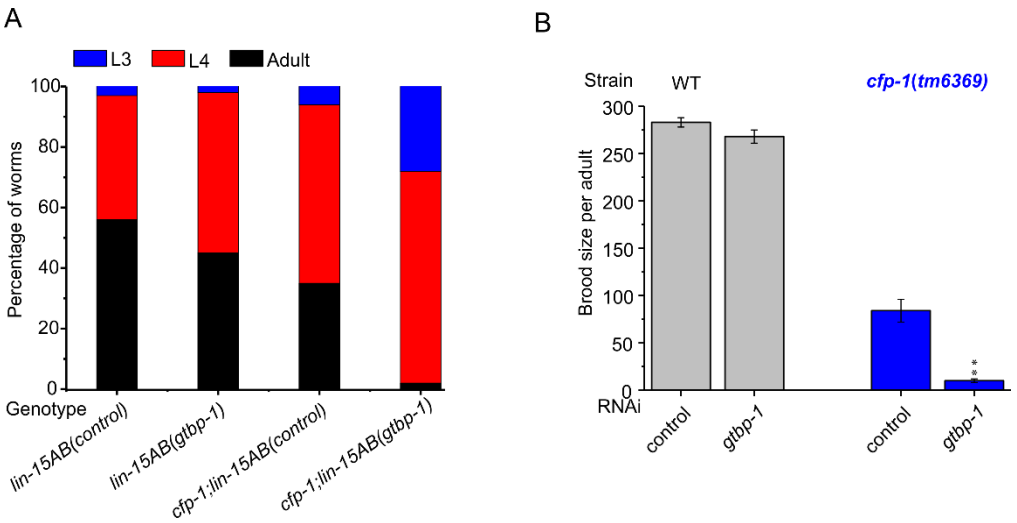


Figure 5.3: CFP-1 and GTBP-1 genetically interact to promote fertility and developmental rate

(A) Developmental progress of *cfp-1(tm6369);lin-15AB(n765)* and *lin-15AB(n765)* mutants grown on RNAi negative control or *gtbp-1* RNAi bacteria, monitored at 68 h at 20 °C. During *gtbp-1* RNAi, development of the *cfp-1(tm6369);lin-15AB(n765)* mutant was delayed compared to the *lin-15AB(n765)* mutant. Two biological replicates were combined for the figure (n > 25 per strain in each replicate). Combined number of animals from two replicates: *lin-15AB* (control) 233, *lin-15AB(gtbp-1)* 90, *cfp-1(tm6369);lin-15AB*(control) 156, *cfp-1(tm6369);lin-15AB(gtbp-1)* 57. (B) Total brood size assay for wild-type and the *cfp-1(tm6369)* mutant during *gtbp-1* RNAi at 20 °C. RNAi knockdown of *gtbp-1* significantly reduced the brood size of the *cfp-1(tm6369)* mutant. Two biological replicates were combined for the boxplot (n=10 in each replicate). P-values were calculated using student t-test: ** = P<0.01. Error bars represent ± standard error of the mean (SEM).

5.3 Discussion

Cfp1 is an epigenetic regulator that is responsible for H3K4me3 modification (Brown et al., 2017). It plays an essential role in cell fate specification and development of an organism (Skalnik, 2010). However, how Cfp1 contributes to this event is not yet clear. In this study, it was found that the loss of function of CFP-1 suppresses the SynMuv phenotype of SynMuv mutants. Interestingly, despite both *cfp-1* and *set-2* genes having a similar role in H3K4me3 deposition and development of *C. elegans*, the loss of function of SET-2 did not suppress the SynMuv phenotype. This suggests a function for CFP-1 outside of the COMPASS complex. Additionally, I observed that *pbrm-1* suppresses the function of *cfp-1*, whereas *gtbp-1, sin-3* or *hda-1* enhances the function of *cfp-1* in vulval induction. Taken together, this study suggests that *cfp-1* interacts genetically with *gtbp-1, sin-3, hda-1* and *pbrm-1* to regulate vulval cell fate in SynMuv mutants.

5.3.1 Role of CFP-1 in vulval cell fates

I observed that the loss of function of CFP-1 suppressed the SynMuv phenotype of SynMuv mutants. Interestingly, loss of function of CFP-1 in wild-type animals does not result in a defect in vulval development. It is possible that CFP-1 is only required for the ectopic vulval cell fates, but not in normal vulval development. SynMuv genes could inhibit the expression of the genes that promote the vulval cell fate in VPC cells, and after the loss of function of SynMuv genes, CFP-1 might promote the expression of SynMuv target genes, thus promoting the vulval cell fate-decision. It is also possible that the CFP-1 might act on distinct genes, different to those misregulated in the loss of SynMuv genes, and transcription of CFP-1 target genes might antagonise the activities of the SynMuv protein.

SynMuv genes are reported to act on RTK/Ras/MAPK signalling to suppress the Muv phenotypes (Cui et al., 2006a, Cui et al., 2006b, Simonet et al., 2007). It is possible that CFP-1 plays a role in the vulval induction by promoting RTK/Ras/MAPK signalling. However, RNAi knockdown of *cfp-1* did not reduce the

Muv phenotype of the Ras gain of function mutants. This suggests that CFP-1 may not act in the RTK/Ras/MAPK signalling pathways. Similar to RTK/Ras/MAPK signalling, RNAi knockdown of *cfp-1* did not reduce the Muv phenotype of the Notch gain of function mutants, suggesting that CFP-1 does not act in the Notch signalling pathway to promote the SynMuv phenotype.

5.3.2 Chromatin regulators regulate the role of CFP-1 in vulval development

SynMuv genes are essential to prevent ectopic vulval induction, and mutations in these genes results in the appearance of the SynMuv phenotype. A number of these genes are chromatin regulators, which are associated with gene expression and repression (Cui et al., 2006b, Fay and Yochem, 2007). For example, some of class B SynMuv genes encode for chromatin regulators which are associated with gene repression, whereas class C SynMuv genes encode for components of the NuA4 complex which are associated with gene expression (Harrison et al., 2006, Solari and Ahringer, 2000). The interplay between chromatin regulators, which are associated with gene expression or repression, regulates the proper cell fates in vulva development.

In this study, CFP-1 was found to antagonise SynMuv genes. Interestingly, it was found that PBRM-1, a subunit of chromatin remodelling complex (SWI/SNF complex), suppresses the function of CFP-1 in vulval induction. While HDA-1 enhanced the function of CFP-1 in the SynMuv phenotype. HDA-1 is the orthologue of mammalian histone deacetylase 1 (HDAC1) which is responsible for removal of acetylation modification. Earlier studies suggested that *hda-1* acts as a classical SynMuv gene (Lu and Horvitz, 1998, Solari and Ahringer, 2000). However, a study conducted by Dufourcq et al. (2002) using the *hda-1* genetic mutant suggested that loss of *hda-1* can induce the Muv phenotype without interacting with class A and B SynMuv genes. Surprisingly, in this study, it was observed that RNAi knockdown of *hda-1* in the *cfp-1(tm6369);lin-15AB(n765)* mutant strongly suppressed the SynMuv phenotype. One possible explanation for this intriguing finding is that HDA-1 may interact with different proteins such as SIN-3, SPR-1 and CHD-3, and function as a

transcriptional activator as well as a repressor to regulate vulval development (Silverstein and Ekwall, 2005, Grzenda et al., 2009).

In support of this, I observed that RNAi knockdown of *sin-3* reduced the SynMuv phenotype of the *cfp-1(tm6369);lin-15AB(n765)* mutant. SIN-3 is the orthologue of mammalian SIN3, and functions in a repressor or corepressor complexes along with HDAC1 (Grzenda et al., 2009). SIN-3 and HDA-1 are the critical subunits of the SIN-3/HDAC1/2 complex, which regulates gene expression. The SIN-3/HDAC1/2 complex plays a role in both gene expression and repression (Silverstein and Ekwall, 2005, Grzenda et al., 2009). It is possible that CFP-1 might interact with the SIN-3/HDAC1/2 complex to regulate the expression of gene/s that promote vulval induction such as *lin-3*. Previous studies have suggested that the SIN-3/HDAC1/2 complex can be recruited by DNA binding proteins (Doetzlhofer et al., 1999). CFP-1 contains a DNA-binding domain and binds to unmethylated CpG islands (CGIs) which are frequently found at mammalian and *C. elegans* promoters (Skalnik, 2010). CFP-1 could help in the recruitment of the SIN-3/HDAC1/2 complex at the *lin-3* promoter. The recruited SIN-3/HDAC1/2 complex could promote *lin-3* expression and also prevent the recruitment of chromatin remodelling complexes such as the SWI/SNF complex.

5.3.3 GTBP-1 interacts genetically with CFP-1 in *C. elegans* development

In this study it was observed that RNAi knockdown of *gtbp-1* suppresses the SynMuv phenotype in the *cfp-1(tm6369);lin-15AB(n765)* mutant. GTBP-1 is a homologue of the mammalian GAP (GTPase-activating proteins) binding protein (G3BP1) that binds to Ras GTPase activating protein GAP (Parker et al., 1996). GAP regulates the RTK/Ras/MAPK signalling by activating the GTPase activity of Ras protein. The previous study has suggested that G3BP1 acts downstream of RTK/Ras/MAPK signalling to promote cell proliferation (Irvine et al., 2004). Furthermore, in gastric cancer cells, it has been observed that overexpression of *G3BP1* results in an increase in the expression of ERK and ATK genes suggesting a positive correlation between *G3BP1* and RTK/Ras/MAPK signalling (Min et al.,

2015, Barnes et al., 2002, Armstrong et al., 2006, Zhang et al., 2018). In the present work, knockdown of *gtbp-1* suppressed the SynMuv phenotype of the *lin-15AB(n765)* mutant, and the suppression was enhanced by the knockdown of *cfp-1*. It is possible that both CFP-1 and GTBP-1 cooperate to regulate RTK/Ras/MAPK signalling. However, RNAi knockdown of *gtbp-1* and/or *cfp-1* did not reduce the Muv phenotype of the Ras gain of function mutants. It is possible that GTBP-1 and CFP-1 may regulate other genes to promote SynMuv development. Further experiments to identify the genes that are targeted by SynMuv proteins, CFP-1 and by GTBP-1 should help to dissect the exact mechanism behind SynMuv suppression as observed through the loss of CFP-1 and/or GTBP-1 function in SynMuv mutants.

Additionally, I found that *gtbp-1* genetically interacts with *cfp-1* to promote fertility and development. The RNAi knockdown of *gtbp-1* in the *cfp-1(tm6369)* mutant resulted in a drastic reduction in brood size and developmentally delayed phenotype. It is possible that CFP-1 and GTBP-1 might interact to regulate the expression of genes involved in reproduction and development of the organism. The identification of the genes that are misregulated in the loss of CFP-1 and/or GTBP-1 function could help to understand the observed interaction in detail.

5.3.4 SET-2/COMPASS independent function of CFP-1

Another key finding of this study was the non-overlapping function of CFP-1 and SET-2. CFP-1 and SET-2 are the key subunits of the Set1/COMPASS complex. Chapter 3 shows that both CFP-1 and SET-2 are required for bulk H3K4me3, and they play a similar role in fertility and development of *C. elegans.* Surprisingly, in this Chapter, it was observed that loss of function of *cfp-1*, but not *set-2*, suppressed the SynMuv phenotype of *lin-15AB(n765)*. Findings from this Chapter suggest that CFP-1 can function outside of the SET-2/COMPASS complex.

Chapter 6

CFP-1 interacts with HDAC1/2 complexes in *C. elegans* development

6.1 Introduction

The interplay between the highly dynamic histone modifications can determine chromatin structure and gene function (Zhang et al., 2015). At enhancer and promoter regions, histones are subject to high turn-over of acetylation or methylation modification which results in either activation or repression of gene expression (Zhang et al., 2015, Clouaire et al., 2014). Acetylation of histones, by conserved histone acetyltransferases (HATs) such as Gcn5, p300/CBP, sRC/p160 and MYST, promotes an open chromatin state that allows active transcription to occur (Marmorstein and Zhou, 2014). Acetylation of histones at promoters is counteracted by a group of histone deacetylases (HDAC) (Crump et al., 2011, Hsu et al., 2012). HDAC can form multiprotein complexes such as SIN3, NuRD, and CoREST to regulate gene expression (Kelly and Cowley, 2013, Gregoretti et al., 2004, Crump et al., 2011, Hsu et al., 2012). These HDAC complexes play key roles in the development of an organism. Histone 3 lysine 4 trimethylation (H3K4me3) at active promoters can act as binding sites for conserved HAT and HDAC multi-protein complexes to modulate the level of acetylated chromatin in active promoter regions (Shilatifard, 2012, Howe et al., 2017, Bochynska et al., 2018). Alternatively, these HATs and HDACs can also be recruited to the active promoter regions by DNA-binding proteins (Robert et al., 2004). For instance, Sp1 can physically interact with CBP1 (HAT) to activate gene expression (Song et al., 2002). Sp1 can also recruit HDAC1 to downregulate target gene expression (Doetzlhofer et al., 1999).

In Chapter 3, it was reported that both CFP-1 and SET-2 are required for the deposition of H3K4me3, fertility and development of *C. elegans*. Furthermore, it was also observed that the loss of function of CFP-1 or SET-2 resulted in a stronger gene induction (Chapter 4). Surprisingly, loss of function of CFP-1, but not SET-2, suppressed the synthetic multivulva (SynMuv) phenotype, an organogenesis defect

caused by aberrant epidermal development (Chapter 5). The genetic screen to identify the genes that can enhance or suppress CFP-1 function in the epidermal development indicated that CFP-1 can interact genetically with SIN-3 and HDA-1 to regulate the SynMuv phenotype. In this Chapter, the genetic interaction between CFP-1 and histone deacetylases was investigated. It was found that CFP-1, but not SET-2, interacts genetically with class I HDAC1/2 complexes.

6.2 Results

6.2.1 RNAi-mediated knockdown of *hda-1* enhances the poor fertility phenotype of *cfp-1(tm6369)*

In the SynMuv screen, it was observed that RNAi-mediated knockdown of *hda-1* suppresses the SynMuv phenotype of *cfp-1(tm6369);lin-15AB(n765)* mutant. This finding suggests that *cfp-1* could interact genetically with *hda-1*. To further investigate the interaction between *cfp-1* and HDACs, *hda-1* was knocked down by RNAi in *cfp-1(tm6369)* mutants and in wild-type animals then the impact of knockdown on fertility was measured. The depletion of *hda-1* levels by RNAi-mediated knockdown resulted in a significant reduction of the average brood size of wild-type *C. elegans* (**Figure 6.1**). Additionally, 50% of the embryos out of the total brood were dead. This finding is consistent with the previous report which suggests that the loss of *hda-1* is embryonic lethal (Shi and Mello, 1998).

Similar to wild-type animals, 50% of the embryos of *cfp-1(tm6369)* mutants upon *hda-1* RNAi were dead and 65% (n = 30) of *cfp-1(tm6369)* mutant animals were sterile. Additionally, RNAi-mediated knockdown of *hda-1* in *cfp-1(tm6369)* mutants dramatically reduced the brood size of *cfp-1(tm6369)* mutants by 87% (**Figure 6.1**). As shown in **Figure 6.1**, a stronger reduction in all the three biological replicates was observed. This result suggests a potential genetic interaction between *cfp-1* and *hda-1*. Additionally, I also used a one sided t-test for genetic interaction, to investigate whether the reduction of *cfp-1(tm6369)* brood size in *hda-1* RNAi was statistically significant to suggests a genetic interaction between CFP-1 and HDA-1 (Baugh et al., 2005, Mcmurchy et al., 2017). The one-sided t-test determined that the average brood size of *cfp-1(tm6369)* during *hda-1* RNAi was significantly lower ((P<0.01) when compared to the *cfp-1(tm6369)* mutant grown on RNAi negative control and the *hda-1* RNAi knockdown in wild-type animals (Baugh et al., 2005, Mcmurchy et al., 2017). This reduction of the brood size of the *cfp-1(tm6369)* mutant during *hda-1* RNAi suggests a genetic interaction between *cfp-1* and *hda-1*.

I also investigated whether *set-2* and *hda-1* genetically interact to promote fertility. Out of the total brood, 33% of the *set-2(bn129)* embryos upon *hda-1* RNAi were dead. As shown in **Figure 6.1**, the average brood size of the *set-2(bn129)* mutant on *hda-1* RNAi was slightly lower compared to RNAi negative control. Even though the average brood size of the *set-2(bn129)* mutant was reduced during *hda-1* RNAi, the reduction was not statistically significant ($P>0.1$), thus suggesting no genetic interaction between *set-2* and *hda-1*. Collectively, these results suggest that *cfp-1,* but not *set-2,* interacts genetically with *hda-1* to promote fertility.

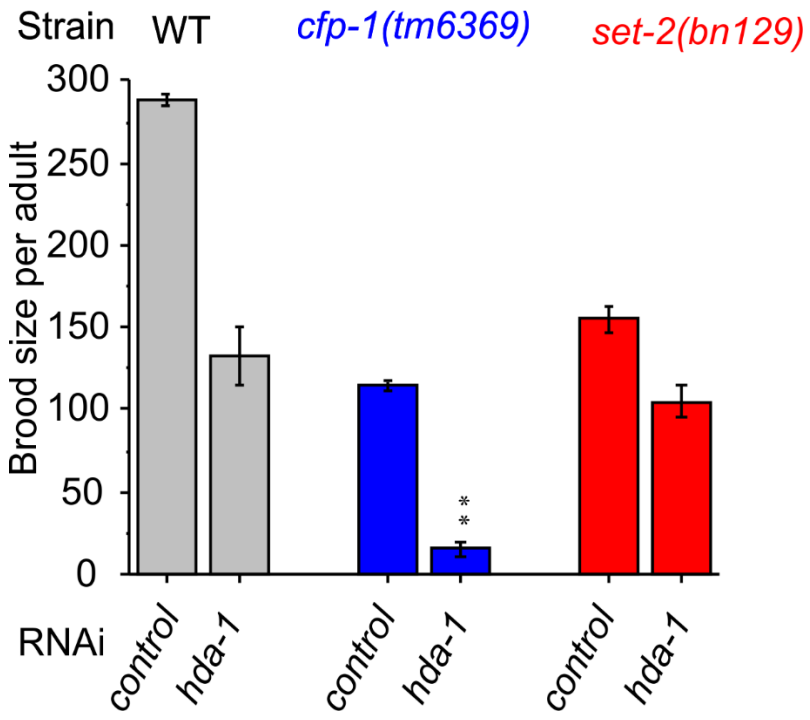


Figure 6.1: RNAi-mediated knockdown of *hda-1* reduces the brood size of *cfp-1(tm6369)*

Brood size assay showing genetic interactions between *cfp-1* and *hda*-1 using RNAi-mediated knockdown. *E. coli* containing empty vector was used as a RNAi negative control (control). The F1 generation on *hda-1* RNAi displayed a higher percentage of embryonic lethality, thus fertility was assayed at P0. RNAi knockdown of *hda-1* resulted in a reduction in brood size in wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants. A null hypothesis t-test (refer to the methods section) showed a genetic interaction between *hda-1* and *cfp-1* but not with *set-2.* Three biological replicates were combined for the figure (n=10 in each experiment). P-values were calculated using a one-tailed student t-test: ** = P<0.01. Error bars represent ± standard error of the mean (SEM).

6.2.2 CFP-1 cooperates with class I HDACs to promote fertility

hda-1 is an orthologue of mammalian HDAC1 (Shi and Mello, 1998). In mammals, HDAC1 often forms heterodimers with HDAC2 for deacetylase activity (Yang and Seto, 2003, De Ruijter et al., 2003, Yang and Seto, 2008). To investigate whether *cfp-1* also interacts genetically with *HDAC2*, *hda-2* or *hda-3,* potential homologues of *HDAC2* were knocked down in the *cfp-1(tm6369)* mutant and the effect of knockdown on fertility was measured (Shi and Mello, 1998).

Interestingly, unlike *hda-1* there was no significant change in the brood size of wild-type worms treated with *hda-3* RNAi*,* as the average brood size during *hda-3* RNAi was 269 ± 7 (average ± standard error of the mean), which was similar to the measured 280 ± 4 during RNAi negative control (**Figure 6.2**). On the other hand, the average brood size of the wild-type during *hda-2* RNAi was increased. However, this increment was not significant (**Figure 6.2**).

Similar to *hda-1*, the average brood size of the *cfp-1(tm6369)* mutants treated with either *hda-2* or *hda-3* RNAi was remarkably reduced in all the three biological replicates and the reduction was statistically significant ($P<0.05$) (**Figure 6.2**). This suggests that *cfp-1* interacts genetically with *hda-2* and *hda-3* to promote fertility. On the other hand, RNAi mediated depletion of *hda-3* levels in *set-2(bn129)* mutant animals did not result in a significant change in the brood size (**Figure 6.2**). Although *hda-2* RNAi increased the average brood size of the *set-2(bn129)* mutants, the increment was again not significant (**Figure 6.2**). Taken together, these findings suggest that *cfp-1* interacts genetically with *hda-2* and *hda-3* to promote fertility.

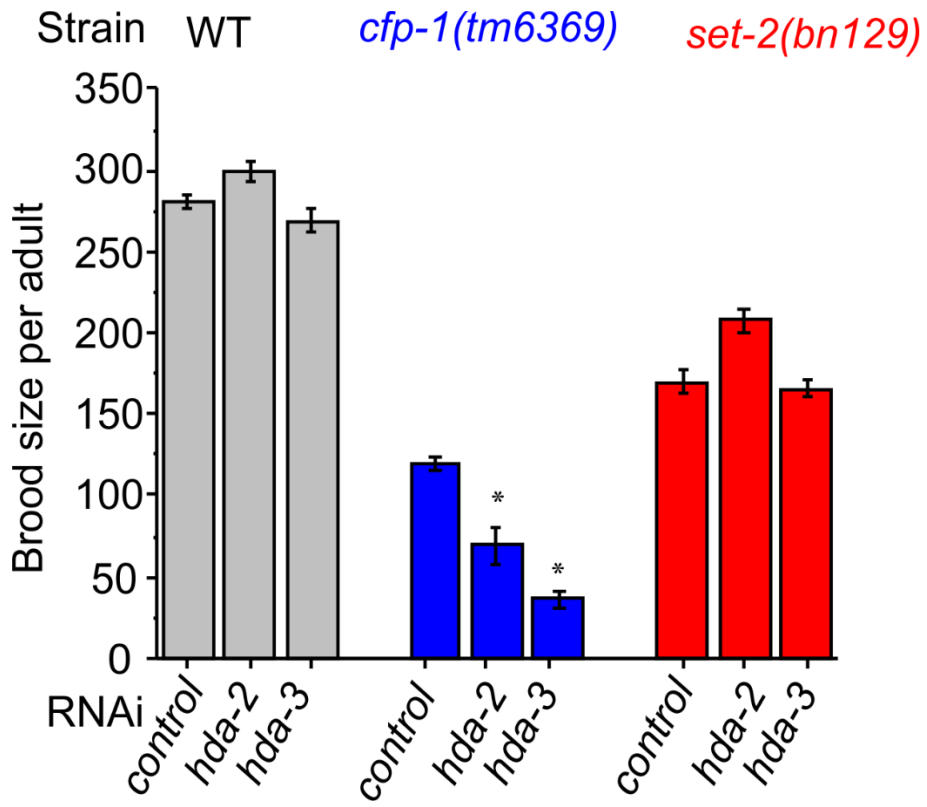


Figure 6.2: CFP-1 interacts genetically with HDAC 1/2 to promote fertility

Brood size assay showing genetic interactions between CFP-1 and HDAC 1/2 using *hda-2* and *hda-3* RNAi. The average brood size of *cfp-1(tm6369)* (blue) was further reduced by RNAi-mediated knockdown of *hda-2* or *hda-3* but had no significant impact on the *set-2(bn129)* mutant (red) brood size. Three biological replicates were combined for the figure (n=10-15 in each experiment). P-values were calculated using a one-tailed student t-test: * = $P<0.05$. Error bars represent ± SEM.

6.2.3 CFP-1 interacts genetically with SIN-3, CHD-3 and SPR-1 complexes

HDAC1 and HDAC2 are found in the SIN-3, NuRD and CoREST complexes, which are each defined by the presence of the major subunits SIN-3, CHD-3/LET-418 and SPR-1, respectively (Solari and Ahringer, 2000, Choy et al., 2007, Bender et al., 2007, Yang and Seto, 2003, De Ruijter et al., 2003, Yang and Seto, 2008). To test which of these HDAC1/2 complexes interact with CFP-1, the major subunits of each of the HDAC1/2 containing complexes were depleted by RNAi knockdown, and the effect on fertility was measured.

The RNAi-mediated knockdown of *sin-3* significantly reduced the brood size of the wild-type. Interestingly, 35% (n=30) of the *cfp-1*(*tm6369*) mutants were sterile upon *sin-3* RNAi. Furthermore, the fertility of the *cfp-1*(*tm6369*) was drastically reduced by 76% (**Figure 6.3**). Additionally, the observed reduction of the brood size was statistically significant ($P<0.01$), indicating a clear genetic interaction between *cfp-1* and *sin-3. sin-3* was also knocked down in the *set-2(bn129)* mutant, and the brood size was measured. The average brood size of the *set-2(bn129)* mutant was slightly reduced during *sin-3* RNAi compared to RNAi negative control (**Figure 6.3**). However, the reduction was not statistically significant ($P>0.1$), suggesting no genetic interaction between *set-2* and *sin-3*.

Next, either *chd-3* or *spr-1* was knocked down by RNAi in wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants. The average brood size of the wild-type upon *chd-3* or *spr-1* RNAi was 288 ± 7 or 278 ± 6, respectively, which was similar to the average brood size during RNAi negative control (293 ± 6) (**Figure 6.3**). Whereas the RNAi-mediated knockdown of *chd-3* or *spr-1* remarkably reduced the brood size of the *cfp-1(tm6369)* mutant to 45 ± 9 or 43 ± 9, respectively (**Figure 6.3**). On the other hand, RNAi-mediated knockdown of *chd-3* or *spr-1* did not reduce the fertility of the *set-2(bn129)* mutant (**Figure 6.3**). Collectively, these results suggest that *cfp-1*, but not *set-2*, interacts genetically with *sin-3, chd-3* and *spr-1*.

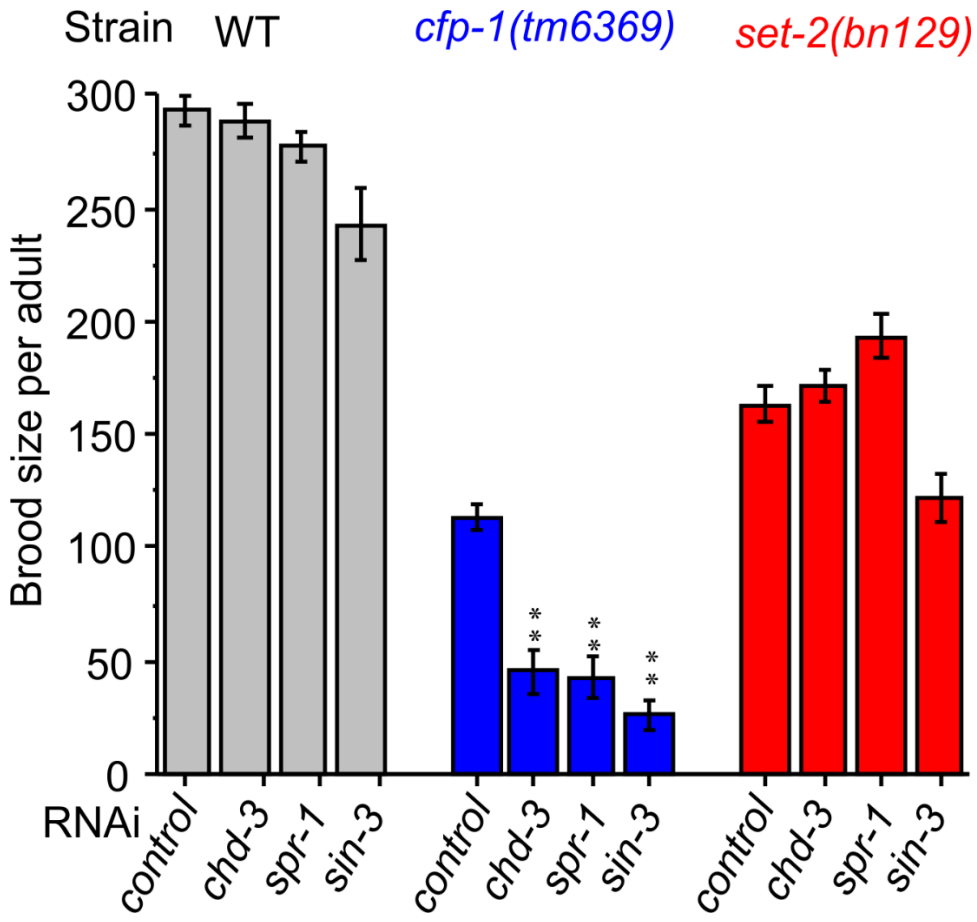


Figure 6.3: CFP-1 interacts with key subunits of HDAC 1/2 complexes to promote fertility

The average brood size on RNAi knockdown of *chd-3* or *spr-1* or *sin-3* or RNAi negative control (control) in wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants. Brood size of *cfp-1(tm6369)* was significantly reduced during RNAi-mediated knockdown of *chd-3*, *spr-1* or *sin-3*. RNAi-mediated knockdown of *chd-3* or *spr-1* had no significant impact on the *set-2(bn129)* mutant brood size. Knockdown of *sin-3* further reduced the brood size of the *set-2(bn129)* mutant, but the reduction was not statistically significant. Three biological replicates were combined for the figure (n=10-15 in each experiment). P-values were calculated using a one-tailed student t-test: ** = P<0.01, * = P<0.05. Error bars represent ± SEM.

6.2.4 Genetic mutants confirmed the genetic interaction between CFP-1 and HDAC1/2 complexes

To further confirm the RNAi results, *cfp-1(tm6369)* and *set-2(bn129)* mutants were crossed with *sin-3(tm1276),* and *spr-1(ok2144)* mutants (**Appendix 9.1 D and E)** and the brood size of these double mutants was measured (**Figure 6.4)**. The average brood size of the *spr-1(ok2144)* was significantly lower than the wild-type (**Figure 6.4 A**). As expected, the brood size of the *cfp-1(tm6369);spr-1(ok2144)* mutant was 12 ± 1, which was significantly lower than the average brood size of the *cfp-1(tm6369)* and *spr-1(ok2144)* single mutants (**Figure 6.4 A**). Furthermore, the reduction was greater compared to *spr-1* RNAi in *cfp-1(tm6369)* mutants (**Figure 6.3** and **Figure 6.4 A**). This result further confirms the genetic interaction between *cfp-1* and *spr-1.* On the other hand, the average brood size of the *set-2(bn129);spr-1(ok2144)* was similar to the average brood size of the *set-2(bn129)* mutant (**Figure 6.4 A**).

Next, the brood size of *sin-3(tm1276), cfp-1(tm6369);sin-3(tm1276)* and *set-2(bn129);sin-3(tm1276)* mutants was measured. It was observed that the brood size of the *sin-3(tm1276)* mutant was significantly lower than the average brood size of the wild-type (**Figure 6.4 B**). Surprisingly, *cfp-1(tm6369);sin-3(tm1276)* mutants were completely sterile (**Figure 6.4 B**). Due to the sterility, the *cfp-1(tm6369);sin-3(tm1276)* mutant could not grow for more than one generation. On the other hand, *set-2(bn129);sin-3(tm1276)* mutants were fertile. The average brood size of the *set-2(bn129);sin-3(tm1276)* mutant was similar to the brood size of the *sin-3(tm1276)* single mutant (**Figure 6.4 B**). Collectively, these results confirm the genetic interaction between *cfp-1* and *sin-3*, and between *cfp-1* and *spr-1.*

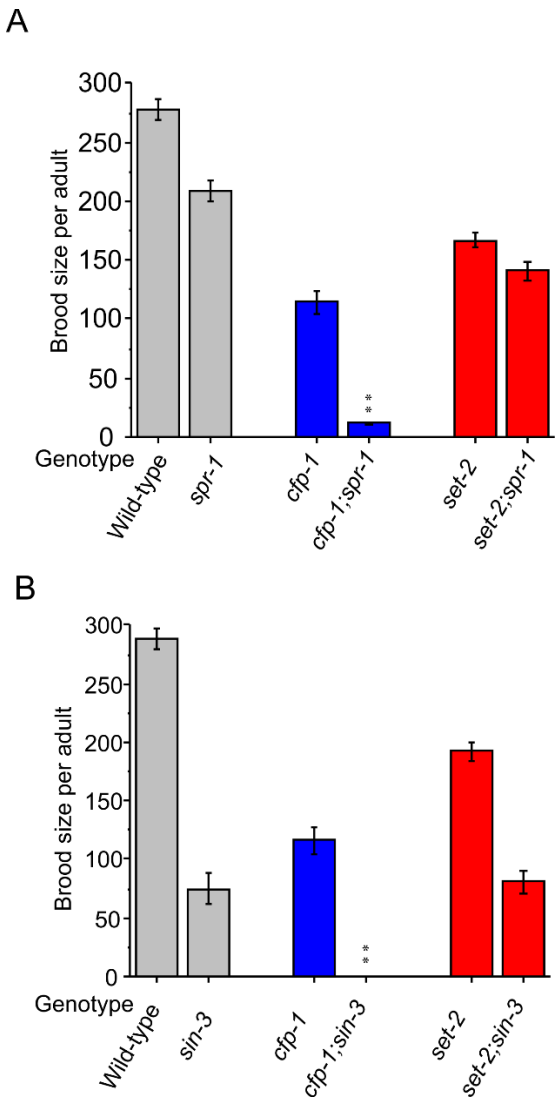


Figure 6.4: Genetic interaction between CFP-1 and HDAC1/2 subunits

Brood size assays showing genetic interactions between *cfp-1* and *sin-3* or *spr-1* using double mutants. (A) The average brood size of *spr-1(ok2144)* mutants. The average brood size of *cfp-1(tm6369);spr-1(ok2144)* was lower than 12, whereas the brood size of *set-2(bn129);spr-1(ok2144)* was similar to *set-2(bn129)* mutant. Two biological replicates were combined for the figure (n=10 in each experiment). (B) Brood size of *set-2(bn129);sin-3(tm1276)* mutants was similar to the brood size of *sin-3(tm1276)* showing no genetic interaction. All of the *cfp-1(tm6369);sin-3(tm1276)* mutants were sterile. Two biological replicates were combined for the figure (n=10 in each experiment). P-values were calculated using a one-tailed student t-test: ** = P<0.01. Error bars represent ± SEM.

6.2.5 TSA reduces the fertility of the *cfp-1(tm6369)* mutant

I observed a decrease in the brood size of the *cfp-1*(*tm6369*) mutant during *hda-1*, *hda-2* or *hda-3* RNAi. To test whether it was due to a reduction in deacetylase activity, *cfp-1*(*tm6369*) mutants were treated with Trichostatin A (TSA). TSA is a chemical drug that specifically inhibits class I/II histone deacetylases, and TSA treated cells have a significant gain in histone acetylation (Crump et al., 2011, Hsu et al., 2012, Vastenhouw et al., 2006). TSA is a toxic chemical. To avoid toxicity issues, the chemical was used at a final concentration of 4 μM, which is known to be non-toxic for wild-type worms (Vastenhouw et al., 2006). TSA was dissolved in dimethyl sulphoxide (DMSO), and DMSO alone was used as the control. Worms were transferred to the plates containing 4 μM TSA or DMSO for the fertility assay. It was found that the brood size of wild-type worms treated with TSA was 261 ± 10, which was similar to 240 ± 8 on no TSA control (**Figure 6.5**). This suggests that the brood size of the wild-type worms was not affected by TSA. Similarly, the average brood size of the *set-2(bn129)* mutant was not affected when treated with TSA (**Figure 6.5**). In contrast, the average brood size of the *cfp-1(tm6369)* mutant was significantly reduced from 119 ± 8 on no TSA control to 80 ± 10 on TSA treatment (**Figure 6.5**). These findings suggest that that decreased brood size of the *cfp-1*(*tm6369*) mutant during HDAC1/2 RNAi could be due to reduced deacetylase activity. Taken together these findings support the genetic interaction between CFP-1 and HDAC1/2 and provide the functional link for the interaction.

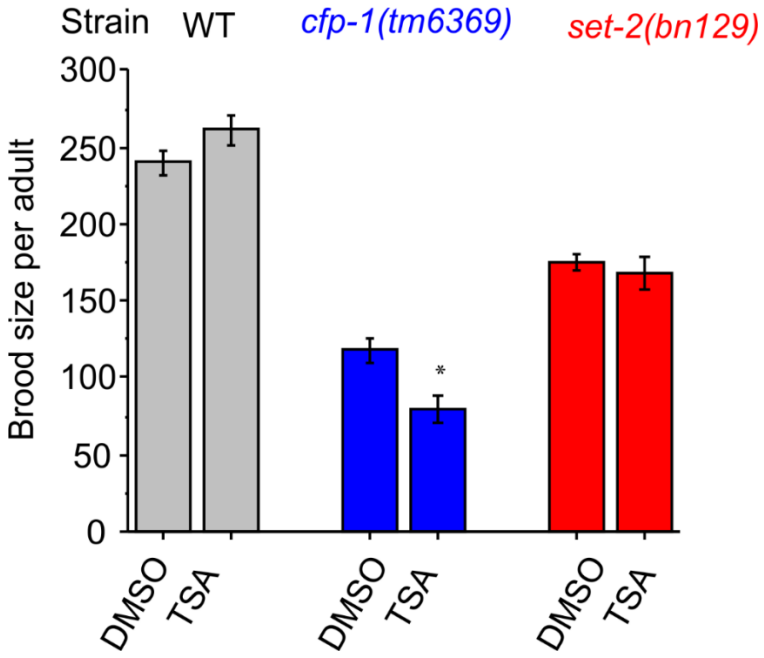


Figure 6.5: TSA treatment further reduces the brood size of *cfp-1(tm6369)*
Brood size of wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants treated with control (DMSO) or Trichostatin A (TSA). The average brood size of the *cfp-1(tm6369)* mutant was significantly reduced when treated with TSA whereas the brood size of the wild-type and the *set-2(bn129)* mutant was not affected by TSA treatment. Three biological replicates were combined for the figure (n=9-10 in each replicate). P-values were calculated using a one-tailed student t-test: * $P<0.05$. Error bars represent ± SEM

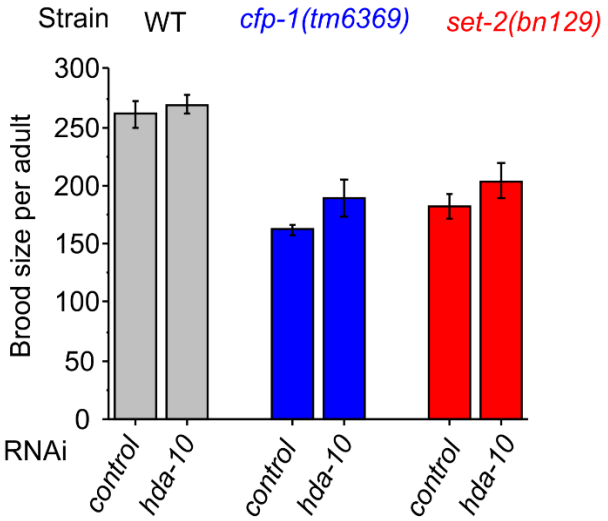
6.2.6 CFP-1 does not interact genetically with class II, III and IV HDACs

HDACs are classified into four groups (class I, class II, class III and class IV) based on the sequence homology and phylogenetic analysis of the yeast histone deacetylase (Shi and Mello, 1998, Gregoretti et al., 2004). *hda-1, hda-2* and *hda-3* are class I HDACs (Shi and Mello, 1998, Gregoretti et al., 2004). To investigate whether *cfp-1* also interacts genetically with other classes of HDACs, *hda-10* of class II, *sir-2.1* or *sir 2.2* of class III and *hda-11* of class IV were knocked down in wild-type, *cfp-1*(*tm6369*) and *set-2(bn129)* mutants, and fertility was measured.

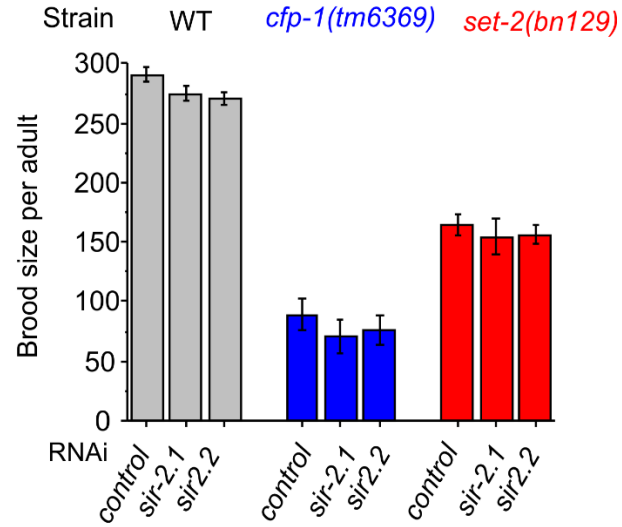
The average brood size of the wild-type treated with *hda-10* RNAi was similar to the brood size during RNAi negative control (**Figure 6.6 A**). This suggests that the knockdown of *hda-10* did not have a significant impact on the brood size of the wild-type. The average brood size of the *set-2(bn129)* and *cfp-1(tm6369)* mutant animals during *hda-10* RNAi knockdown was increased. However, the gain was not statistically significant (**Figure 6.6 A**). This suggests that *cfp-1* and *set-2* do not interact with *hda-10,* or possibly with class II HDACs.

Similarly, knockdown of class III HDACs, *sir-2.1* or *sir-2.2* did not have a significant impact on the brood size of wild-type, *cfp-1*(*tm6369*) and *set-2(bn129)* mutants (**Figure 6.6 B**). Next, *hda-11*, the only class IV HDAC, was knocked down and the brood size of wild-type, *cfp-1*(*tm6369*) and *set-2(bn129)* mutants was measured. Similar to *hda-10*, *sir-2.1* and *sir-2.2*, RNAi-mediated knockdown of *hda-11* did not have a significant impact on the fertility of the wild-type and the *set-2(bn129)* mutant (**Figure 6.6 C**). On the other hand, the average brood size of the *cfp-1*(*tm6369*) mutant on *hda-11* RNAi was reduced compared to RNAi negative control (**Figure 6.6 C**). However, the observed decrease was not significant to suggest a genetic interaction between *cfp-1* and *hda-11*. Collectively, these results suggest that *cfp-1* and *set-2* may not interact with class II, III and IV HDACs.

A


Strain
WT
cfp-1(tm6369)
set-2(bn129)
Brood size per adult
300
250
200
150
100
50
0
RNAi
control
hda-10
control
hda-10
control
hda-10

B


Strain
WT
cfp-1(tm6369)
set-2(bn129)
Brood size per adult
300
250
200
150
100
50
0
RNAi
control
sir-2.1
sir2.2
control
sir-2.1
sir2.2
control
sir-2.1
sir2.2

C

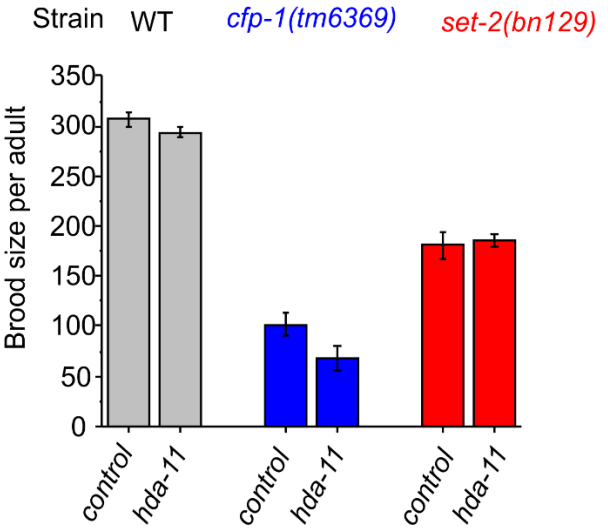

Strain
WT
cfp-1(tm6369)
set-2(bn129)
Brood size per adult
350
300
250
200
150
100
50
0
control
hda-11
control
hda-11
control
hda-11

Figure 6.6: CFP-1 does not interact with class II, III or IV HDACs

Brood size assays showing no genetic interactions between CFP-1 and class II, III or IV HDACs using RNAi knockdown (A-C). (A) The average brood size of wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants on a class II HDAC, *hda-10* RNAi. RNAi knockdown of *hda-10* had no impact on the brood size of the wild-type, whereas the brood size of *cfp-1(tm6369)* and *set-2(bn129)* mutants was slightly increased but not significantly. (B) The average brood size of wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants on class III HDACs, *sir-2.1* and *sir 2.2* RNAi. RNAi-mediated knockdown of *sir-2.1* or *sir 2.2* did not have an impact on the brood size of the wild-type and *set-2(bn129)* mutants. RNAi-mediated knockdown of *sir-2.1* or *sir 2.2* slightly reduced the brood size of *cfp-1(tm6369),* but the reduction was not significant. (C) The average brood size of wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutants on class IV HDACs, *hda-11* RNAi. Brood size of *cfp-1(tm6369)* was reduced during *hda-11* RNAi compared to RNAi negative control. However, the reduction was not significant. RNAi-mediated knockdown of *hda-11* did not have an impact on the average brood size of wild-type animals and the *set-2(bn129)* mutant. Two biological replicates were combined for figures (n=10-15 in each experiment). P-values were calculated using a one-tailed student t-test: * $P<0.05$. Error bars represent ± SEM

6.2.7 Investigating the genetic interaction between CFP-1 and HATs

Acetylation state of histones is controlled by HATs and HDACs. Here, it was observed that CFP-1 interacts genetically with HDAC1/2 complexes. To investigate, whether CFP-1 also interacts with HATs, I either knocked down *cbp-1* or *hat-1* by RNAi or crossed the *mys-4(tm3161)* mutant (**Appendix 9.1 F)** with *cfp-1(tm6369)* and *set-2(bn129)* mutants and measured the effect on fertility. *cbp-1* is a *C. elegans* homologue of mammalian histone acetyltransferase CBP and p300 (Shi and Mello, 1998). *mys-4* is the orthologue of mammalian histone acetyltransferase KAT6A (Erdelyi et al., 2017, Lau et al., 2016). *hat-1* is the orthologue of mammalian histone acetyltransferase HAT1 (Wenzel et al., 2011).

RNAi mediated knockdown of *hat-1* resulted in a significant reduction in the average brood size of wild-type animals compared to RNAi negative control (**Figure 6.7 A**). Interestingly, even though the average brood size of wild-type animals were reduced during the *hat-1* RNAi, the average brood size of *cfp-1(tm6369)* mutant animals was not reduced. Likewise, the brood size of *set-2(bn129)* mutants was also not affected upon *hat-1* RNAi (**Figure 6.7 A**).

On the other hand, the average brood size of *mys-4(tm3161)* mutants was 188 ± 10, which was significantly lower than the 265 ± 6 of wild-type animals (**Figure 6.7 B**). The brood size of *set-2(bn129);mys-4(tm3161)* double mutants was 123 ± 10 which was similar to the 115 ± 7 of *set-2(bn129)* mutants (**Figure 6.7 B**). Surprisingly, the brood size of *cfp-1(tm6369);mys-4(tm3161)* double mutants was significantly reduced compared to *cfp-1(tm6369)* and *mys-4(tm3161)* single mutants (**Figure 6.7 B**). In comparison to the wild-type, the average brood size of the *mys-4(tm3161)* mutant was also reduced significantly. Therefore, It is possible that the observed reduction in the brood size of *cfp-1(tm6369);mys-4(tm3161)* double mutants could be additive. Further genetic and biochemical experiments will help to investigate the interaction between *cfp-1* and *mys-4* in detail.

Next, I investigated the interaction between *cbp-1* and *cfp-1*. *cbp-1* RNAi resulted in a larval arrest in wild-type, *set-2(bn129)* and *cfp-1(tm6369)* mutants. Therefore, to

investigate the potential interaction between *cbp-1* and *cfp-1* or *set-2,* the *cbp-1*(*ku258*) gain of function mutant was used (Eastburn and Han, 2005). I observed that the brood size of *cbp-1*(*ku258*) mutants was 94 ± 12 which was significantly lower than the wild-type average brood size (269 ± 10) (**Figure 6.7 C**). The average brood size of the *cbp-1*(*ku258*) mutant did not change significantly during *cfp-1* or *set-2* RNAi (**Figure 6.7 C**). *cbp-1*(*ku258*) mutants also have an egg laying defective phenotype (Egl) (Eastburn and Han, 2005). I also investigated whether knockdown of *cfp-1* or *set-2* on *cbp-1*(*ku258*) enhance or suppress the Egl phenotype. As shown in **Table 6.1**, the percentage of the Egl phenotype of *cbp-1*(*ku258*) grown on *cfp-1* or *set-2* RNAi was similar to RNAi negative control. This suggests that neither *cfp-1* nor *set-2* enhance or suppress the Egl phenotype of *cbp-1*(*ku258*) mutants.

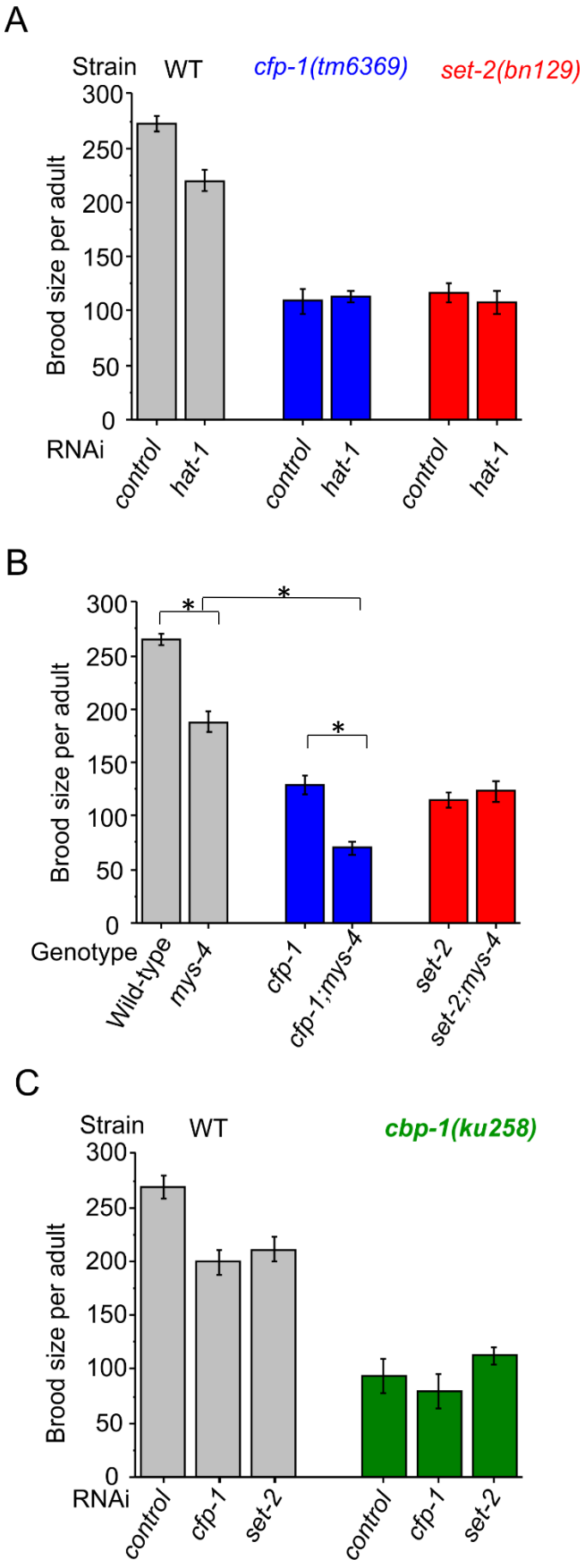

A
Strain
WT
cfp-1(tm6369)
set-2(bn129)
Brood size per adult
300
250
200
150
100
50
0
RNAi
control
hat-1
control
hat-1
control
hat-1
B
*
*
*
Brood size per adult
300
250
200
150
100
50
0
Genotype
Wild-type
mys-4
cfp-1
cfp-1;mys-4
set-2
set-2;mys-4
C
Strain
WT
cbp-1(ku258)
Brood size per adult
300
250
200
150
100
50
0
RNAi
control
cfp-1
set-2
control
cfp-1
set-2

Figure 6.7: Genetic interaction between CFP-1 and HATs

(A) Brood size of wild-type (grey), *cfp-1(tm6369)* (blue) and *set-2(bn129)* (red) mutant on *hat-1* RNAi. RNAi-mediated knockdown of *hat-1* significantly reduced the brood size of the wild-type. Interestingly, *hat-1* RNAi did not reduce the brood size of *cfp-1(tm6369)* and *set-2(bn129)* mutants. Two biological replicates were combined for the figure (n=15 in each experiment) (B) Brood size of wild-type (grey), *cfp-1(tm6369)* (blue), *set-2(bn129)* (red), *cfp-1(tm6369);mys-4(tm3161)* (blue) and *set-2(bn129);mys-4(tm3161)* (red) mutants at 20 °C. The *cfp-1(tm6369);mys-4(tm3161)* mutant exhibited a reduced brood size compared to single mutants, however the difference in brood size was not significant. Two biological replicates were combined for the figure (n=10 in each experiment). (C) Brood size of *cbp-1(ku258)* (green) on RNAi negative control, *cfp-1* and *set-2* RNAi. Brood size of *cbp-1(ku258)* on RNAi of *cfp-1* or *set-2* was similar to EV. Two biological replicates were combined for the figure (n=9-10 in each experiment). P-values were calculated using a student t-test: * = $P<0.05$. Error bars represent ± SEM

Table 6.1: Egg-laying defective (Egl) phenotype of the *cbp-1(ku258)* mutant
cfp-1 or *set-2* was knocked down by RNAi and the Egl phenotype was scored based on the presence of bags of eggs or hatched larva inside the mother. Two biological replicates were combined for the table. n is total animals from two biological replicates.

RNAi	Percentage of *cbp-1(ku258)* animals with the Egl phenotype	n
EV	25	99
cfp-1	26	109
set-2	28	235

6.2.8 The poor fertility and developmental delay of *cfp-1(tm6369)* or *set-2(bn129)* mutants is enhanced in *cfp-1(tm6369);set-2(bn129)* double mutants.

Genetic interaction results suggested that only *cfp-1* but not *set-2* interacts genetically with HDAC1/2 complexes to promote fertility. This is indicating that *cfp-1* and *set-2* may regulate fertility independently. To further investigate this, *cfp-1(tm6369);set-2(bn129)* mutants were generated, and the fertility was assayed. If both *cfp-1* and *set-2* act in the same pathway, then we expect that the average brood size of the *cfp-1(tm6369);set-2(bn129)* double mutant would be equal to the average brood size of either of the single mutants. Interestingly, the average brood size of *cfp-1(tm6369);set-2(bn129)* mutants was significantly lower than the average brood size of *cfp-1(tm6369)* and *set-2(bn129)* single mutants (**Figure 6.8 A**).

Additionally, the growth rate of *cfp-1(tm6369);set-2(bn129)* mutants as they develop from embryo to adult was measured and compared to single mutants. It was observed that after 68 hours, 100% of wild-type animals reached the adult stage, whereas, 72% of *cfp-1(tm6369)* and 69% of *set-2(bn129)* animals were adults after the same time period (**Figure 6.8 B**). Interestingly, only 10% of the *cfp-1(tm6369);set-2(bn129)* mutant reached the adult stage at the same time. Moreover, 59% and 31% of *cfp-1(tm6369);set-2(bn129)* mutants were in the L4 and L3 stage, respectively (**Figure 6.8 A**). This result suggests that the development of *cfp-1(tm6369);set-2(bn129)* mutants is delayed compared to single mutants.

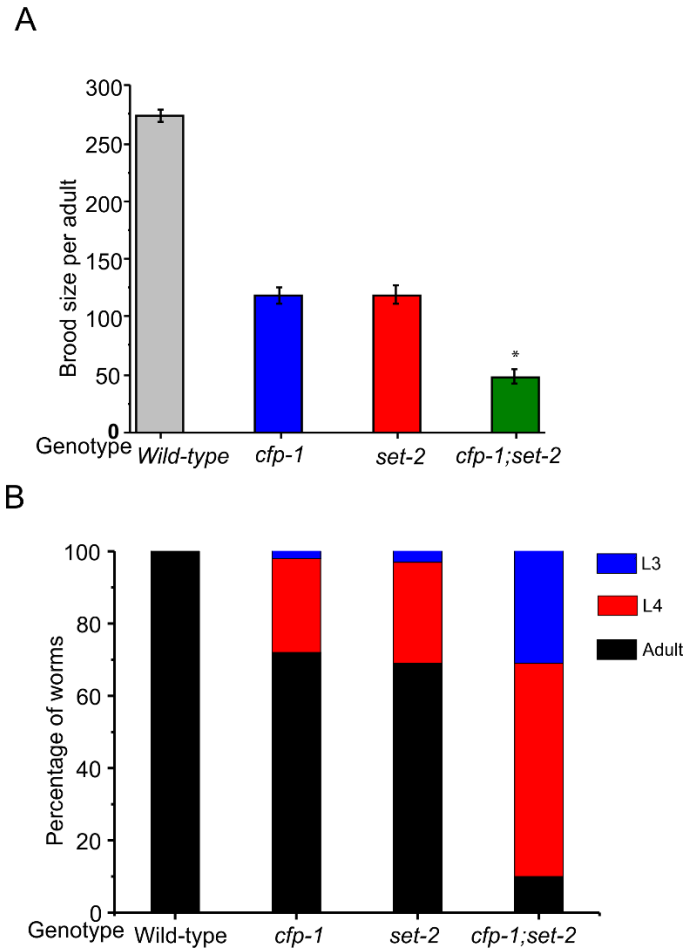


Figure 6.8: Loss of CFP-1 and SET-2 resulted in a further reduction of fertility and enhanced the developmental delay phenotype.

(A) The average brood size of wild-type, *cfp-1(tm6369)*, *set-2(bn129)* and *cfp-1(tm6369);set-2(bn129)* mutants at 20 °C. Double mutant have significantly reduced brood size compared to single mutants. Three biological replicates were combined for the figure (n=10 in each experiment). P-values were calculated using a one-tailed student t-test: * = P<0.05 (B) Developmental progress of wild-type, *cfp-1(tm6369)*, *set-2(bn129)* and *cfp-1(tm6369);set-2(bn129)* embryo monitored at 68 h at 20 °C. The development of the *cfp-1(tm6369);set-2(bn129)* mutant is delayed compared to *cfp-1(tm6369)* and *set-2(bn129)* single mutants. The figure is the average of two independent experiments (n>30 per strain in each experiment. Combined number of animals from two replicates: WT (129), *cfp-1(tm6369)* (73), *set-2(bn129)* (103) and *cfp-1(tm6369);set-2(bn129*) (171)). Error bars represent ± SEM.

6.3 Discussion

CFP-1 and SET-2 are major subunits of the COMPASS complex and are responsible for the bulk of H3K4me3 (Simonet et al., 2007, Li and Kelly, 2011, Xiao et al., 2011). In the previous chapters, it was observed that CFP-1 and SET-2 are important for H3K4me3 modification and play a repressive role in gene induction, thus suggesting that CFP-1 and SET-2 function together. However, in Chapter 5, it was observed that only a loss of CFP-1, but not SET-2 function, suppresses the SynMuv phenotype, suggesting a potential COMPASS independent role of CFP-1. In this Chapter, it was reported that CFP-1, but not SET-2, interacts genetically with HDAC1/2 complexes to promote fertility. These findings suggest that CFP-1 could function outside of the Set1/COMPASS complex.

6.3.1 CFP-1 interacts with HDAC complexes

The interplay between different chromatin modifiers and modifications plays a crucial role in the regulation of gene expression (Zhang et al., 2015, Clouaire et al., 2014). At promoter regions, H3K4me3 marked histones are subject to high turn-over of acetylation modifications, a consequence of the continuous opposing actions of histone acetyltransferases (HATs) and histone de-acetyltransferases (HDACs) (Zhang et al., 2015, Clouaire et al., 2014, Crump et al., 2011, Hsu et al., 2012). Previous studies have suggested the importance of interactions between acetylation and the COMPASS complex in gene regulation (Zhang et al., 2015, Clouaire et al., 2014). In this study, it was found that CFP-1 interacts genetically with HDAC1/2 complexes. RNAi-mediated knockdown of *hda-1, hda-2* and *hda-3* strongly enhance the poor fertility phenotype of the *cfp-1(tm6369)* mutant. HDAC1/2 mostly exists in multiprotein complexes such as SIN-3, CHD-3 and SPR-1 complexes (Solari and Ahringer, 2000, Choy et al., 2007, Bender et al., 2007, Yang and Seto, 2003, De Ruijter et al., 2003, Yang and Seto, 2008). Similar to class I HDAC, RNAi knockdown of key subunits (*sin-3, chd-3 and spr-1)* of SIN-3, CHD-3 and SPR-1 complexes, respectively, further reduced the brood size of the *cfp-1(tm6369)* mutant. Additionally, it was found that the average brood size of *cfp-1(tm6369)* was significantly lowered when treated with TSA, which is a class I/II HDACs specific

inhibitor. This suggests that the decreased brood size of *cfp-1(tm6369)* mutants during HDAC1/2 RNAi could be due to reduced deacetylase activity.

I next investigated whether CFP-1 also interacts genetically with class II, III or IV HDACs. Surprisingly, it was found that RNAi knockdown of HDACs of class II (*hda-10*), class III (*sir-2.1* and *sir-2.2*) and class IV (*hda-11*) did not significantly reduce the brood size of *cfp-1(tm6369)* and *set-2(bn129)* mutants. This suggests that the interactions between CFP-1 and HDACs could be class I HDACs specific. Taken together these findings indicate that CFP-1 interacts genetically with class I HDAC1/2 complexes.

Previous findings have suggested that HDACs could either be recruited by binding histone modifications such as H3K4me3 or through DNA binding proteins (Shilatifard, 2012, Howe et al., 2017, Bochynska et al., 2018, Robert et al., 2004, Song et al., 2002, Doetzlhofer et al., 1999). It is possible that the interaction between CFP-1 and HDAC1/2 could help in the recruitment of HDAC1/2 complexes at the promoter region. CFP-1 contains a DNA-binding domain and binds to the CpG regions (Brown et al., 2017, Skalnik, 2010). CFP-1 could aid in the recruitment of HDAC1/2 complexes at CpG enriched promoters to modulate histone acetylation, thus regulating the expression of genes involved in reproduction and development.

6.3.2 The interaction between CFP-1 and HAT-1

Acetylation modification of lysine residues of histone tails is involved in various biological events. HATs catalyse the acetylation and HDACs remove this modification. In this study, it was observed that depletion of class I HDACs by RNAi resulted in a significant reduction of the average brood size of *cfp-1(tm6369)* mutants. Additionally, the average brood size of *cfp-1(tm6369)* mutants was significantly lowered during the treatment with class I/II HDAC inhibitor (TSA). This suggests a putative role of CFP-1 in histone deacetylation.

On the other hand, depletion of *hat-1* levels by RNAi in wild-type animals significantly reduced the brood size compared to RNAi negative control. Interestingly, the average brood size of *cfp-1(tm6369)* mutants during *hat-1* RNAi was similar to RNAi negative control. This suggests that the loss of *cfp-1* function masked or rescued the effect of *hat-1* RNAi in the fertility of wild-type animals. The opposing effects of *cfp-1* with *hat-1* and the enhanced reduction of the brood size of *cfp-1(tm6369)* mutants in class I HDACs RNAi and TSA treatment further supports the role of CFP-1 in histone deacetylation. Further investigation using other phenotypes and biochemical assays could help to understand interactions between CFP-1 and HAT-1, and CFP-1 and HDACs.

6.3.3 SET-2 and HATs

It was also observed that despite having a significant reduction in the brood size of wild-type animals during *hat-1* RNAi, the average brood size of *set-2(bn129)* mutant animals during *hat-1* RNAi was similar to RNAi negative control. Additionally, the average brood size of *mys-4(tm3161)* mutants was significantly lower than the wild-type, whereas the average brood size of *set-2(bn129);mys-4(tm3161)* double mutants was similar to *set-2(bn129)* mutants.

The results suggest that the loss of *set-2* function may mask the effect caused by the loss of HATs function. This indicates that SET-2 may have some sort of interaction with HATs. Further research will help to understand the interaction in detail.

6.3.4 The interaction between CFP-1 and HDACs is independent of SET-2

RNAi-mediated knockdown *hda-1*, *hda-2*, *hda-3*, *sin-3*, *spr-1* or *chd-3* in *set-2(bn129)* mutants did not cause a significant reduction in the brood size. Moreover, the brood size of *set-2(bn129)* mutants on *hda-1*, *hda-2*, *hda-3*, *sin-3*, *spr-1* or *chd-3* RNAi was similar to RNAi negative control. It is possible that they might act on the same pathway to regulate fertility. Alternatively, different responses of the *cfp-1(tm6369)* and *set-2(bn129)* mutants to the same RNAi conditions could be due to

differential sensitivity to RNAi. *set-2(bn129)* mutants could be RNAi resistant. Thus there was no significant drop in the brood size of *set-2(bn129)* mutants in RNAi of *hda-1*, *hda-2*, *hda-3*, *sin-3*, *spr-1* and *chd-3*. However, in an RNAi sensitivity assay, it was found that both *cfp-1(tm6369)* and *set-2(bn129)* mutants responded similarly to the tested RNAi (Pokhrel et al., 2019). Additionally, the putative null allele of *sin-3(tm1276)* or *spr-1(ok2144)* mutants were crossed with *cfp-1(tm6369)* and *set-2(bn129)* and the brood size was measured. Similar to RNAi results, there was no significant reduction in the brood size of *set-2(bn129);sin-3(tm1276)* or *set-2(bn129);spr-1(ok2144)* mutants. These results suggest that the observed genetic interaction of *cfp-1*, but not *set-2*, with HDAC1/2 complexes is not due to RNAi sensitivity of *cfp-1(tm6369)* and *set-2(bn129)* mutants.

The observation that the brood size of *set-2(bn129)* mutants during RNAi-mediated knockdown of HDAC1/2 complexes did not result in a significant decrease, indicated a potential independent function of CFP-1 and SET-2 for fertility. The fertility and developmental defects of the *cfp-1* and *set-2* single mutants and the enhanced effect in the double mutant indicate that the CFP-1 and SET-2 each may have a unique and partially redundant role in *C. elegans* development. It is possible that CFP-1, together with SET-2 regulates the expression of one set of genes involved in fertility and development, and CFP-1 together with HDAC1/2 complexes, regulates another set of genes involved in fertility and development. Based on the findings from this study, I propose that CFP-1 could interact with both Set1/COMPASS and/or HDAC1/2 complexes in a context-dependent manner (**Figure 6.9**).

In summary, in this study, I observed that RNAi-mediated knockdown of *hda-1, hda-2, hda-3* and key subunits of HDAC1/2 complexes significantly reduced the brood size of *cfp-1(tm6369)* mutants. This suggests that CFP-1 and HDAC1/2 complexes genetically interacts to promote fertility. However, further investigation is required to understand the mechanism of interaction.

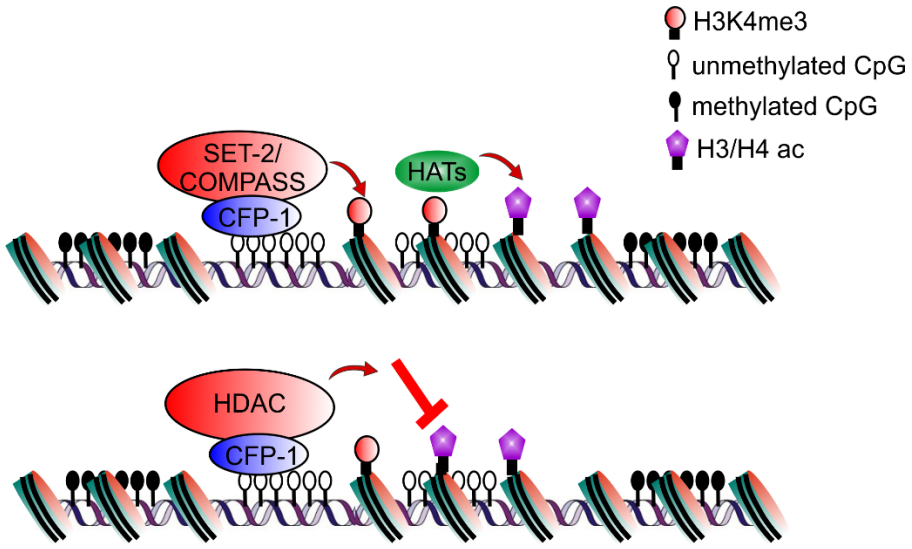


Figure 6.9: Proposed model indicating that CFP-1 cooperates with SET-2/COMPASS and/or with HDACs in a context-dependent manner

The canonical function of CFP-1 is to recruit the SET-2/COMPASS complex to promoter regions by binding onto unmethylated CpG islands. In the non-canonical function of CFP-1: CFP-1 could also recruit HDAC complexes to promoter regions in order to deacetylate the histones. Based on the physiological condition, CFP-1 could either interact with the COMPASS complex or with HDAC complexes.

Chapter 7

Discussion and future work

7.1 Discussion

In this study, I investigated the contribution of CFP-1 towards H3K4me3, the consequences caused by removing CFP-1 and identified the potential genetic regulators of the CFP-1. In order to perform this, *C. elegans* was used as a model organism. Findings from Chapter 3 suggest that CFP-1 is required for the maintenance of global H3K4me3 levels and it plays an essential role in fertility and development of *C. elegans*. Findings from Chapter 4 revealed the role of CFP-1 and SET-2 for stress-induced gene expression. Results in Chapter 5 show that CFP-1 is required for vulval development in SynMuv mutants. Unexpectedly, the last two chapters of the thesis (Chapter 5 and Chapter 6) uncovered that CFP-1 interacts genetically with class I HDAC1/2 complexes during *C. elegans* development, and this interaction is independent of the SET-2/COMPASS complex. Herein, I will discuss my observations as well as potential further studies.

7.1.1 Role of H3K4me3 in *C. elegans* development.

Previous studies in different organisms suggest that the role of CFP1 and SET1 to deposit H3K4me3 at active genes is conserved (Lee and Skalnik, 2005, Ardehali et al., 2011, Shilatifard, 2012, Yu et al., 2017, Clouaire et al., 2014, Clouaire et al., 2012). Consistent with previous findings, in this study, I observed that the levels of H3K4me3 were significantly reduced in *cfp-1(tm6369)* and *set-2(bn129) C. elegans* mutants. This reduction was accompanied by a significant reduction in the fertility of both mutants and delay in development. This suggests that H3K4me3 plays a crucial role in fertility and development. Interestingly, in Chapter 6 it was observed that the brood size of double *cfp-1(tm6369);set-2(bn129)* mutants displayed enhanced fertility defects and a developmentally delayed phenotype. Despite having a very low level of H3K4me3 in single mutants, the enhanced phenotypic defects in

the double mutants indicates that H3K4me3 may not be entirely responsible for the reduced fertility in both mutants. However, it is possible that the residual H3K4me3 present in both the *set-2(bn129)* and the *cfp-1(tm6369)* mutants can contribute to the fertility, and in *cfp-1(tm6369);set-2(bn129)* mutants, the H3K4me3 may be further reduced or completely removed thus resulting in enhanced fertility defect and developmentally delayed phenotype. Alternatively, the enhanced phenotypic defects could be due to the independent function of CFP-1 and SET-2 in *C. elegans* development. It is possible that CFP-1 and SET-2 may have a unique and partially redundant role in *C. elegans* development. In line with this, Chapter 6 shows that CFP-1, but not SET-2, interacts with HDACs to promote fertility.

7.1.2 Role of H3K4me3 in gene expression

Previous studies have shown that the levels of H3K4me3 are correlated with gene expression (Pena et al., 2008, Ardehali et al., 2011, Wysocka et al., 2006). Despite H3K4me3 being associated with active gene expression, loss of H3K4me3 causes very little change in global gene expression levels (Howe et al., 2017, Clouaire et al., 2014). Moreover, recent studies indicate a repressive role of H3K4me3 in gene expression (Margaritis et al., 2012, Zhou and Zhou, 2011, Lorenz et al., 2014, Mikheyeva et al., 2014). In concordance with this, I observed hyper-induction of stress-inducible genes in *cfp-1(tm6369)* and *set-2(bn129)* mutants which have extremely low levels of H3K4me3.

Additionally, in both *cfp-1(tm6369)* and *set-2(bn129)* mutants, the average brood size was drastically reduced, and both displayed a developmentally delayed phenotype. This implies that H3K4me3 modification may be essential for the transcription of genes that are required for fertility and development. Many downstream effectors might play a role in the H3K4me3-dependent transcriptional regulation. For example, TAF3, a key subunit of TFIID, binds to H3K4me3. TAF3 plays a crucial role in transcriptional initiation. The decrease in the H3K4me3 impairs the recruitment of TAF3, thereby causing a defect in the initiation of transcription (Lauberth et al., 2013). Thus, it is possible that H3K4me3 interacts with the transcriptional machinery to regulate gene expression.

7.1.3 Functions of CFP-1 independent of the SET-2/COMPASS complex

Previous studies suggest that CFP-1 and SET-2 may also have independent functions (Adam et al., 2018, Sommermeyer et al., 2013). In mammals, the loss of *Cfp1* results in a global decrease in cytosine methylation levels, whereas no such results have been reported for the loss of *Set1* (Butler et al., 2008, Boyes and Bird, 1992). Additionally, the gene expression profile of *Cfp1* and *Set1* mutants in yeast suggests that they regulate the expression of some set of genes together and also regulates the expression of a few other genes independently to each other (Mikheyeva et al., 2014). Similarly, a recent study in yeast suggests that CFP-1 can exist in the meiotic DSB formation complex (Adam et al., 2018, Sommermeyer et al., 2013). These findings suggest that CFP-1 in *C. elegans* could also have COMPASS independent functions. Indeed, I observed that a *cfp-1* loss of function mutant suppressed the synthetic Muv phenotype of *lin-15AB(n765)* mutants whereas a *set-2* loss of function mutant did not exhibit this phenotype. Additionally, my results show that CFP-1 interacts genetically with class I HDAC1/2 complexes to promote fertility whereas SET-2 does not. These findings strongly advocate that CFP-1 can work independent of the SET-2/COMPASS complex.

7.1.4 CFP-1 plays a role in organogenesis

CFP-1 is require for pre- and post-embryonic development and for cell fate specification (Carlone and Skalnik, 2001, Carlone et al., 2005, Young et al., 2006). However, early embryonic death due to the loss of Cfp1 in mammals prevents an assessment of its role during cell fate specification. In this study, vulval development in *C. elegans* was used as a model to study the function of CFP-1 in cell fate decisions (Cui et al., 2006b, Harrison et al., 2006, Hill and Sternberg, 1993). Findings from Chapter 6 suggest that *cfp-1* promotes the conversion of vulval precursor cells to adopt vulval fates in SynMuv mutants, which results in the Muv phenotype. This suggests that *cfp-1*antagonises the function of SynMuv genes.

The observed antagonism between *cfp-1* and SynMuv genes could also be conserved in mammals, including humans. It could be important for cancer. Some of the SynMuv genes, such as *lin-35* (Rb homologue in humans) are tumour suppressor genes (Ferguson and Horvitz, 1989, Lu and Horvitz, 1998). Mutation or depletion of Rb is associated with the majority of human cancers (Du and Searle, 2009, Knudsen and Wang, 2010). Here, I observed that the loss of function of CFP-1 suppresses the organogenesis defect associated with SynMuv mutants. It is possible that inhibition of the human homologue of CFP-1 might suppress the effects of a loss of function Rb mutant in humans and hence eliminate or reduce the consequences associated with oncogenic mutations. CFP-1 might be a target for therapeutic intervention.

7.1.5 CFP-1, an epigenetic regulator

Most of the promoters in the humans and *C. elegans* genomes are highly enriched with CpG islands (Chen et al., 2014, Thomson et al., 2010, Brown et al., 2017). These promoters are found to be highly occupied target (HOT) regions bound by many transcription factors (Chen et al., 2014, Li et al., 2016a). It suggests that CpG island promoters can function as a regulatory hub for chromatin regulators and consequently might modulate transcriptional activities. CFP-1 binds to CpG islands and aids in the recruitment of the COMPASS complex for the deposition of H3K4me3 modifications (Chen et al., 2014, Thomson et al., 2010, Brown et al., 2017). Chapter 3 reports that the loss of CFP-1 function results in a dramatic reduction of H3K4me3, and the level of reduction was similar to *set-2* mutants suggesting that CFP-1 functions together with SET-2 to regulate H3K4me3.

At promoter regions, histones with H3K4me3 modifications are subject to high turn-over of acetylation modifications which is mediated by the opposing actions of histone acetyltransferases (HATs) and HDACs (Zhang et al., 2015, Clouaire et al., 2014, Crump et al., 2011, Hsu et al., 2012). HATs deposit the acetylation modification and are often associated with gene expression. In Chapter 6, I found that RNAi knock down of *hat-1* reduced the average brood size of wild-type animals

whereas the RNAi knock down of *hat-1* did not reduce the average brood size of *cfp-1(tm6369)* mutants. This suggests that CFP-1 may oppose the action of HAT-1.

HDACs remove the acetyl group from histones which results in a compacted, transcriptionally silent closed chromatin structure. HDACs are often associated with gene repression and play an important role in maintaining ESCs pluripotency and differentiation. They are also found to be aberrantly expressed in different diseases including cancer (Jamaladdin et al., 2014, Haberland et al., 2009). Here, I found that the RNAi knockdown of key subunits of HDAC1/2 complexes enhanced the poor fertility phenotype of *cfp-1(tm6369)* mutants. Additionally, results show that SIN-3 and HDA-1 promote the function of CFP-1 during vulval development in SynMuv mutants. These findings suggest that CFP-1 interacts genetically with HDAC1/2 complexes in *C. elegans* development. The observed genetic interaction could also be conserved in mammals and may play a role in various biological events such as gene expression, DNA repair, apoptosis, stress responses, genomic stability, cell cycle, etc. (Roos and Krumm, 2016, Ropero and Esteller, 2007). CFP-1 contains a DNA binding domain, and it is possible that CFP-1 could help in the recruitment of the HDAC complex to promote the role of HDAC in gene expression.

Chapter 6 shows that PBRM-1 antagonises the function of CFP-1 in vulval development. PBRM-1 is a subunit of the SWI/SNF complex, a multiprotein complex that uses ATP to alter the interaction between DNA and nucleosomes (Dimova et al., 1999, Moreira and Holmberg, 1999, Murphy et al., 1999). The SWI/SNF complex is associated with both gene expression and repression (Dimova et al., 1999, Moreira and Holmberg, 1999, Murphy et al., 1999). Acetylation of histones H3K4 and H3K14 aids in the recruitment of SWI/SNF chromatin-remodelling complexes (Agalioti et al., 2002, Peserico and Simone, 2011). On the other hand, the NuRD complex, which harbours the ATP-dependent helicase CHD3 and histone deacetylases HDAC1 and HDAC2, antagonises SWI/SNF function to control chromatin state and transcription (Mohd-Sarip et al., 2017, Gao et al., 2009). In this study, it was found that CFP-1 interacts genetically with HDAC1/2 complexes, including the NuRD complex, in *C. elegans* development. It is possible

that CFP-1 interferes in the recruitment of the SWI/SNF complex by recruiting HDAC complexes to regulate gene expression.

Collectively, these findings suggest that CFP-1 could act as an epigenetic regulator that functions within the COMPASS complex, and also interacts with HDAC1/2 complexes and PBRM-1. It is possible that CFP-1 functions as an epigenetic regulator that controls the histone modification landscape by interacting with other chromatin regulators.

7.2 Future direction

This project was designed to study the function of CFP-1 in *C. elegans* development, gene induction and to identify novel genetic regulators of CFP-1. Findings from this study suggest that CFP-1 is required for the fertility and development of *C. elegans*. Further experiments such as the gene expression profile of *cfp-1* mutants and the gene ontology analysis of affected genes will help to identify the biological pathways affected in the loss of CFP-1 function.

I also found that in *cfp-1(tm6369)* or *set-2(bn129)* mutants, the expression of stress-inducible genes was significantly upregulated, suggesting that loss of H3K4me3 promotes gene induction. H3K4me3 could play a role in the activation of stress-inducible genes such as HSF-1 genes. For example, H3K4me3 might restrict the nuclear trans localisation of HSF-1. Thus, in the loss of H3K4me3, the rate of nuclear translocation of HSF-1 could have increased, leading to increased gene expression. This can be investigated by using a reported HSF-1 nuclear localisation assay in wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants. It is also possible that H3K4me3 could play an inhibitory role in the recruitment of HSF-1 and other transcription factors. Thus, in the loss of H3K4me3 rate of recruitment of HSF-1 or other TFs could increase during stress. Future experiments such as ChIP-seq assays to map HSF-1 binding sites in wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants could help to identify whether COMPASS mutants or H3K4me3 play a role in the recruitment of HSF-1.

Alternatively, CFP-1 or SET-2 might also control the transcription initiation of stress-inducible genes. One of the features of stress-inducible genes such as heat shock response genes is the presence of paused Pol II. CFP-1 or SET-2 might promote promoter pausing of Pol II, and act as a switch to allow only limited transcription to be initiated. This can be investigated by mapping the occupancy of RNA Pol II in *cfp-1(tm6369)* and *set-2(bn129)* mutants. To do so, I can use AMA-1 (AMA-1 is a large subunit of Pol II and is required for transcription) antibody, and use antibodies against Phospho Ser-5 Pol II (initiating Pol II) and Phospho Ser-2 Pol II (elongating Pol II). This will show us how H3K4me3 affects RNA Pol II activity

during stress. Furthermore, it will also help us to identify genes with similar patterns. Gene ontology analysis of those genes can be done to identify the biological pathways that could be linked with the stress response.

This study also identifies the function of CFP-1 outside the SET-2/COMPASS complex in *C. elegans* development. I found that CFP-1 interacts genetically with HDAC1/2 complexes in *C. elegans* development. Further experiments will help us to understand the interaction in detail. For example, ChIP of components of the HDAC1/2 complex in *cfp-1(tm6369)* and *set-2(bn129)* mutants could provide insight on whether CFP-1 could help in the recruitment of HDAC1/2 complex. Similarly, co-immunoprecipitation experiments will reveal the physical interaction between CFP-1 and HDAC1/2 complex.

CFP-1 might associate with other proteins. Therefore, to identify proteins only associated with CFP-1 but not SET-2, immunoprecipitation in combination with mass spectrometry (IP-MS) could be used to determine the identity of CFP-1-specific associated proteins. One of the challenging aspects to perform IP-MS is the unavailability of antibodies for both *C. elegans* CFP-1 and SET-2 proteins. To overcome this CFP-1 and SET-2 could be FLAG-tagged using clustered regularly interspaced short palindromic repeats (CRISPR) reagents. Immunoprecipitation followed by MS should reveal the proteins associated with CFP-1, but not SET-2. To identify the functional relationship between newly identified proteins and CFP-1, genetic dissections assays using *C. elegans* mutants can be performed.

As discussed above, CFP-1 might act as an epigenetic regulator that interacts with other chromatin regulators to controls the histone modification landscape. To investigate this, histone modifications in the absence of CFP-1 can be explored using MS. This will reveal the histone post-translational modifications (PTMs) associated with *cfp-1(tm6369)* mutants and also any altered modification compared to the wild-type and/or *set-2(bn129)* mutants. Furthermore, to investigate the localisation of any altered modifications, chromatin immunoprecipitation of specific modifications can be performed. This will reveal the change of a specific histone PTM which could be either at enhancer or promoter regions. This could help us to

find the promoters or enhancers regulated by CFP-1 with or without SET-2.

Next, I could investigate the SET-2/COMPASS independent function of CFP-1 at the transcription level. To do so, I could perform RNAseq of both mutants and identify the genes that are regulated only by CFP-1 but not by SET-2. I can use gene ontology analysis to identify the function of affected genes.

7.3 Conclusion

Epigenetic regulators control fundamental biological processes within the nucleus. Mis-regulation of epigenetic regulators is associated with developmental defects and many human diseases including cancer. Therefore, it is crucial to understand how epigenetic regulators are regulated and their roles in controlling gene expression in metazoan development. The work presented in this thesis focusses on an epigenetic regulator, CFP-1, and its importance in *C. elegans* development.

The key findings of this study are:

1. CFP-1 plays an important role in the regulation of fertility and the development of the organism.

2. CFP-1 can play a repressive role in the induction of heat and salt-inducible genes.

3. CFP-1 interacts genetically with other chromatin regulators in ectopic vulval cell fate specification.

4. CFP-1 interacts genetically with Class 1 HDAC1/2 complexes to promote fertility in *C. elegans.* This novel interaction is independent of SET-2, suggesting a COMPASS-independent function of CFP-1.

Since CFP-1 is evolutionarily conserved from yeast to human, it is possible that the observed genetic interactions between CFP-1 and other chromatin regulators identified in this work using *C. elegans,* could also be conserved in humans. Further research to understand the mechanism of interaction between CFP-1 and these regulators could help to understand the role of CFP-1 in cell fate specification and in different diseases including cancer.

References

Acquaviva, L., Szekvolgyi, L., Dichtl, B., Dichtl, B. S., De La Roche Saint Andre, C., Nicolas, A. & Geli, V. 2013. The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. *Science,* 339(6116), pp.215-218.

Adam, C., Guérois, R., Citarella, A., Verardi, L., Adolphe, F., Béneut, C., Sommermeyer, V., Ramus, C., Govin, J., Couté, Y. & Borde, V. 2018. The PHD finger protein Spp1 has distinct functions in the Set1 and the meiotic DSB formation complexes. *PLOS Genetics,* 14(2), pp.e1007223.

Adelman, K. & Lis, J. T. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nature Reviews Genetics,* 13(10), pp.720-731.

Agalioti, T., Chen, G. & Thanos, D. 2002. Deciphering the transcriptional histone acetylation code for a human gene. *Cell,* 111(3), pp.381-392.

Agger, K., Cloos, P. A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A. E. & Helin, K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature,* 449(7163), pp.731-734.

Ahringer, J. & Gasser, S. M. 2018. Repressive Chromatin in Caenorhabditis elegans: Establishment, Composition, and Function. *Genetics,* 208(2), pp.491-511.

Allis, Berger, S. L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhattar, R., Shilatifard, A., Workman, J. & Zhang, Y. 2007. New nomenclature for chromatin-modifying enzymes. *Cell,* 131(4), pp.633-636.

Allis & Jenuwein, T. 2016. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics,* 17(8), pp.487-500.

Allshire, R. C. & Madhani, H. D. 2018. Ten principles of heterochromatin formation and function. *Nature Reviews Molecular Cell Biology,* 19(4), pp.229-244.

Andersen, E. C. 2007. *The synthetic multivulva genes and their suppressors regulate opposing cell fates through chromatin remodeling.* Ph.D. Thesis, Massachuetts Institute of Technology.

Ang, Rivas, R. N., Ribeiro, A. J. S., Srivas, R., Rivera, J., Stone, N. R., Pratt, K., Mohamed, T. M. A., Fu, J. D., Spencer, C. I., Tippens, N. D., Li, M., Narasimha, A., Radzinsky, E., Moon-Grady, A. J., Yu, H., Pruitt, B. L., Snyder, M. P. & Srivastava, D. 2016. Disease Model of GATA4 Mutation Reveals Transcription Factor Cooperativity in Human Cardiogenesis. *Cell,* 167(7), pp.1734-1749.

Ang, Tsai, S.-Y., Lee, D.-F., Monk, J., Su, J., Ratnakumar, K., Ding, J., Ge, Y., Darr, H., Chang, B., Wang, J., Rendl, M., Bernstein, E., Schaniel, C. & Lemischka, I. R. 2011. Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell,* 145(2), pp.183-197.

Annelie Strålfors, K. E. 2011. Heterochromatin and Euchromatin—Organization, Boundaries, and Gene Regulation. *Reviews in Cell Biology and Molecular Medicine.* Wiley-VCHa Verlag GmbH & Co. KGaA, pp.171–189.

Araki, Y., Wang, Z., Zang, C., Wood, W. H., 3rd, Schones, D., Cui, K., Roh, T. Y., Lhotsky, B., Wersto, R. P., Peng, W., Becker, K. G., Zhao, K. & Weng, N. P. 2009. Genome-wide analysis of histone methylation reveals chromatin state-based regulation of gene transcription and function of memory CD8+ T cells. *Immunity,* 30(6), pp.912-925.

Ardehali, M. B., Mei, A., Zobeck, K. L., Caron, M., Lis, J. T. & Kusch, T. 2011. Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *EMBO Journal,* 30(14), pp.2817-2828.

Armstrong, L., Hughes, O., Yung, S., Hyslop, L., Stewart, R., Wappler, I., Peters, H., Walter, T., Stojkovic, P., Evans, J., Stojkovic, M. & Lako, M. 2006. The role of PI3K/AKT, MAPK/ERK and NFkappabeta signalling in the maintenance of human embryonic stem cell pluripotency and viability highlighted by transcriptional profiling and functional analysis. *Human Molecular Genetics,* 15(11), pp.1894-1913.

Aroian, R. V., Koga, M., Mendel, J. E., Ohshima, Y. & Sternberg, P. W. 1990. The let-23 gene necessary for Caenorhabditis elegans vulval induction encodes a tyrosine kinase of the EGF receptor subfamily. *Nature,* 348(6303), pp.693-699.

Avvakumov, N. & Cote, J. 2007. The MYST family of histone acetyltransferases and their intimate links to cancer. *Oncogene,* 26(37), pp.5395-5407.

Ayton, P. M., Chen, E. H. & Cleary, M. L. 2004. Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. *Molecular and Cellular Biology,* 24(23), pp.10470-10478.

Bannister & Kouzarides, T. 2011. Regulation of chromatin by histone modifications. *Cell Research,* 21(3), pp.381-395.

Bannister, Schneider, R., Myers, F. A., Thorne, A. W., Crane-Robinson, C. & Kouzarides, T. 2005. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *Journal of Biological Chemistry,* 280(18), pp.17732-17736.

Bannister, Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. & Kouzarides, T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature,* 410(6824), pp.120-124.

Barnes, C. J., Li, F., Mandal, M., Yang, Z., Sahin, A. A. & Kumar, R. 2002. Heregulin induces expression, ATPase activity, and nuclear localization of G3BP, a Ras signaling component, in human breast tumors. *Cancer Research,* 62(5), pp.1251-1255.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. & Zhao, K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell,* 129(4), pp.823-837.

Baugh, L. R., Wen, J. C., Hill, A. A., Slonim, D. K., Brown, E. L. & Hunter, C. P. 2005. Synthetic lethal analysis of Caenorhabditis elegans posterior embryonic patterning genes identifies conserved genetic interactions. *Genome Biology,* 6(5), pp.R45.

Baumann, K. 2017. GENOME ORGANIZATION: A vision of 3D chromatin organization. *Nature Reviews Molecular Cell Biology,* 18(9), pp.532.

Beilharz, T. H., Harrison, P. F., Miles, D. M., See, M. M., Le, U. M., Kalanon, M., Curtis, M. J., Hasan, Q., Saksouk, J., Margaritis, T., Holstege, F., Geli, V. & Dichtl, B. 2017. Coordination of Cell Cycle Progression and Mitotic Spindle Assembly Involves Histone H3 Lysine 4 Methylation by Set1/COMPASS. *Genetics,* 205(1), pp.185-199.

Beira, J. V. & Paro, R. 2016. The legacy of Drosophila imaginal discs. *Chromosoma,* 125(4), pp.573-592.

Beitel, G. J., Clark, S. G. & Horvitz, H. R. 1990. Caenorhabditis elegans ras gene let-60 acts as a switch in the pathway of vulval induction. *Nature,* 348(6301), pp.503-509.

Beitel, G. J., Tuck, S., Greenwald, I. & Horvitz, H. R. 1995. The Caenorhabditis elegans gene lin-1 encodes an ETS-domain protein and defines a branch of the vulval induction pathway. *Genes & Development,* 9(24), pp.3149-3162.

Benayoun, B. A., Pollina, E. A., Ucar, D., Mahmoudi, S., Karra, K., Wong, E. D., Devarajan, K., Daugherty, A. C., Kundaje, A. B., Mancini, E., Hitz, B. C., Gupta, R., Rando, T. A., Baker, J. C., Snyder, M. P., Cherry, J. M. & Brunet, A. 2014. H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell,* 158(3), pp.673-688.

Bender, A. M., Kirienko, N. V., Olson, S. K., Esko, J. D. & Fay, D. S. 2007. lin-35/Rb and the CoREST ortholog spr-1 coordinately regulate vulval morphogenesis and gonad development in C. elegans. *Developmental Biology,* 302(2), pp.448-462.

Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T. & Schreiber, S. L. 2002. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proceedings of the National Academy of Sciences of the United States of America,* 99(13), pp.8695-8700.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M. & Plath, K. 2006. A bivalent chromatin structure marks key developmental genes in embryonicstem cells. *Cell,* 125(2), pp.315-326.

Bertero, A., Madrigal, P., Galli, A., Hubner, N. C., Moreno, I., Burks, D., Brown, S., Pedersen, R. A., Gaffney, D., Mendjan, S., Pauklin, S. & Vallier, L. 2015.

Activin/nodal signaling and NANOG orchestrate human embryonic stem cell fate decisions by controlling the H3K4me3 chromatin mark. *Genes & Development,* 29(7), pp.702-717.

Bian, C., Xu, C., Ruan, J., Lee, K. K., Burke, T. L., Tempel, W., Barsyte, D., Li, J., Wu, M., Zhou, B. O., Fleharty, B. E., Paulson, A., Allali-Hassani, A., Zhou, J. Q., Mer, G., Grant, P. A., Workman, J. L., Zang, J. & Min, J. 2011. Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO Journal,* 30(14), pp.2829-2842.

Birke, M., Schreiner, S., Garcia-Cuellar, M. P., Mahr, K., Titgemeyer, F. & Slany, R. K. 2002. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Research,* 30(4), pp.958-965.

Bledau, A. S., Schmidt, K., Neumann, K., Hill, U., Ciotta, G., Gupta, A., Torres, D. C., Fu, J., Kranz, A., Stewart, A. F. & Anastassiadis, K. 2014. The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Development,* 141(5), pp.1022-1035.

Bochynska, A., Luscher-Firzlaff, J. & Luscher, B. 2018. Modes of Interaction of KMT2 Histone H3 Lysine 4 Methyltransferase/COMPASS Complexes with Chromatin. *Cells,* 7(3), pp.1-31.

Boyes, J. & Bird, A. 1992. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO Journal,* 11(1), pp.327-333.

Brenner, S. 1974. The genetics of Caenorhabditis elegans. *Genetics,* 77(1), pp.71-94.

Brown, D. A., Di Cerbo, V., Feldmann, A., Ahn, J., Ito, S., Blackledge, N. P., Nakayama, M., Mcclellan, M., Dimitrova, E., Turberfield, A. H., Long, H. K., King, H. W., Kriaucionis, S., Schermelleh, L., Kutateladze, T. G., Koseki, H. & Klose, R. J. 2017. The SET1 Complex Selects Actively Transcribed Target Genes via Multivalent Interaction with CpG Island Chromatin. *Cell reports,* 20(10), pp.2313-2327.

Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y. & Allis, C. D. 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell,* 84(6), pp.843-851.

Brunmeir, R., Lagger, S. & Seiser, C. 2009. Histone deacetylase HDAC1/HDAC2-controlled embryonic development and cell differentiation. *International Journal of Developmental Biology,* 53(2-3), pp.275-289.

Brunquell, J., Morris, S., Lu, Y., Cheng, F. & Westerheide, S. D. 2016. The genome-wide role of HSF-1 in the regulation of gene expression in Caenorhabditis elegans. *BMC Genomics,* 17(1), pp.559-577.

Buratowski, S. 2009. Progression through the RNA Polymerase II CTD Cycle. *Molecular Cell,* 36(4), pp.541-546.

Butler, J. S., Lee, J. H. & Skalnik, D. G. 2008. CFP1 interacts with DNMT1 independently of association with the Setd1 Histone H3K4 methyltransferase complexes. *DNA Cell Biology,* 27(10), pp.533-543.

Canovas, S., Cibelli, J. B. & Ross, P. J. 2012. Jumonji domain-containing protein 3 regulates histone 3 lysine 27 methylation during bovine preimplantation development. *Proceedings of the National Academy of Sciences of the United States of America,* 109(7), pp.2400-2405.

Cao, K., Collings, C. K., Marshall, S. A., Morgan, M. A., Rendleman, E. J., Wang, L., Sze, C. C., Sun, T., Bartom, E. T. & Shilatifard, A. 2017. SET1A/COMPASS and shadow enhancers in the regulation of homeotic gene expression. *Genes & Development,* 31(8), pp.787-801.

Carlone, D. L., Hart, S. R., Ladd, P. D. & Skalnik, D. G. 2002. Cloning and characterization of the gene encoding the mouse homologue of CpG binding protein. *Gene,* 295(1), pp.71-77.

Carlone, D. L., Lee, J. H., Young, S. R., Dobrota, E., Butler, J. S., Ruiz, J. & Skalnik, D. G. 2005. Reduced genomic cytosine methylation and defective cellular differentiationin embryonic stem cells lacking CpG binding protein. *Molecular and Cellular Biology,* 25(12), pp.4881-4891.

Carlone, D. L. & Skalnik, D. G. 2001. CpG binding protein is crucial for early embryonic development. *Molecular and Cellular Biology,* 21(22), pp.7601-7606.

Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W. J., Anderson, S., Yates, J., Washburn, M. P. & Workman, J. L. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell,* 123(4), pp.581-592.

Carrozza, M. J., Utley, R. T., Workman, J. L. & Cote, J. 2003. The diverse functions of histone acetyltransferase complexes. *Trends in Genetics,* 19(6), pp.321-329.

Castelli-Gair, J. & Akam, M. 1995. How the Hox gene Ultrabithorax specifies two different segments: the significance of spatial and temporal regulation within metameres. *Development,* 121(9), pp.2973-2982.

Cayrol, C., Lacroix, C., Mathe, C., Ecochard, V., Ceribelli, M., Loreau, E., Lazar, V., Dessen, P., Mantovani, R., Aguilar, L. & Girard, J. P. 2007. The THAP-zinc finger protein THAP1 regulates endothelial cell proliferation through modulation of pRB/E2F cell-cycle target genes. *Blood,* 109(2), pp.584-594.

Ceol, C. J. & Horvitz, H. R. 2001. dpl-1 DP and efl-1 E2F act with lin-35 Rb to antagonize Ras signaling in C. elegans vulval development. *Molecular Cell,* 7(3), pp.461-473.

Ceol, C. J. & Horvitz, H. R. 2004. A new class of C elegans synMuv genes implicates a Tip60/NuA4-like HAT complex as a negative regulator of ras signaling. *Developmental Cell,* 6(4), pp.563-576.

Ceol, C. J., Stegmeier, F., Harrison, M. M. & Horvitz, H. R. 2006. Identification and classification of genes that act antagonistically to let-60 Ras signaling in Caenorhabditis elegans vulval development. *Genetics,* 173(2), pp.709-726.

Chakravarty, S., Pathak, S. S., Maitra, S., Khandelwal, N., Chandra, K. B. & Kumar, A. 2014. Chapter Four - Epigenetic Regulatory Mechanisms in Stress-Induced Behavior. In: PANDEY, S. C. ed. *International Review of Neurobiology.* Academic Press, pp.117-154.

Chang, C., Hopper, N. A. & Sternberg, P. W. 2000. Caenorhabditis elegans SOS-1 is necessary for multiple RAS-mediated developmental signals. *EMBO Journal,* 19(13), pp.3283-3294.

Chang, S., Mckinsey, T. A., Zhang, C. L., Richardson, J. A., Hill, J. A. & Olson, E. N. 2004. Histone deacetylases 5 and 9 govern responsiveness of the heart to a subset of stress signals and play redundant roles in heart development. *Molecular and Cellular Biology,* 24(19), pp.8467-8476.

Chang, S., Young, B. D., Li, S., Qi, X., Richardson, J. A. & Olson, E. N. 2006. Histone deacetylase 7 maintains vascular integrity by repressing matrix metalloproteinase 10. *Cell,* 126(2), pp.321-334.

Chen, H., Lin, R. J., Schiltz, R. L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M. L., Nakatani, Y. & Evans, R. M. 1997. Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell,* 90(3), pp.569-580.

Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., Rodriguez, B., Xi, Y., Xia, Z., Chen, X., Shi, X., Wang, Q. & Li, W. 2015. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nature Genetics,* 47(10), pp.1149-1157.

Chen, L. S., Krause, M., Sepanski, M. & Fire, A. 1994. The Caenorhabditis Elegans Myod Homolog Hlh-1 Is Essential for Proper Muscle Function and Complete Morphogenesis. *Development,* 120(6), pp.1631-1641.

Chen, N. & Greenwald, I. 2004. The lateral signal for LIN-12/Notch in C elegans vulval development comprises redundant secreted and transmembrane DSL proteins. *Developmental Cell,* 6(2), pp.183-192.

Chen, R. A., Stempor, P., Down, T. A., Zeiser, E., Feuer, S. K. & Ahringer, J. 2014. Extreme HOT regions are CpG-dense promoters in C. elegans and humans. *Genome Research,* 24(7), pp.1138-1146.

Cho, Y. W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G. R., Copeland, T. D., Kalkum, M. & Ge, K. 2007. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4methyltransferase complex. *Journal of Biological Chemistry,* 282(28), pp.20395-20406.

Choe, K. P. & Strange, K. 2007. Molecular and genetic characterization of osmosensing and signal transduction in the nematode Caenorhabditis elegans. *The FEBS journal,* 274(22), pp.5782-5789.

Choy, S. W., Wong, Y. M., Ho, S. H. & Chow, K. L. 2007. C. elegans SIN-3 and its associated HDAC corepressor complex act as mediators of male sensory ray development. *Biochemical and Biophysical Research Communications,* 358(3), pp.802-807.

Christensen, S., Kodoyianni, V., Bosenberg, M., Friedman, L. & Kimble, J. 1996. lag-1, a gene required for lin-12 and glp-1 signaling in Caenorhabditis elegans, is homologous to human CBF1 and Drosophila Su(H). *Development,* 122(5), pp.1373-1383.

Chun, K. T., Li, B., Dobrota, E., Tate, C., Lee, J. H., Khan, S., Haneline, L., Hogenesch, H. & Skalnik, D. G. 2014. The epigenetic regulator CXXC finger protein 1 is essential for murine hematopoiesis. *PLOS One,* 9(12), pp.e113745.

Cierpicki, T., Risner, L. E., Grembecka, J., Lukasik, S. M., Popovic, R., Omonkowska, M., Shultis, D. D., Zeleznik-Le, N. J. & Bushweller, J. H. 2010. Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. *Nature Structural & Molecular Biology,* 17(1), pp.62-68.

Clark, S. G., Lu, X. & Horvitz, H. R. 1994. The Caenorhabditis elegans locus lin-15, a negative regulator of a tyrosine kinase signaling pathway, encodes two different proteins. *Genetics,* 137(4), pp.987-997.

Clark, S. G., Stern, M. J. & Horvitz, H. R. 1992. C. elegans cell-signalling gene sem-5 encodes a protein with SH2 and SH3 domains. *Nature,* 356(6367), pp.340-344.

Clouaire, T., Webb, S. & Bird, A. 2014. Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells. *Genome Biology,* 15(9), pp.451-467.

Clouaire, T., Webb, S., Skene, P., Illingworth, R., Kerr, A., Andrews, R., Lee, J. H., Skalnik, D. & Bird, A. 2012. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes & Development,* 26(15), pp.1714-1728.

Corsi, A. K. 2006. A Biochemist's Guide to C. elegans. *Analytical biochemistry,* 359(1), pp.1-17.

Corsi, A. K., Wightman, B. & Chalfie, M. 2015. A Transparent Window into Biology: A Primer on Caenorhabditis elegans. *Genetics,* 200(2), pp.387-407.

Coustham, V., Bedet, C., Monier, K., Schott, S., Karali, M. & Palladino, F. 2006. The C. elegans HP1 homologue HPL-2 and the LIN-13 zinc finger protein form a complex implicated in vulval development. *Developmental Biology,* 297(2), pp.308-322.

Couteau, F., Guerry, F., Muller, F. & Palladino, F. 2002. A heterochromatin protein 1 homologue in Caenorhabditis elegans acts in germline and vulval development. *EMBO Reports,* 3(3), pp.235-241.

Crump, N. T., Hazzalin, C. A., Bowers, E. M., Alani, R. M., Cole, P. A. & Mahadevan, L. C. 2011. Dynamic acetylation of all lysine-4 trimethylated histone H3 is evolutionarily conserved and mediated by p300/CBP. *Proceedings of the National Academy of Sciences of the United States of America,* 108(19), pp.7814-7819.

Cui, M., Chen, J., Myers, T. R., Hwang, B. J., Sternberg, P. W., Greenwald, I. & Han, M. 2006a. SynMuv genes redundantly inhibit lin-3/EGF expression to prevent inappropriate vulval induction in C. elegans. *Developmental Cell,* 10(5), pp.667-672.

Cui, M., Kim, E. B. & Han, M. 2006b. Diverse chromatin remodeling genes antagonize the Rb-involved SynMuv pathways in C. elegans. *PLOS Genetics,* 2(5), pp.e74.

Cunha, P. M., Sandmann, T., Gustafson, E. H., Ciglar, L., Eichenlaub, M. P. & Furlong, E. E. 2010. Combinatorial binding leads to diverse regulatory responses: Lmd is a tissue-specific modulator of Mef2 activity. *PLOS Genetics,* 6(7), pp.e1001014.

Dahl, J. A., Jung, I., Aanes, H., Greggains, G. D., Manaf, A., Lerdrup, M., Li, G. Q., Kuan, S., Li, B., Lee, A. Y., Preissl, S., Jermstad, I., Haugen, M. H., Suganthan, R., Bjoras, M., Hansen, K., Dalen, K. T., Fedorcsak, P., Ren, B. & Klungland, A.

2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature,* 537(7621), pp.548-553.

Dai, Y. & Faller, D. V. 2008. Transcription Regulation by Class III Histone Deacetylases (HDACs)—Sirtuins. *Translational Oncogenomics,* 3, pp.53-65.

Davison, E. M., Harrison, M. M., Walhout, A. J. M., Vidal, M. & Horvitz, H. R. 2005. lin-8, which antagonizes Caenorhabditis elegans Ras-mediated vulval induction, encodes a novel nuclear protein that interacts with the LIN-35 Rb protein. *Genetics,* 171(3), pp.1017-1031.

Davison, E. M., Saffer, A. M., Huang, L. S., Demodena, J., Sternberg, P. W. & Horvitz, H. R. 2011. The LIN-15A and LIN-56 Transcriptional Regulators Interact to Negatively Regulate EGF/Ras Signaling in Caenorhabditis elegans Vulval Cell-Fate Determination. *Genetics,* 187(3), pp.803-815.

De Ruijter, A. J., Van Gennip, A. H., Caron, H. N., Kemp, S. & Van Kuilenburg, A. B. 2003. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochemical Journal,* 370(Pt 3), pp.737-749.

Deaton, A. M. & Bird, A. 2010. CpG islands and the regulation of transcription. *Genes & Development,* 25(10), pp.1010-1022.

Degner, J. F., Pai, A. A., Pique-Regi, R., Veyrieras, J. B., Gaffney, D. J., Pickrell, J. K., De Leon, S., Michelini, K., Lewellen, N., Crawford, G. E., Stephens, M., Gilad, Y. & Pritchard, J. K. 2012. DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature,* 482(7385), pp.390-394.

Dehe, P. M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R., Rodriguez-Gil, A., Mkandawire, M., Landsberg, K., Shevchenko, A., Shevchenko, A., Rosaleny, L. E., Tordera, V., Chavez, S., Stewart, A. F. & Geli, V. 2006. Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *Journal of Biological Chemistry,* 281(46), pp.35404-35412.

Dejosez, M., Krumenacker, J. S., Zitur, L. J., Passeri, M., Chu, L. F., Zhou, S. Y., Thomson, J. A. & Zwaka, T. P. 2008. Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. *Cell,* 133(7), pp.1162-1174.

Delcuve, G. P., Khan, D. H. & Davie, J. R. 2012. Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. *Clinical Epigenetics,* 4(1), pp.1-13.

Deng, C., Li, Y., Zhou, L., Cho, J., Patel, B., Terada, N., Li, Y., Bungert, J., Qiu, Y. & Huang, S. 2016. HoxBlinc RNA recruits Set1/MLL complexes to activate Hox gene expression patterns and mesoderm lineage development. *Cell reports,* 14(1), pp.103-114.

Denslow, S. A. & Wade, P. A. 2007. The human Mi-2/NuRD complex and gene regulation. *Oncogene,* 26(37), pp.5433-5438.

Dimova, D., Nackerdien, Z., Furgeson, S., Eguchi, S. & Osley, M. A. 1999. A role for transcriptional repressors in targeting the yeast Swi/Snf complex. *Molecular Cell,* 4(1), pp.75-83.

Doetzlhofer, A., Rotheneder, H., Lagger, G., Koranda, M., Kurtev, V., Brosch, G., Wintersberger, E. & Seiser, C. 1999. Histone deacetylase 1 can repress transcription by binding to Sp1. *Molecular and Cellular Biology,* 19(8), pp.5504-5511.

Doi, M., Hirayama, J. & Sassone-Corsi, P. 2006. Circadian regulator CLOCK is a histone acetyltransferase. *Cell,* 125(3), pp.497-508.

Dong, X. & Weng, Z. 2013. The correlation between histone modifications and gene expression. *Epigenomics,* 5(2), pp.113-116.

Dong, X. J., Tsuji, J., Labadorf, A., Roussos, P., Chen, J. F., Myers, R. H., Akbarian, S. & Weng, Z. P. 2015. The Role of H3K4me3 in Transcriptional Regulation Is Altered in Huntington's Disease. *PLOS One,* 10(12), pp.e0144398.

Dover, J., Schneider, J., Tawiah-Boateng, M. A., Wood, A., Dean, K., Johnston, M. & Shilatifard, A. 2002. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *Journal of Biological Chemistry,* 277(32), pp.28368-28371.

Doyon, Y., Cayrou, C., Ullah, M., Landry, A. J., Cote, V., Selleck, W., Lane, W. S., Tan, S., Yang, X. J. & Cote, J. 2006. ING tumor suppressor proteins are critical

regulators of chromatin acetylation required for genome expression and perpetuation. *Molecular Cell,* 21(1), pp.51-64.

Doyon, Y., Selleck, W., Lane, W. S., Tan, S. & Cote, J. 2004. Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Molecular and Cellular Biology,* 24(5), pp.1884-1896.

Dreijerink, K. M., Varier, R. A., Van Beekum, O., Jeninga, E. H., Hoppener, J. W., Lips, C. J., Kummer, J. A., Kalkhoven, E. & Timmers, H. T. 2009. The multiple endocrine neoplasia type 1 (MEN1) tumor suppressor regulates peroxisome proliferator-activated receptor gamma-dependent adipocyte differentiation. *Molecular and Cellular Biology,* 29(18), pp.5060-5069.

Du, W. & Searle, J. S. 2009. The rb pathway and cancer therapeutics. *Current Drug Targets,* 10(7), pp.581-589.

Dufourcq, P., Victor, M., Gay, F., Calvo, D., Hodgkin, J. & Shi, Y. 2002. Functional requirement for histone deacetylase 1 in Caenorhabditis elegans gonadogenesis. *Molecular and Cellular Biology,* 22(9), pp.3024-3034.

Eastburn, D. J. & Han, M. 2005. A gain-of-function allele of cbp-1, the Caenorhabditis elegans ortholog of the mammalian CBP/p300 gene, causes an increase in histone acetyltransferase activity and antagonism of activated Ras. *Molecular and Cellular Biology,* 25(21), pp.9427-9434.

Eissenberg, J. C. & Shilatifard, A. 2010. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Developmental Biology,* 339(2), pp.240-249.

Engelen, E., Brandsma, J. H., Moen, M. J., Signorile, L., Dekkers, D. H., Demmers, J., Kockx, C. E., Ozgur, Z., Van, I. W. F., Van Den Berg, D. L. & Poot, R. A. 2015. Proteins that bind regulatory regions identified by histone modification chromatin immunoprecipitations and mass spectrometry. *Nature Communications,* 20(6), pp.7155.

Erdelyi, P., Wang, X., Suleski, M. & Wicky, C. 2017. A Network of Chromatin Factors Is Regulating the Transition to Postembryonic Development in Caenorhabditis elegans. *G3-Genes Genomes Genetics,* 7(2), pp.343-353.

Ermolaeva, M. A., Segref, A., Dakhovnik, A., Ou, H. L., Schneider, J. I., Utermohlen, O., Hoppe, T. & Schumacher, B. 2013. DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance. *Nature,* 501(7467), pp.416-420.

Fang, L., Zhang, J., Zhang, H., Yang, X., Jin, X., Zhang, L., Skalnik, D. G., Jin, Y., Zhang, Y., Huang, X., Li, J. & Wong, J. 2016. H3K4 Methyltransferase Set1a Is A Key Oct4 Coactivator Essential for Generation of Oct4 Positive Inner Cell Mass. *Stem Cells,* 34(3), pp.565-580.

Faucher, D. & Wellinger, R. J. 2010. Methylated H3K4, a transcription-associated histone modification, is involved in the DNA damage response pathway. *PLOS Genetics,* 6(8), pp.e1001082.

Fay, D. 2006. Genetic mapping and manipulation: chapter 1--Introduction and basics. *WormBook*, pp.1-12.

Fay, D. S. & Yochem, J. 2007. The SynMuv genes of Caenorhabditis elegans in vulval development and beyond. *Developmental biology,* 306(1), pp.1-9.

Ferguson, E. L. & Horvitz, H. R. 1989. The multivulva phenotype of certain Caenorhabditis elegans mutants results from defects in two functionally redundant pathways. *Genetics,* 123(1), pp.109-121.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature,* 391(6669), pp.806-811.

Fisher, K., Southall, S. M., Wilson, J. R. & Poulin, G. B. 2010. Methylation and demethylation activities of a C. elegans MLL-like complex attenuate RAS signalling. *Developmental Biology,* 341(1), pp.142-153.

Frye, R. A. 2000. Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochemical and Biophysical Research Communications,* 273(2), pp.793-798.

Gao, H., Lukin, K., Ramirez, J., Fields, S., Lopez, D. & Hagman, J. 2009. Opposing effects of SWI/SNF and Mi-2/NuRD chromatin remodeling complexes on

epigenetic reprogramming by EBF and Pax5. *Proceedings of the National Academy of Sciences of the United States of America,* 106(27), pp.11258-11263.

Gao, L., Cueto, M. A., Asselbergs, F. & Atadja, P. 2002. Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. *Journal of Biological Chemistry,* 277(28), pp.25748-25755.

Gartner, A., Boag, P. R. & Blackwell, T. K. 2008. Germline survival and apoptosis. *WormBook*, pp.1-20.

Gauthier, K. & Rocheleau, C. E. 2017. C. elegans Vulva Induction: An In Vivo Model to Study Epidermal Growth Factor Receptor Signaling and Trafficking. *Methods in Molecular Biology,* 1652, pp.43-61.

Gomez-Pastor, R., Burchfiel, E. T. & Thiele, D. J. 2018. Regulation of heat shock transcription factors and their roles in physiology and disease. *Nature Reviews Molecular Cell Biology,* 19(1), pp.4-19.

Gonzalez, I. & Busturia, A. 2009. High levels of dRYBP induce apoptosis in Drosophila imaginal cells through the activation of reaper and the requirement of trithorax, dredd and dFADD. *Cell Research,* 19(6), pp.747-757.

Goo, Y. H., Sohn, Y. C., Kim, D. H., Kim, S. W., Kang, M. J., Jung, D. J., Kwak, E., Barlev, N. A., Berger, S. L. & Chow, V. T. 2003. Activating signal cointegrator 2 belongs to a novel steady-state complex thatcontains a subset of trithorax group proteins. *Molecular and Cellular Biology,* 23(1), pp.140-149.

Greenwald, I. 1998. LIN-12/Notch signaling: lessons from worms and flies. *Genes & Development,* 12(12), pp.1751-1762.

Greer, E. L., Maures, T. J., Hauswirth, A. G., Green, E. M., Leeman, D. S., Maro, G. S., Han, S., Banko, M. R., Gozani, O. & Brunet, A. 2010. Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in C. elegans. *Nature,* 466(7304), pp.383-387.

Gregoretti, I. V., Lee, Y. M. & Goodson, H. V. 2004. Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *Journal of Molecular Biology,* 338(1), pp.17-31.

Grewal, S. I. S. & Jia, S. T. 2007. Heterochromatin revisited. *Nature Reviews Genetics,* 8(1), pp.35-46.

Grzenda, A., Lomberk, G., Zhang, J. S. & Urrutia, R. 2009. Sin3: master scaffold and transcriptional corepressor. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1789(6-8), pp.443-450.

Guenther, M. G., Jenner, R. G., Chevalier, B., Nakamura, T., Croce, C. M., Canaani, E. & Young, R. A. 2005. Global and Hox-specific roles for the MLL1 methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America,* 102(24), pp.8603-8608.

Guillemette, B., Drogaris, P., Lin, H.-H. S., Armstrong, H., Hiragami-Hamada, K., Imhof, A., Bonneil, É., Thibault, P., Verreault, A. & Festenstein, R. J. 2011. H3 Lysine 4 Is Acetylated at Active Gene Promoters and Is Regulated by H3 Lysine 4 Methylation. *PLOS Genetics,* 7(3), pp.e1001354.

Guss, K. A., Nelson, C. E., Hudson, A., Kraus, M. E. & Carroll, S. B. 2001. Control of a genetic regulatory network by a selector gene. *Science,* 292(5519), pp.1164-1167.

Haberland, M., Montgomery, R. L. & Olson, E. N. 2009. The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nature Reviews Genetics,* 10(1), pp.32-42.

Haberle, V. & Stark, A. 2018. Eukaryotic core promoters and the functional basis of transcription initiation. *Nature Reviews Molecular Cell Biology,* 19(10), pp.621-637.

Hallson, G., Hollebakken, R. E., Li, T., Syrzycka, M., Kim, I., Cotsworth, S., Fitzpatrick, K. A., Sinclair, D. A. & Honda, B. M. 2012. dSet1 is the main H3K4 di- and tri-methyltransferase throughout Drosophila development. *Genetics,* 190(1), pp.91-100.

Han, J., Zhou, H., Li, Z., Xu, R. M. & Zhang, Z. 2007a. Acetylation of lysine 56 of histone H3 catalyzed by RTT109 and regulated by ASF1 is required for replisome integrity. *Journal of Biological Chemistry,* 282(39), pp.28587-28596.

Han, J., Zhou, H., Li, Z., Xu, R. M. & Zhang, Z. 2007b. The Rtt109-Vps75 histone acetyltransferase complex acetylates non-nucleosomal histone H3. *Journal of Biological Chemistry,* 282(19), pp.14158-14164.

Han, M., Aroian, R. V. & Sternberg, P. W. 1990. The let-60 locus controls the switch between vulval and nonvulval cell fates in Caenorhabditis elegans. *Genetics,* 126(4), pp.899-913.

Han, M., Golden, A., Han, Y. & Sternberg, P. W. 1993. C. elegans lin-45 raf gene participates in let-60 ras-stimulated vulval differentiation. *Nature,* 363(6425), pp.133-140.

Hanna, C. W., Taudt, A., Huang, J., Gahurova, L., Kranz, A., Andrews, S., Dean, W., Stewart, A. F., Colome-Tatche, M. & Kelsey, G. 2018. MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nature Structural & Molecular Biology,* 25(1), pp.73-82.

Harrison, M. M., Ceol, C. J., Lu, X. & Horvitz, H. R. 2006. Some C. elegans class B synthetic multivulva proteins encode a conserved LIN-35 Rb-containing complex distinct from a NuRD-like complex. *Proceedings of the National Academy of Sciences of the United States of America,* 103(45), pp.16782-16787.

Hayakawa, T. & Nakayama, J. 2011. Physiological roles of class I HDAC complex and histone demethylase. *Journal of Biomedicine and Biotechnology,* 2011, pp.129383.

He, C., Li, F., Zhang, J., Wu, J. & Shi, Y. 2013. The methyltransferase NSD3 has chromatin-binding motifs, PHD5-C5HCH, that are distinct from other NSD (nuclear receptor SET domain) family members in their histone H3 recognition. *Journal of biological chemistry,* 288(7), pp.4692-4703.

He, X. 2004. Wnt signaling went derailed again: a new track via the LIN-18 receptor? *Cell,* 118(6), pp.668-670.

Henderson, S. T., Gao, D., Lambie, E. J. & Kimble, J. 1994. Lag-2 May Encode a Signaling Ligand for the Glp-1 and Lin-12 Receptors of C. elegans. *Development,* 120(10), pp.2913-2924.

Hentze, N., Le Breton, L., Wiesner, J., Kempf, G. & Mayer, M. P. 2016. Molecular mechanism of thermosensory function of human heat shock transcription factor Hsf1. *Elife,* 5, pp.e11576.

Herbette, M., Mercier, M. G., Michal, F., Cluet, D., Burny, C., Yvert, G., Robert, V. J. & Palladino, F. 2017. The C. elegans SET-2/SET1 histone H3 Lys4 (H3K4) methyltransferase preserves genome stability in the germline. *DNA Repair,* 57, pp.139-150.

Herz, H. M., Mohan, M., Garruss, A. S., Liang, K., Takahashi, Y. H., Mickey, K., Voets, O., Verrijzer, C. P. & Shilatifard, A. 2012. Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophilahomolog of mammalian Mll3/Mll4. *Genes & Development,* 26(23), pp.2604-2620.

Hill, R. J. & Sternberg, P. W. 1992. The gene lin-3 encodes an inductive signal for vulval development in C. elegans. *Nature,* 358(6386), pp.470-476.

Hill, R. J. & Sternberg, P. W. 1993. Cell fate patterning during C. elegans vulval development. *Development*, pp.9-18.

Hillers, K. J. & Villeneuve, A. M. 2009. Analysis of meiotic recombination in Caenorhabditis elegans. *Methods in Molecular Biology,* 557, pp.77-97.

Ho, L. & Crabtree, G. R. 2010. Chromatin remodelling during development. *Nature,* 463(7280), pp.474-484.

Hoffman, S., Martin, D., Melendez, A. & Bargonetti, J. 2014. C. elegans CEP-1/p53 and BEC-1 are involved in DNA repair. *PLOS One,* 9(2), pp.e88828.

Hofmann, E. R., Milstein, S., Boulton, S. J., Ye, M. J., Hofmann, J. J., Stergiou, L., Gartner, A., Vidal, M. & Hengartner, M. O. 2002. Caenorhabditis elegans HUS-1 is a DNA damage checkpoint protein required for genome stability and EGL-1-mediated apoptosis. *Current Biology,* 12(22), pp.1908-1918.

Holmquist, G., Gray, M., Porter, T. & Jordan, J. 1982. Characterization of Giemsa dark- and light-band DNA. *Cell,* 31(1), pp.121-129.

Horvitz, H. R. & Sulston, J. E. 1980. Isolation and genetic characterization of cell-lineage mutants of the nematode Caenorhabditis elegans. *Genetics,* 96(2), pp.435-454.

Hoskins, R., Hajnal, A. F., Harp, S. A. & Kim, S. K. 1996. The C. elegans vulval induction gene lin-2 encodes a member of the MAGUK family of cell junction proteins. *Development,* 122(1), pp.97-111.

Howe, F. S., Fischl, H., Murray, S. C. & Mellor, J. 2017. Is H3K4me3 instructive for transcription activation? *Bioessays,* 39(1), pp.1-12.

Hsu, D. W., Chubb, J. R., Muramoto, T., Pears, C. J. & Mahadevan, L. C. 2012. Dynamic acetylation of lysine-4-trimethylated histone H3 and H3 variant biology in a simple multicellular eukaryote. *Nucleic Acids Research,* 40(15), pp.7247-7256.

Hsu, P. L., Li, H., Lau, H. T., Leonen, C., Dhall, A., Ong, S. E., Chatterjee, C. & Zheng, N. 2018. Crystal Structure of the COMPASS H3K4 Methyltransferase Catalytic Module. *Cell,* 174(5), pp.1106-1116.

Huang, L. S., Tzou, P. & Sternberg, P. W. 1994. The lin-15 locus encodes two negative regulators of Caenorhabditis elegans vulval development. *Molecular Biology of the Cell,* 5(4), pp.395-411.

Huang, Y., Fang, J., Bedford, M. T., Zhang, Y. & Xu, R. M. 2006. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science,* 312(5774), pp.748-751.

Hughes, C. M., Rozenblatt-Rosen, O., Milne, T. A., Copeland, T. D., Levine, S. S., Lee, J. C., Hayes, D. N., Shanmugam, K. S., Bhattacharjee, A. & Biondi, C. A. 2004. Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. *Molecular Cell,* 13(4), pp.587-597.

Hung, T., Binda, O., Champagne, K. S., Kuo, A. J., Johnson, K., Chang, H. Y., Simon, M. D., Kutateladze, T. G. & Gozani, O. 2009. ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Molecular Cell,* 33(2), pp.248-256.

Hwang, W. W., Venkatasubrahmanyam, S., Ianculescu, A. G., Tong, A., Boone, C. & Madhani, H. D. 2003. A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Molecular Cell,* 11(1), pp.261-266.

Ingham, P. & Whittle, R. 1980. Trithorax: A new homoeotic mutation of Drosophila melanogaster causing transformations of abdominal and thoracic imaginal segments. *Molecular and General Genetics,* 179(3), pp.607-614.

Ingham, P. W. 1981. Trithorax: A new homoeotic mutation ofDrosophila melanogaster : II. The role oftrx (+) after embryogenesis. *Wilehm Roux Archives of Devlopmental Biology,* 190(6), pp.365-369.

Innis, J. W. 1997. Role of HOX genes in human development. *Current Opinion in Pediatrics,* 9(6), pp.617-622.

Inoue, T., Oz, H. S., Wiland, D., Gharib, S., Deshpande, R., Hill, R. J., Katz, W. S. & Sternberg, P. W. 2004. C. elegans LIN-18 is a Ryk ortholog and functions in parallel to LIN-17/Frizzled in Wnt signaling. *Cell,* 118(6), pp.795-806.

Irvine, K., Stirling, R., Hume, D. & Kennedy, D. 2004. Rasputin, more promiscuous than ever: a review of G3BP. *International Journal of Developmental Biology,* 48(10), pp.1065-1077.

Ito, S., Greiss, S., Gartner, A. & Derry, W. B. 2010. Cell-nonautonomous regulation of C. elegans germ cell death by kri-1. *Current Biology,* 20(4), pp.333-338.

Jamaladdin, S., Kelly, R. D., O'regan, L., Dovey, O. M., Hodson, G. E., Millard, C. J., Portolano, N., Fry, A. M., Schwabe, J. W. & Cowley, S. M. 2014. Histone deacetylase (HDAC) 1 and 2 are essential for accurate cell division and the pluripotency of embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America,* 111(27), pp.9840-9845.

Jaramillo-Lambert, A., Harigaya, Y., Vitt, J., Villeneuve, A. & Engebrecht, J. 2010. Meiotic errors activate checkpoints that improve gamete quality without triggering apoptosis in male germ cells. *Current Biology,* 20(23), pp.2078-2089.

John, S., Sabo, P. J., Johnson, T. A., Sung, M. H., Biddie, S. C., Lightman, S. L., Voss, T. C., Davis, S. R., Meltzer, P. S., Stamatoyannopoulos, J. A. & Hager, G. L. 2008. Interaction of the glucocorticoid receptor with the chromatin landscape. *Molecular Cell,* 29(5), pp.611-624.

John, S., Sabo, P. J., Thurman, R. E., Sung, M. H., Biddie, S. C., Johnson, T. A., Hager, G. L. & Stamatoyannopoulos, J. A. 2011. Chromatin accessibility pre-

determines glucocorticoid receptor binding patterns. *Nature Genetics,* 43(3), pp.264-268.

Jones, D., Russnak, R. H., Kay, R. J. & Candido, E. P. 1986. Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis elegans that is flanked by repetitive elements. *Journal of Biological Chemistry,* 261(26), pp.12006-12015.

Joshi, A. A. & Struhl, K. 2005. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Molecular Cell,* 20(6), pp.971-978.

Kaech, S. M., Whitfield, C. W. & Kim, S. K. 1998. The LIN-2/LIN-7/LIN-10 complex mediates basolateral membrane localization of the C. elegans EGF receptor LET-23 in vulval epithelial cells. *Cell,* 94(6), pp.761-771.

Kalkhoven, E. 2004. CBP and p300: HATs for different occasions. *Biochemical Pharmacology,* 68(6), pp.1145-1155.

Katz, W. S., Hill, R. J., Clandinin, T. R. & Sternberg, P. W. 1995. Different levels of the C. elegans growth factor LIN-3 promote distinct vulval precursor fates. *Cell,* 82(2), pp.297-307.

Katz, W. S., Lesa, G. M., Yannoukakos, D., Clandinin, T. R., Schlessinger, J. & Sternberg, P. W. 1996. A point mutation in the extracellular domain activates LET-23, the Caenorhabditis elegans epidermal growth factor receptor homolog. *Molecular and Cellular Biology,* 16(2), pp.529-537.

Kawasaki, H., Schiltz, L., Chiu, R., Itakura, K., Taira, K., Nakatani, Y. & Yokoyama, K. K. 2000. ATF-2 has intrinsic histone acetyltransferase activity which is modulated by phosphorylation. *Nature,* 405(6783), pp.195-200.

Kelly, R. D. & Cowley, S. M. 2013. The physiological roles of histone deacetylase (HDAC) 1 and 2: complex co-stars with multiple leading parts. *Biochemical Society Transactions,* 41(3), pp.741-749.

Kenyon, C. 1986. A gene involved in the development of the posterior body region of C. elegans. *Cell,* 46(3), pp.477-487.

Kim, J., Guermah, M., Mcginty, R. K., Lee, J. S., Tang, Z., Milne, T. A., Shilatifard, A., Muir, T. W. & Roeder, R. G. 2009. RAD6-Mediated transcription-coupled

H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell,* 137(3), pp.459-471.

Kim, T. & Buratowski, S. 2009. Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. *Cell,* 137(2), pp.259-272.

Kimble, J. 1981. Alterations in cell lineage following laser ablation of cells in the somatic gonad of Caenorhabditis elegans. *Developmental Biology,* 87(2), pp.286-300.

Kimble, J. & Hirsh, D. 1979. The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. *Developmental Biology,* 70(2), pp.396-417.

Kleff, S., Andrulis, E. D., Anderson, C. W. & Sternglanz, R. 1995. Identification of a gene encoding a yeast histone H4 acetyltransferase. *Journal of Biological Chemistry,* 270(42), pp.24674-24677.

Klymenko, T. & Müller, J. 2004. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Reports,* 5(4), pp.373-377.

Knudsen, E. S. & Wang, J. Y. J. 2010. Targeting the RB-pathway in Cancer Therapy. *Clinical Cancer Research,* 16(4), pp.1094-1099.

Koga, M. & Ohshima, Y. 1995. Mosaic analysis of the let-23 gene function in vulval induction of Caenorhabditis elegans. *Development,* 121(8), pp.2655-2666.

Korenjak, M., Taylor-Harding, B., Binne, U. K., Satterlee, J. S., Stevaux, O., Aasland, R., White-Cooper, H., Dyson, N. & Brehm, A. 2004. Native E2F/RBF complexes contain Myb-interacting proteins and repress transcription of developmentally controlled E2F target genes. *Cell,* 119(2), pp.181-193.

Kornberg, R. D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science,* 184(4139), pp.868-871.

Kornberg, R. D. & Lorch, Y. 1999. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell,* 98(3), pp.285-294.

Kornfeld, K., Guan, K. L. & Horvitz, H. R. 1995. The Caenorhabditis elegans gene mek-2 is required for vulval induction and encodes a protein similar to the protein kinase MEK. *Genes & Development,* 9(6), pp.756-768.

Kouzarides, T. 2007. Chromatin modifications and their function. *Cell,* 128(4), pp.693-705.

Krogan, N. J., Dover, J., Khorrami, S., Greenblatt, J. F., Schneider, J., Johnston, M. & Shilatifard, A. 2002. COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *Journal of Biological Chemistry,* 277(13), pp.10753-10755.

Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., Greenblatt, J. F. & Shilatifard, A. 2003. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Molecular Cell,* 11(3), pp.721-729.

Kruger, N. J. 1994. The Bradford method for protein quantitation. *Methods in Molecular Biology,* 32, pp.9-15.

Kultz, D. 2007. Osmotic stress sensing and signaling in animals. *The FEBS journal,* 274(22), pp.5781-5781.

Kutscher, L. M. & Shaham, S. 2014. Forward and reverse mutagenesis in C. elegans. *WormBook*, pp.1-26.

Labbadia, J. & Morimoto, R. I. 2015. The biology of proteostasis in aging and disease. *Annual Review of Biochemistry,* 84, pp.435-464.

Lackner, M. R., Kornfeld, K., Miller, L. M., Horvitz, H. R. & Kim, S. K. 1994. A Map Kinase Homolog, Mpk-1, Is Involved in Ras-Mediated Induction of Vulvar Cell Fates in Caenorhabditis elegans. *Genes & Development,* 8(2), pp.160-173.

Lagger, G., O'carroll, D., Rembold, M., Khier, H., Tischler, J., Weitzer, G., Schuettengruber, B., Hauser, C., Brunmeir, R., Jenuwein, T. & Seiser, C. 2002. Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. *EMBO Journal,* 21(11), pp.2672-2681.

Lai, E. C. 2004. Notch signaling: control of cell communication and cell fate. *Development,* 131(5), pp.965-973.

Lamitina, S. T., Morrison, R., Moeckel, G. W. & Strange, K. 2004. Adaptation of the nematode Caenorhabditis elegans to extreme osmotic stress. *American Journal of Physiology-Cell Physiology,* 286(4), pp.C785-791.

Lamitina, T., Huang, C. G. & Strange, K. 2006. Genome-wide RNAi screening identifies protein damage as a regulator of osmoprotective gene expression. *Proceedings of the National Academy of Sciences of the United States of America,* 103(32), pp.12173-12178.

Lau, A. C., Zhu, K. P., Brouhard, E. A., Davis, M. B. & Csankovszki, G. 2016. An H4K16 histone acetyltransferase mediates decondensation of the X chromosome in C. elegans males. *Epigenetics & Chromatin,* 19(9), pp.44-66.

Lauberth, S. M., Nakayama, T., Wu, X., Ferris, A. L., Tang, Z., Hughes, S. H. & Roeder, R. G. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell,* 152(5), pp.1021-1036.

Lavender, C. A., Cannady, K. R., Hoffman, J. A., Trotter, K. W., Gilchrist, D. A., Bennett, B. D., Burkholder, A. B., Burd, C. J., Fargo, D. C. & Archer, T. K. 2016. Downstream Antisense Transcription Predicts Genomic Features That Define the Specific Chromatin Environment at Mammalian Promoters. *PLOS Genetics,* 12(8), pp.e1006224.

Lee, H. H. & Frasch, M. 2005. Nuclear integration of positive Dpp signals, antagonistic Wg inputs and mesodermal competence factors during Drosophila visceral mesoderm induction. *Development,* 132(6), pp.1429-1442.

Lee, J., Kim, D. H., Lee, S., Yang, Q. H., Lee, D. K., Lee, S. K., Roeder, R. G. & Lee, J. W. 2009a. A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. *Proceedings of the National Academy of Sciences of the United States of America,* 106(21), pp.8513-8518.

Lee, J. H. & Skalnik, D. G. 2002. CpG-binding protein is a nuclear matrix- and euchromatin-associated protein localized to nuclear speckles containing human

trithorax. Identification of nuclear matrix targeting signals. *Journal of Biological Chemistry,* 277(44), pp.42259-42267.

Lee, J. H. & Skalnik, D. G. 2005. CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. *Journal of Biological Chemistry,* 280(50), pp.41725-41731.

Lee, J. H., Tate, C. M., You, J. S. & Skalnik, D. G. 2007a. Identification and characterization of the human Set1B histone H3-Lys4methyltransferase complex. *Journal of Biological Chemistry,* 282(18), pp. 13419-13428.

Lee, J. H., Voo, K. S. & Skalnik, D. G. 2001. Identification and characterization of the DNA binding domain of CpG-binding protein. *Journal of Biological Chemistry,* 276(48), pp.44669-44676.

Lee, J. S., Shukla, A., Schneider, J., Swanson, S. K., Washburn, M. P., Florens, L., Bhaumik, S. R. & Shilatifard, A. 2007b. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell,* 131(6), pp.1084-1096.

Lee, K. K. & Workman, J. L. 2007. Histone acetyltransferase complexes: one size doesn't fit all. *Nature Reviews Molecular Cell Biology,* 8(4), pp.284-295.

Lee, S., Kim, D. H., Goo, Y. H., Lee, Y. C., Lee, S. K. & Lee, J. W. 2009b. Crucial roles for interactions between MLL3/4 and INI1 in nuclear receptortransactivation. *Molecular Endocrinology,* 23(5), pp.610-619.

Lee, S., Lee, D. K., Dou, Y., Lee, J., Lee, B., Kwak, E., Kong, Y. Y., Lee, S. K., Roeder, R. G. & Lee, J. W. 2006. Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. *Proceedings of the National Academy of Sciences of the United States of America,* 103(42), pp.15392-15397.

Lee, T. I. & Young, R. A. 2013. Transcriptional regulation and its misregulation in disease. *Cell,* 152(6), pp.1237-1251.

Lefevre, G. M., Patel, S. R., Kim, D., Tessarollo, L. & Dressler, G. R. 2010. Altering a Histone H3K4 Methylation Pathway in Glomerular Podocytes Promotes a Chronic Disease Phenotype. *PLOS Genetics,* 6(10), pp.e1001142.

Lehner, B., Crombie, C., Tischler, J., Fortunato, A. & Fraser, A. G. 2006. Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. *Nature Genetics,* 38(8), pp.896-903.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. & Peters, A. H. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Current Biology,* 13(14), pp.1192-1200.

Lesch, B. J. & Page, D. C. 2014. Poised chromatin in the mammalian germ line. *Development,* 141(19), pp.3619-3626.

Levine, M. 2011. Paused RNA polymerase II as a developmental checkpoint. *Cell,* 145(4), pp.502-511.

Lewis, P. W., Beall, E. L., Fleischer, T. C., Georlette, D., Link, A. J. & Botchan, M. R. 2004. Identification of a Drosophila Myb-E2F2/RBF transcriptional repressor complex. *Genes & Development,* 18(23), pp.2929-2940.

Li, E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Reviews Genetics,* 3(9), pp.662-673.

Li, H., Liu, F., Ren, C., Bo, X. & Shu, W. 2016a. Genome-wide identification and characterisation of HOT regions in the human genome. *BMC Genomics,* 17(1), pp.1-16.

Li, H. T., Ilin, S., Wang, W. K., Duncan, E. M., Wysocka, J., Allis, C. D. & Patel, D. J. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature,* 442(7098), pp.91-95.

Li, J., Chauve, L., Phelps, G., Brielmann, R. M. & Morimoto, R. I. 2016b. E2F coregulates an essential HSF developmental program that is distinct from the heat-shock response. *Genes & Development,* 30(18), pp.2062-2075.

Li, T. & Kelly, W. G. 2011. A role for Set1/MLL-related components in epigenetic regulation of the Caenorhabditis elegans germ line. *PLOS Genetics,* 7(3), pp.e1001349.

Li, T. G. & Kelly, W. G. 2014. A role for WDR5 in TRA-1/Gli mediated transcriptional control of the sperm/oocyte switch in C elegans. *Nucleic Acids Research,* 42(9), pp.5567-5581.

Li, X. Y., Thomas, S., Sabo, P. J., Eisen, M. B., Stamatoyannopoulos, J. A. & Biggin, M. D. 2011. The role of chromatin accessibility in directing the widespread, overlapping patterns of Drosophila transcription factor binding. *Genome Biology,* 12(4), pp.R34.

Lints, R. a. H., D.H. 2009. Reproductive system, germ line. *In WormAtlas*.

Lipsick, J. S. 2004. synMuv verite - Myb comes into focus. *Genes & Development,* 18(23), pp.2837-2844.

Liu, T., Liu, P. Y. & Marshall, G. M. 2009. The critical role of the class III histone deacetylase SIRT1 in cancer. *Cancer Research,* 69(5), pp.1702-1705.

Liu, X., Lee, C. K., Granek, J. A., Clarke, N. D. & Lieb, J. D. 2006. Whole-genome comparison of Leu3 binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. *Genome Research,* 16(12), pp.1517-1528.

Lorenz, D. R., Meyer, L. F., Grady, P. J., Meyer, M. M. & Cam, H. P. 2014. Heterochromatin assembly and transcriptome repression by Set1 in coordination with a class II histone deacetylase. *Elife,* 3, pp.e04506.

Lorenz, D. R., Mikheyeva, I. V., Johansen, P., Meyer, L., Berg, A., Grewal, S. I. & Cam, H. P. 2012. CENP-B cooperates with Set1 in bidirectional transcriptional silencing and genome organization of retrotransposons. *Molecular and Cellular Biology,* 32(20), pp.4215-4225.

Lu, J., Mckinsey, T. A., Nicol, R. L. & Olson, E. N. 2000. Signal-dependent activation of the MEF2 transcription factor by dissociation from histone deacetylases. *Proceedings of the National Academy of Sciences of the United States of America,* 97(8), pp.4070-4075.

Lu, X. W. & Horvitz, H. R. 1998. lin-35 and lin-53, two genes that antagonize a C. elegans Ras pathway, encode proteins similar to Rb and its binding protein RbAp48. *Cell,* 95(7), pp.981-991.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature,* 389(6648), pp.251-260.

Luo, Y., Jian, W., Stavreva, D., Fu, X., Hager, G., Bungert, J., Huang, S. & Qiu, Y. 2009. Trans-regulation of histone deacetylase activities through acetylation. *Journal of Biological Chemistry,* 284(50), pp.34901-34910.

Luo, Z. J., Lin, C. Q. & Shilatifard, A. 2012. The super elongation complex (SEC) family in transcriptional control. *Nature Reviews Molecular Cell Biology,* 13(9), pp.543-547.

Lussi, Y. C., Mariani, L., Friis, C., Peltonen, J., Myers, T. R., Krag, C., Wong, G. & Salcini, A. E. 2016. Impaired removal of H3K4 methylation affects cell fate determination and gene transcription. *Development,* 143(20), pp.3751-3762.

Mahadevan, J. & Skalnik, D. G. 2016. Efficient differentiation of murine embryonic stem cells requires the binding of CXXC finger protein 1 to DNA or methylated histone H3-Lys4. *Gene,* 594(1), pp.1-9.

Majdzadeh, N., Wang, L., Morrison, B. E., Bassel-Duby, R., Olson, E. N. & D'mello, S. R. 2008. HDAC4 inhibits cell-cycle progression and protects neurons from cell death. *Developmental Neurobiology,* 68(8), pp.1076-1092.

Malek, R., Gajula, R. P., Williams, R. D., Nghiem, B., Simons, B. W., Nugent, K., Wang, H., Taparra, K., Lemtiri-Chlieh, G., Yoon, A. R., True, L., An, S. S., Deweese, T. L., Ross, A. E., Schaeffer, E. M., Pienta, K. J., Hurley, P. J., Morrissey, C. & Tran, P. T. 2017. TWIST1-WDR5-Hottip Regulates Hoxa9 Chromatin to Facilitate Prostate Cancer Metastasis. *Cancer Research,* 77(12), pp.3181-3193.

Margaritis, T., Oreal, V., Brabers, N., Maestroni, L., Vitaliano-Prunier, A., Benschop, J. J., Hooff, S., Leenen, D., Dargemont, C., Geli, V. & Holstege, F. C. 2012. Two distinct repressive mechanisms for histone 3 lysine 4 methylation

through promoting 3'-end antisense transcription. *PLOS Genetics,* 8(9), pp.e1002952.

Marmorstein, R. & Zhou, M. M. 2014. Writers and readers of histone acetylation: structure, mechanism, and inhibition. *Cold Spring Harbor Perspectives in Biology,* 6(7), pp.1-25.

Mcmullen, P. D., Aprison, E. Z., Winter, P. B., Amaral, L. A., Morimoto, R. I. & Ruvinsky, I. 2012. Macro-level modeling of the response of C. elegans reproduction to chronic heat stress. *PLOS Computational Biology,* 8(1), pp.e1002338.

Mcmurchy, A. N., Stempor, P., Gaarenstroom, T., Wysolmerski, B., Dong, Y., Aussianikava, D., Appert, A., Huang, N., Kolasinska-Zwierz, P., Sapetschnig, A., Miska, E. A. & Ahringer, J. 2017. A team of heterochromatin factors collaborates with small RNA pathways to combat repetitive elements and germline stress. *Elife,* 6, pp.e21666.

Melendez, A. & Greenwald, I. 2000. Caenorhabditis elegans lin-13, a member of the LIN-35 Rb class of genes involved in vulval development, encodes a protein with zinc fingers and an LXCXE motif. *Genetics,* 155(3), pp.1127-1137.

Mello, C. C., Draper, B. W. & Priess, J. R. 1994. The Maternal Genes Apx-1 and Glp-1 and Establishment of Dorsal-Ventral Polarity in the Early C. elegans Embryo. *Cell,* 77(1), pp.95-106.

Mikheyeva, I. V., Grady, P. J., Tamburini, F. B., Lorenz, D. R. & Cam, H. P. 2014. Multifaceted genome control by Set1 Dependent and Independent of H3K4 methylation and the Set1C/COMPASS complex. *PLOS Genetics,* 10(10), pp.e1004740.

Miller, Gallegos, M. E., Morisseau, B. A. & Kim, S. K. 1993. lin-31, a Caenorhabditis elegans HNF-3/fork head transcription factor homolog, specifies three alternative cell fates in vulval development. *Genes & Development,* 7(6), pp.933-947.

Miller, Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F. & Shilatifard, A. 2001. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proceedings of the*

National Academy of Sciences of the United States of America, 98(23), pp.12902-12907.

Min, L., Ruan, Y., Shen, Z., Jia, D., Wang, X., Zhao, J., Sun, Y. & Gu, J. 2015. Overexpression of Ras-GTPase-activating protein SH3 domain-binding protein 1 correlates with poor prognosis in gastric cancer patients. *Histopathology,* 67(5), pp.677-688.

Mizzen, C. A., Yang, X.-J., Kokubo, T., Brownell, J. E., Bannister, A. J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S. L., Kouzarides, T., Nakatani, Y. & Allis, C. D. 1996. The TAFII250 Subunit of TFIID Has Histone Acetyltransferase Activity. *Cell,* 87(7), pp.1261-1270.

Mo, R., Rao, S. M. & Zhu, Y. J. 2006. Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. *Journal of Biological Chemistry,* 281(23), pp.15714-15720.

Mohan, M., Herz, H. M., Smith, E. R., Zhang, Y., Jackson, J., Washburn, M. P., Florens, L., Eissenberg, J. C. & Shilatifard, A. 2011. The COMPASS family of H3K4 methylases in Drosophila. *Molecular and Cellular Biology,* 31(21), pp.4310-4318.

Mohan, M., Lin, C., Guest, E. & Shilatifard, A. 2010. Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis. *Nature Reviews Cancer,* 10(10), pp.721-728.

Mohd-Sarip, A., Teeuwssen, M., Bot, A. G., De Herdt, M. J., Willems, S. M., Baatenburg De Jong, R. J., Looijenga, L. H. J., Zatreanu, D., Bezstarosti, K., Van Riet, J., Oole, E., Van Ijcken, W. F. J., Van De Werken, H. J. G., Demmers, J. A., Fodde, R. & Verrijzer, C. P. 2017. DOC1-Dependent Recruitment of NURD Reveals Antagonism with SWI/SNF during Epithelial-Mesenchymal Transition in Oral Cancer Cells. *Cell reports,* 20(1), pp.61-75.

Moreira, J. M. A. & Holmberg, S. 1999. Transcriptional repression of the yeast CHA1 gene requires the chromatin-remodeling complex RSC. *EMBO Journal,* 18(10), pp.2836-2844.

Morishita, M., Mevius, D. & Di Luccio, E. 2014. In vitro histone lysine methylation by NSD1, NSD2/MMSET/WHSC1 and NSD3/WHSC1L. *BMC Structural Biology,* 14(25), pp.1-13.

Murphy, D. J., Hardy, S. & Engel, D. A. 1999. Human SWI-SNF component BRG1 represses transcription of the c-fos gene. *Molecular and Cellular Biology,* 19(4), pp.2724-2733.

Niwa, H. 2018. The principles that govern transcription factor network functions in stem cells. *Development,* 145(6), pp.dev157420.

Orphanides, G. & Reinberg, D. 2002. A unified theory of gene expression. *Cell,* 108(4), pp.439-451.

Ou, H. D., Phan, S., Deerinck, T. J., Thor, A., Ellisman, M. H. & O'shea, C. C. 2017. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science,* 357(6349), pp.1-13.

Paggetti, J., Largeot, A., Aucagne, R., Jacquel, A., Lagrange, B., Yang, X. J., Solary, E., Bastie, J. N. & Delva, L. 2010. Crosstalk between leukemia-associated proteins MOZ and MLL regulates HOX gene expression in human cord blood CD34+ cells. *Oncogene,* 29(36), pp.5019-5031.

Palacios, A., Moreno, A., Oliveira, B. L., Rivera, T., Prieto, J., Garcia, P., Fernandez-Fernandez, M. R., Bernado, P., Palmero, I. & Blanco, F. J. 2010. The dimeric structure and the bivalent recognition of H3K4me3 by the tumor suppressor ING4 suggests a mechanism for enhanced targeting of the HBO1 complex to chromatin. *Journal of Molecular Biology,* 396(4), pp.1117-1127.

Pan, X., Lei, B., Zhou, N., Feng, B., Yao, W., Zhao, X., Yu, Y. & Lu, H. 2012. Identification of novel genes involved in DNA damage response by screening a genome-wide Schizosaccharomyces pombe deletion library. *BMC Genomics,* 13(1), pp.662-677.

Papatsenko, D. & Levine, M. 2005. Quantitative analysis of binding motifs mediating diverse spatial readouts of the Dorsal gradient in the Drosophila embryo. *Proceedings of the National Academy of Sciences of the United States of America,* 102(14), pp.4966-4971.

Parker, F., Maurier, F., Delumeau, I., Duchesne, M., Faucher, D., Debussche, L., Dugue, A., Schweighoffer, F. & Tocque, B. 1996. A ras-GTPase-activating protein SH3-domain-binding protein. *Molecular and Cellular Biology,* 16(6), pp.2561-2569.

Pasini, D., Malatesta, M., Jung, H. R., Walfridsson, J., Willer, A., Olsson, L., Skotte, J., Wutz, A., Porse, B., Jensen, O. N. & Helin, K. 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), pp.4958-4969.

Pedersen, M. T., Agger, K., Laugesen, A., Johansen, J. V., Cloos, P. A., Christensen, J. & Helin, K. 2014. The demethylase JMJD2C localizes to H3K4me3-positive transcription start sites and is dispensable for embryonic development. *Molecular and Cellular Biology,* 34(6), pp.1031-1045.

Pena, P. V., Hom, R. A., Hung, T., Lin, H., Kuo, A. J., Wong, R. P., Subach, O. M., Champagne, K. S., Zhao, R., Verkhusha, V. V., Li, G., Gozani, O. & Kutateladze, T. G. 2008. Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor. *Journal of Molecular Biology,* 380(2), pp.303-312.

Perissi, V., Jepsen, K., Glass, C. K. & Rosenfeld, M. G. 2010. Deconstructing repression: evolving models of co-repressor action. *Nature Reviews Genetics,* 11(2), pp.109-123.

Peserico, A. & Simone, C. 2011. Physical and functional HAT/HDAC interplay regulates protein acetylation balance. *Journal of Biomedicine and Biotechnology,* 2011, pp.371832-371842.

Pferdehirt, R. R., Kruesi, W. S. & Meyer, B. J. 2011. An MLL/COMPASS subunit functions in the C. elegans dosage compensation complex to target X chromosomes for transcriptional regulation of gene expression. *Genes & Development,* 25(5), pp.499-515.

Philip Ingham, R. W. 1980. Trithorax: A new homoeotic mutation of Drosophila melanogaster causing transformations of abdominal and thoracic imaginal segments. *Molecular and General Genetics,* 179(3), pp.607-614.

Pirrotta, V. 1998. Polycombing the genome: PcG, trxG, and chromatin silencing. *Cell,* 93(3), pp.333-336.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K. & Young, R. A. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell,* 122(4), pp.517-527.

Pokhrel, B., Chen, Y. & Biro, J. J. 2019. CFP-1 interacts with HDAC1/2 complexes in C. elegans development. *The FEBS journal,* 286(13), pp.2490-2504.

Portela, A. & Esteller, M. 2010. Epigenetic modifications and human disease. *Nature Biotechnology,* 28(10), pp.1057-1068.

Potthoff, M. J. & Olson, E. N. 2007. MEF2: a central regulator of diverse developmental programs. *Development,* 134(23), pp.4131-4140.

Poulin, G., Dong, Y., Fraser, A. G., Hopper, N. A. & Ahringer, J. 2005. Chromatin regulation and sumoylation in the inhibition of Ras-induced vulval development in Caenorhabditis elegans. *EMBO Journal,* 24(14), pp.2613-2623.

Prahlad, V., Cornelius, T. & Morimoto, R. I. 2008. Regulation of the cellular heat shock response in Caenorhabditis elegans by thermosensory neurons. *Science,* 320(5877), pp.811-814.

Qu, Q., Takahashi, Y. H., Yang, Y., Hu, H., Zhang, Y., Brunzelle, J. S., Couture, J. F., Shilatifard, A. & Skiniotis, G. 2018. Structure and Conformational Dynamics of a COMPASS Histone H3K4 Methyltransferase Complex. *Cell,* 174(5), pp.1117-1126

Reddy, K. C. & Villeneuve, A. M. 2004. C. elegans HIM-17 links chromatin modification and competence for initiation of meiotic recombination. *Cell,* 118(4), pp.439-452.

Reinke, V., Krause, M. & Okkema, P. 2013. Transcriptional regulation of gene expression in C. elegans. *WormBook*, pp.1-34.

Ritossa 1962. A new puffing pattern induced by temperature shock and DNP in drosophila. *Experientia,* 18(12), pp.3.

Robert, F., Pokholok, D. K., Hannett, N. M., Rinaldi, N. J., Chandy, M., Rolfe, A., Workman, J. L., Gifford, D. K. & Young, R. A. 2004. Global position and recruitment of HATs and HDACs in the yeast genome. *Molecular Cell,* 16(2), pp.199-209.

Robert, V. J., Mercier, M. G., Bedet, C., Janczarski, S., Merlet, J., Garvis, S., Ciosk, R. & Palladino, F. 2014. The SET-2/SET1 histone H3K4 methyltransferase maintains pluripotency in the Caenorhabditis elegans germline. *Cell reports,* 9(2), pp.443-450.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R. & Stewart, A. F. 2001. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue andmethylates histone 3 lysine 4. *EMBO Journal,* 20(24), pp.7137-7148.

Roos, W. P. & Krumm, A. 2016. The multifaceted influence of histone deacetylases on DNA damage signalling and DNA repair. *Nucleic Acids Research,* 44(21), pp.10017-10030.

Ropero, S. & Esteller, M. 2007. The role of histone deacetylases (HDACs) in human cancer. *Molecular Oncology,* 1(1), pp.19-25.

Roth, S. Y., Denu, J. M. & Allis, C. D. 2001. Histone acetyltransferases. *Annual Review of Biochemistry,* 70, pp.81-120.

Sachs, M., Onodera, C., Blaschke, K., Ebata, K. T., Song, J. S. & Ramalho-Santos, M. 2013. Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline in vivo. *Cell reports,* 3(6), pp.1777-1784.

Salghetti, S. E., Caudy, A. A., Chenoweth, J. G. & Tansey, W. P. 2001. Regulation of transcriptional activation domain function by ubiquitin. *Science,* 293(5535), pp.1651-1653.

Salz, T., Li, G., Kaye, F., Zhou, L., Qiu, Y. & Huang, S. 2014. hSETD1A regulates Wnt target genes and controls tumor growth of colorectal cancer cells. *Cancer Research,* 74(3), pp.775-786.

Samarakkody, A., Abbas, A., Scheidegger, A., Warns, J., Nnoli, O., Jokinen, B., Zarns, K., Kubat, B., Dhasarathy, A. & Nechaev, S. 2015. RNA polymerase II

pausing can be retained or acquired during activation of genes involved in the epithelial to mesenchymal transition. *Nucleic Acids Research,* 43(8), pp.3938-3949.

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J. & Kouzarides, T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature,* 419(6905), pp.407-411.

Sapountzi, V. & Cote, J. 2011. MYST-family histone acetyltransferases: beyond chromatin. *Cellular and Molecular Life Sciences,* 68(7), pp.1147-1156.

Schlichter, A. & Cairns, B. R. 2005. Histone trimethylation by Set1 is coordinated by the RRM, autoinhibitory, and catalytic domains. *EMBO Journal,* 24(6), pp.1222-1231.

Schneider, J., Dover, J., Johnston, M. & Shilatifard, A. 2004. Global proteomic analysis of S. cerevisiae (GPS) to identify proteins required for histone modifications. *Methods in Enzymology,* 377, pp.227-234.

Schneider, J., Wood, A., Lee, J. S., Schuster, R., Dueker, J., Maguire, C., Swanson, S. K., Florens, L., Washburn, M. P. & Shilatifard, A. 2005. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Molecular Cell,* 19(6), pp.849-856.

Schock, F., Purnell, B. A., Wimmer, E. A. & Jackle, H. 1999. Common and diverged functions of the Drosophila gene pair D-Sp1 and buttonhead. *Mechanisms of Development,* 89(1-2), pp.125-132.

Schumacher, B., Hofmann, K., Boulton, S. & Gartner, A. 2001. The C. elegans homolog of the p53 tumor suppressor is required for DNA damage-induced apoptosis. *Current Biology,* 11(21), pp.1722-1727.

Schwer, B. & Verdin, E. 2008. Conserved metabolic regulatory functions of sirtuins. *Cell Metabolism,* 7(2), pp.104-112.

Sedkov, Y., Benes, J. J., Berger, J. R., Riker, K. M., Tillib, S., Jones, R. S. & Mazo, A. 1999. Molecular genetic analysis of the Drosophila trithorax-related gene which encodes a novel SET domain protein. *Mechanisms of Development,* 82(1-2), pp.171-179.

Sedkov, Y., Cho, E., Petruk, S., Cherbas, L., Smith, S. T., Jones, R. S., Cherbas, P., Canaani, E., Jaynes, J. B. & Mazo, A. 2003. Methylation at lysine 4 of histone H3 in ecdysone-dependent development of Drosophila. *Nature,* 426(6962), pp.78-83.

Seto, E. & Yoshida, M. 2014. Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harbor Perspectives in Biology,* 6(4), pp.1-26.

Sha, Q. Q., Dai, X. X., Jiang, J. C., Yu, C., Jiang, Y., Liu, J. P., Ou, X. H., Zhang, S. Y. & Fan, H. Y. 2018. CFP1 coordinates histone H3 lysine-4 trimethylation and meiotic cell cycle progression in mouse oocytes. *Nature Communications,* 9(1), pp.3477.

Shandilya, J. & Roberts, S. G. E. 2012. The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1819(5), pp.391-400.

Shaye, D. D. & Greenwald, I. 2002. Endocytosis-mediated downregulation of LIN-12/Notch upon Ras activation in Caenorhabditis elegans. *Nature,* 420(6916), pp.686-690.

Shearn, A. 1989. The ash-1, ash-2 and trithorax genes of Drosophila melanogaster are functionally related. *Genetics,* 121(3), pp.517-525.

Shearn, A., Hersperger, E. & Hersperger, G. 1987. Genetic studies of mutations at two loci of Drosophila melanogaster which cause a wide variety of homeotic transformations. *Roux's Archives of Developmental Biology,* 196(4), pp.231-242.

Shi, X., Finkelstein, A., Wolf, A. J., Wade, P. A., Burton, Z. F. & Jaehning, J. A. 1996. Paf1p, an RNA polymerase II-associated factor in Saccharomyces cerevisiae, may have both positive and negative roles in transcription. *Molecular and Cellular Biology,* 16(2), pp.669-676.

Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M. R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B. R., Ayer, D. E., Kutateladze, T. G., Shi, Y., Côté, J., Chua, K. F. & Gozani, O. 2006. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature,* 442(7098), pp.96-99.

Shi, Y., Lee, J. S. & Galvin, K. M. 1997. Everything you have ever wanted to know about Yin Yang 1. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1332(2), pp.F49-66.

Shi, Y. & Mello, C. 1998. A CBP/p300 homolog specifies multiple differentiation pathways in Caenorhabditis elegans. *Genes & Development,* 12(7), pp.943-955.

Shibata, Y., Uchida, M., Takeshita, H., Nishiwaki, K. & Sawa, H. 2012. Multiple functions of PBRM-1/Polybromo- and LET-526/Osa-containing chromatin remodeling complexes in C. elegans development. *Developmental Biology,* 361(2), pp.349-357.

Shilatifard, A. 2008. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Current Opinion in Cell Biology,* 20(3), pp.341-348.

Shilatifard, A. 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual Review of Biochemistry,* 81, pp.65-95.

Shopland, L. S., Hirayoshi, K., Fernandes, M. & Lis, J. T. 1995. HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites. *Genes & Development,* 9(22), pp.2756-2769.

Siebold, A. P., Banerjee, R., Tie, F., Kiss, D. L., Moskowitz, J. & Harte, P. J. 2010. Polycomb Repressive Complex 2 and Trithorax modulate Drosophila longevity and stress resistance. *Proceedings of the National Academy of Sciences of the United States of America,* 107(1), pp.169-174.

Sikorski, T. W. & Buratowski, S. 2009. The basal initiation machinery: beyond the general transcription factors. *Current Opinion in Cell Biology,* 21(3), pp.344-351.

Silverstein, R. A. & Ekwall, K. 2005. Sin3: a flexible regulator of global gene expression and genome stability. *Current Genetics,* 47(1), pp.1-17.

Simonet, T., Dulermo, R., Schott, S. & Palladino, F. 2007. Antagonistic functions of SET-2/SET1 and HPL/HP1 proteins in C. elegans development. *Developmental Biology,* 312(1), pp.367-383.

Sims, R. J., Millhouse, S., Chen, C. F., Lewis, B. A., Erdjument-Bromage, H., Tempst, P., Manley, J. L. & Reinberg, D. 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Molecular Cell,* 28(4), pp.665-676.

Simske, J. S., Kaech, S. M., Harp, S. A. & Kim, S. K. 1996. LET-23 receptor localization by the cell junction protein LIN-7 during C. elegans vulval induction. *Cell,* 85(2), pp.195-204.

Simske, J. S. & Kim, S. K. 1995. Sequential signalling during Caenorhabditis elegans vulval induction. *Nature,* 375(6527), pp.142-146.

Sin, O., Michels, H. & Nollen, E. A. 2014. Genetic screens in Caenorhabditis elegans models for neurodegenerative diseases. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1842(10), pp.1951-1959.

Singh, N. & Han, M. 1995. sur-2, a novel gene, functions late in the let-60 ras-mediated signaling pathway during Caenorhabditis elegans vulval induction. *Genes & Development,* 9(18), pp.2251-2265.

Skalnik, D. G. 2010. The epigenetic regulator Cfp1. *Biomolecular Concepts,* 1(5-6), pp.325-334.

Slaughter, M. J., Shanle, E. K., Mcfadden, A. W., Hollis, E. S., Suttle, L. E., Strahl, B. D. & Davis, I. J. 2018. PBRM1 bromodomains variably influence nucleosome interactions and cellular function. *Journal of Biological Chemistry,* 293(35), pp.13592-13603.

Smith, E., Lin, C. Q. & Shilatifard, A. 2011. The super elongation complex (SEC) and MLL in development and disease. *Genes & Development,* 25(7), pp.661-672.

Snutch, T. P., Heschl, M. F. & Baillie, D. L. 1988. The Caenorhabditis elegans hsp70 gene family: a molecular genetic characterization. *Gene,* 64(2), pp.241-255.

Solari, F. & Ahringer, J. 2000. NURD-complex genes antagonise Ras-induced vulval development in Caenorhabditis elegans. *Current Biology,* 10(4), pp.223-226.

Sommermeyer, V., Beneut, C., Chaplais, E., Serrentino, M. E. & Borde, V. 2013. Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in

promoters by tethering histone H3K4 methylation sites to chromosome axes. *Molecular Cell,* 49(1), pp.43-54.

Song, C. Z., Keller, K., Murata, K., Asano, H. & Stamatoyannopoulos, G. 2002. Functional interaction between coactivators CBP/p300, PCAF, and transcription factor FKLF2. *Journal of Biological Chemistry,* 277(9), pp.7029-7036.

Spencer, T. E., Jenster, G., Burcin, M. M., Allis, C. D., Zhou, J., Mizzen, C. A., Mckenna, N. J., Onate, S. A., Tsai, S. Y., Tsai, M. J. & O'malley, B. W. 1997. Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature,* 389(6647), pp.194-198.

Spensberger, D., Vermeulen, M., Le Guezennec, X., Beekman, R., Van Hoven, A., Bindels, E., Stunnenberg, H. & Delwel, R. 2008. Myeloid transforming protein Evi1 interacts with methyl-CpG binding domain protein 3 and inhibits in vitro histone deacetylation by Mbd3/Mi-2/NuRD. *Biochemistry,* 47(24), pp.6418-6426.

Spitz, F. & Furlong, E. E. 2012. Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics,* 13(9), pp.613-626.

Stassen, M. J., Bailey, D., Nelson, S., Chinwalla, V. & Harte, P. J. 1995. The Drosophila trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mechanisms of Development,* 52(2-3), pp.209-223.

Stathopoulos, A., Van Drenth, M., Erives, A., Markstein, M. & Levine, M. 2002. Whole-genome analysis of dorsal-ventral patterning in the Drosophila embryo. *Cell,* 111(5), pp.687-701.

Sternberg, P. W. 2005. Vulval development. *WormBook*, pp.1-28.

Sternberg, P. W. & Horvitz, H. R. 1986. Pattern formation during vulval development in C. elegans. *Cell,* 44(5), pp.761-772.

Sulston, J. E. & Horvitz, H. R. 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental Biology,* 56(1), pp.110-156.

Sulston, J. E. & White, J. G. 1980. Regulation and cell autonomy during postembryonic development of Caenorhabditis elegans. *Developmental Biology,* 78(2), pp.577-597.

Sundaram, M. V. 2006. RTK/Ras/MAPK signaling. *WormBook*, pp.1-19.

Sze, C. C., Cao, K., Collings, C. K., Marshall, S. A., Rendleman, E. J., Ozark, P. A., Chen, F. X., Morgan, M. A., Wang, L. & Shilatifard, A. 2017. Histone H3K4 methylation-dependent and -independent functions of Set1A/COMPASS in embryonic stem cell self-renewal and differentiation. *Genes & Development,* 31(17), pp.1732-1737.

Tang, Z., Chen, W.-Y., Shimada, M., Nguyen, U. T. T., Kim, J., Sun, X.-J., Sengoku, T., Mcginty, R. K., Fernandez, J. P., Muir, T. W. & Roeder, R. G. 2013. SET1 and p300 Act Synergistically, through Coupled Histone Modifications, in Transcriptional Activation by p53. *Cell,* 154(2), pp.297-310.

Taplick, J., Kurtev, V., Kroboth, K., Posch, M., Lechner, T. & Seiser, C. 2001. Homo-oligomerisation and nuclear localisation of mouse histone deacetylase 1. *Journal of Molecular Biology,* 308(1), pp.27-38.

Tate, C. M., Lee, J. H. & Skalnik, D. G. 2010. CXXC finger protein 1 restricts the Setd1A histone H3K4 methyltransferase complex to euchromatin. *The FEBS journal,* 277(1), pp.210-223.

Taunton, J., Hassig, C. A. & Schreiber, S. L. 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science,* 272(5260), pp.408-411.

Taverna, S. D., Ilin, S., Rogers, R. S., Tanny, J. C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L. P., Chait, B. T., Patel, D. J., Aitchison, J. D., Tackett, A. J. & Allis, C. D. 2006. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Molecular Cell,* 24(5), pp.785-796.

Tax, F. E. & Thomas, J. H. 1994. Cell Cell Interactions - Receiving Signals in the Nematode Embryo. *Current Biology,* 4(10), pp.914-916.

Tessarz, P. & Kouzarides, T. 2014. Histone core modifications regulating nucleosome structure and dynamics. *Nature Reviews Molecular Cell Biology,* 15(11), pp.703-708.

Teves, S. S. & Henikoff, S. 2013. The heat shock response: A case study of chromatin dynamics in gene regulation. *Biochemistry and Cell Biology,* 91(1), pp.42-48.

Thiel, A. T., Feng, Z., Pant, D. K., Chodosh, L. A. & Hua, X. 2013. The trithorax protein partner menin acts in tandem with EZH2 to suppress C/EBPalpha and differentiation in MLL-AF9 leukemia. *Haematologica,* 98(6), pp.918-927.

Thiel, A. T., Huang, J., Lei, M. & Hua, X. 2012. Menin as a hub controlling mixed lineage leukemia. *Bioessays,* 34(9), pp.771-780.

Thomas, J. H., Ceol, C. J., Schwartz, H. T. & Horvitz, H. R. 2003. New genes that interact with lin-35 Rb to negatively regulate the let-60 ras pathway in Caenorhabditis elegans. *Genetics,* 164(1), pp.135-151.

Thomas, J. H., Stern, M. J. & Horvitz, H. R. 1990. Cell interactions coordinate the development of the C. elegans egg-laying system. *Cell,* 62(6), pp.1041-1052.

Thomson, J. P., Skene, P. J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A. R., Deaton, A., Andrews, R., James, K. D., Turner, D. J., Illingworth, R. & Bird, A. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature,* 464(7291), pp.1082-1086.

Tie, F., Banerjee, R., Conrad, P. A., Scacheri, P. C. & Harte, P. J. 2012. Histone Demethylase UTX and Chromatin Remodeler BRM Bind Directly to CBP and Modulate Acetylation of Histone H3 Lysine 27. *Molecular and Cellular Biology,* 32(12), pp.2323-2334.

Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E. & Schreiber, S. L. 1998. Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature,* 395(6705), pp.917-921.

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G. & Reuter, G. 1994. The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO Journal,* 13(16), pp.3822-3831.

Van De Lagemaat, L. N., Flenley, M., Lynch, M. D., Garrick, D., Tomlinson, S. R., Kranc, K. R. & Vernimmen, D. 2018. CpG binding protein (CFP1) occupies open

chromatin regions of active genes, including enhancers and non-CpG islands. *Epigenetics Chromatin,* 11(1), pp.1-18.

Vandamme, J., Lettier, G., Sidoli, S., Di Schiavi, E., Nørregaard Jensen, O. & Salcini, A. E. 2012. The C. elegans H3K27 Demethylase UTX-1 Is Essential for Normal Development, Independent of Its Enzymatic Activity. *PLOS Genetics,* 8(5), pp.e1002647.

Vandamme, J. & Salcini, A. E. 2013. Catalytic-independent roles of UTX-1 in C. elegans development. *Worm,* 2(2), pp.1-5.

Vastenhouw, N. L., Brunschwig, K., Okihara, K. L., Muller, F., Tijsterman, M. & Plasterk, R. H. 2006. Gene expression: long-term gene silencing by RNAi. *Nature,* 442(7105), pp.882-892.

Verdin, E., Dequiedt, F. & Kasler, H. G. 2003. Class II histone deacetylases: versatile regulators. *Trends in Genetics,* 19(5), pp.286-293.

Vicent, G. P., Zaurin, R., Nacht, A. S., Li, A., Font-Mateu, J., Le Dily, F., Vermeulen, M., Mann, M. & Beato, M. 2009. Two chromatin remodeling activities cooperate during activation of hormone responsive promoters. *PLOS Genetics,* 5(7), pp.e1000567.

Vitaliano-Prunier, A., Menant, A., Hobeika, M., Geli, V., Gwizdek, C. & Dargemont, C. 2008. Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. *Nature Cell Biology,* 10(11), pp.1365-1371.

Voo, K. S., Carlone, D. L., Jacobsen, B. M., Flodin, A. & Skalnik, D. G. 2000. Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1. *Molecular and Cellular Biology,* 20(6), pp.2108-2121.

Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B. R., Protacio, A., Flynn, R. A., Gupta, R. A., Wysocka, J., Lei, M., Dekker, J., Helms, J. A. & Chang, H. Y. 2011a. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature,* 472(7341), pp.120–124.

Wang, P., Bowl, M. R., Bender, S., Peng, J., Farber, L., Chen, J., Ali, A., Zhang, Z., Alberts, A. S., Thakker, R. V., Shilatifard, A., Williams, B. O. & Teh, B. T. 2008. Parafibromin, a component of the human PAF complex, regulates growth factors and is required for embryonic development and survival in adult mice. *Molecular and Cellular Biology,* 28(9), pp.2930-2940.

Wang, P., Lin, C., Smith, E. R., Guo, H., Sanderson, B. W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L. M., Ge, K., Krumlauf, R. & Shilatifard, A. 2009. Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Molecular and Cellular Biology,* 29(22), pp.6074-6085.

Wang, S., Fisher, K. & Poulin, G. B. 2011b. Lineage specific trimethylation of H3 on lysine 4 during C. elegans early embryogenesis. *Developmental Biology,* 355(2), pp.227-238.

Wannissorn, N. 2016. *blmp-1 is a Ras-cooperating Tumor Suppressor Gene in Caenorhabditis elegans.* Ph.D. Thesis, University of Toronto.

Weiner, A., Chen, H. V., Liu, C. L., Rahat, A., Klien, A., Soares, L., Gudipati, M., Pfeffner, J., Regev, A., Buratowski, S., Pleiss, J. A., Friedman, N. & Rando, O. J. 2012. Systematic dissection of roles for chromatin regulators in a yeast stress response. *PLOS Biology,* 10(7), pp.e1001369.

Wenzel, D., Palladino, F. & Jedrusik-Bode, M. 2011. Epigenetics in C. elegans: facts and challenges. *Genesis,* 49(8), pp.647-661.

Whitfield, C. W., Benard, C., Barnes, T., Hekimi, S. & Kim, S. K. 1999. Basolateral localization of the Caenorhabditis elegans epidermal growth factor receptor in epithelial cells by the PDZ protein LIN-10. *Molecular Biology of the Cell,* 10(6), pp.2087-2100.

Wilkinson, H. A., Fitzgerald, K. & Greenwald, I. 1994. Reciprocal changes in expression of the receptor lin-12 and its ligand lag-2 prior to commitment in a C. elegans cell fate decision. *Cell,* 79(7), pp.1187-1198.

Wimmer, E. A., Frommer, G., Purnell, B. A. & Jackle, H. 1996. buttonhead and D-Sp1: a novel Drosophila gene pair. *Mechanisms of Development,* 59(1), pp.53-62.

Wolkow, C. a. a. H., D.H 2015. Introduction to the Dauer Larva, Overview. *In WormAtlas*.

Wood, A., Schneider, J., Dover, J., Johnston, M. & Shilatifard, A. 2003. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *Journal of Biological Chemistry,* 278(37), pp.34739-34742.

Wu, M., Wang, P. F., Lee, J. S., Martin-Brown, S., Florens, L., Washburn, M. & Shilatifard, A. 2008. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Molecular and Cellular Biology,* 28(24), pp.7337-7344.

Wu, Y. & Han, M. 1994. Suppression of activated Let-60 ras protein defines a role of Caenorhabditis elegans Sur-1 MAP kinase in vulval differentiation. *Genes & Development,* 8(2), pp.147-159.

Wu, Y., Han, M. & Guan, K. L. 1995. MEK-2, a Caenorhabditis elegans MAP kinase kinase, functions in Ras-mediated vulval induction and other developmental events. *Genes & Development,* 9(6), pp.742-755.

Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., Wu, C. & Allis, C. D. 2006. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature,* 442(7098), pp.86-90.

Xiao, Y., Bedet, C., Robert, V. J., Simonet, T., Dunkelbarger, S., Rakotomalala, C., Soete, G., Korswagen, H. C., Strome, S. & Palladino, F. 2011. Caenorhabditis elegans chromatin-associated proteins SET-2 and ASH-2 are differentially required for histone H3 Lys 4 methylation in embryos and adult germ cells. *Proceedings of the National Academy of Sciences of the United States of America,* 108(20), pp.8305-8310.

Yang, X. J. & Seto, E. 2003. Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. *Current Opinion in Genetics & Development,* 13(2), pp.143-153.

Yang, X. J. & Seto, E. 2008. The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nature Reviews Molecular Cell Biology,* 9(3), pp.206-218.

Yang, Y. W., Flynn, R. A., Chen, Y., Qu, K., Wan, B., Wang, K. C., Lei, M. & Chang, H. Y. 2014. Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. *Elife,* 3, pp.e02046.

Yanginlar, C. & Logie, C. 2018. HDAC11 is a regulator of diverse immune functions. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms,* 1861(1), pp.54-59.

Yoo, A. S., Bais, C. & Greenwald, I. 2004. Crosstalk between the EGFR and LIN-12/Notch pathways in C. elegans vulval development. *Science,* 303(5658), pp.663-666.

Young, S. R., Mumaw, C., Marrs, J. A. & Skalnik, D. G. 2006. Antisense targeting of CXXC finger protein 1 inhibits genomic cytosine methylation and primitive hematopoiesis in zebrafish. *Journal of Biological Chemistry,* 281(48), pp.37034-37044.

Young, S. R. & Skalnik, D. G. 2007. CXXC finger protein 1 is required for normal proliferation and differentiation of the PLB-985 myeloid cell line. *DNA Cell Biology,* 26(2), pp.80-90.

Yu, C., Fan, X., Sha, Q. Q., Wang, H. H., Li, B. T., Dai, X. X., Shen, L., Liu, J., Wang, L., Liu, K., Tang, F. & Fan, H. Y. 2017. CFP1 Regulates Histone H3K4 Trimethylation and Developmental Potential in Mouse Oocytes. *Cell reports,* 20(5), pp.1161-1172.

Zaret, K. S. & Carroll, J. S. 2011. Pioneer transcription factors: establishing competence for gene expression. *Genes & Development,* 25(21), pp.2227-2241.

Zhang, C. H., Wang, J. X., Cai, M. L., Shao, R., Liu, H. & Zhao, W. L. 2018. The roles and mechanisms of G3BP1 in tumour promotion. *Journal of Drug Targeting,* 27(3), pp.1-6.

Zhang, C. L., Mckinsey, T. A., Chang, S., Antos, C. L., Hill, J. A. & Olson, E. N. 2002. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell,* 110(4), pp.479-488.

Zhang, L., Wang, J., Pan, Y., Jin, J., Sang, J., Huang, P. & Shao, G. 2014. Expression of histone H3 lysine 4 methylation and its demethylases in the developing mouse testis. *Cell and Tissue Research,* 358(3), pp.875-883.

Zhang, T., Cooper, S. & Brockdorff, N. 2015. The interplay of histone modifications - writers that read. *EMBO Reports,* 16(11), pp.1467-1481.

Zhang, Y., Chen, D., Smith, M. A., Zhang, B. & Pan, X. 2012. Selection of reliable reference genes in Caenorhabditis elegans for analysis of nanotoxicity. *PLOS One,* 7(3), pp.e31849.

Zhang, Y., Leroy, G., Seelig, H. P., Lane, W. S. & Reinberg, D. 1998. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell,* 95(2), pp.279-289.

Zhang, Y. & Reinberg, D. 2001. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development,* 15(18), pp.2343-2360.

Zheng, S., Wyrick, J. J. & Reese, J. C. 2010. Novel trans-tail regulation of H2B ubiquitylation and H3K4 methylation by the N terminus of histone H2A. *Molecular and Cellular Biology,* 30(14), pp.3635-3645.

Zhou, B. O. & Zhou, J. Q. 2011. Recent transcription-induced histone H3 lysine 4 (H3K4) methylation inhibits gene reactivation. *Journal of Biological Chemistry,* 286(40), pp.34770-34776.

Zhou, Q., Li, T. D. & Price, D. H. 2012. RNA Polymerase II Elongation Control. *Annual Review of Biochemistry,* 81, pp.119-143.

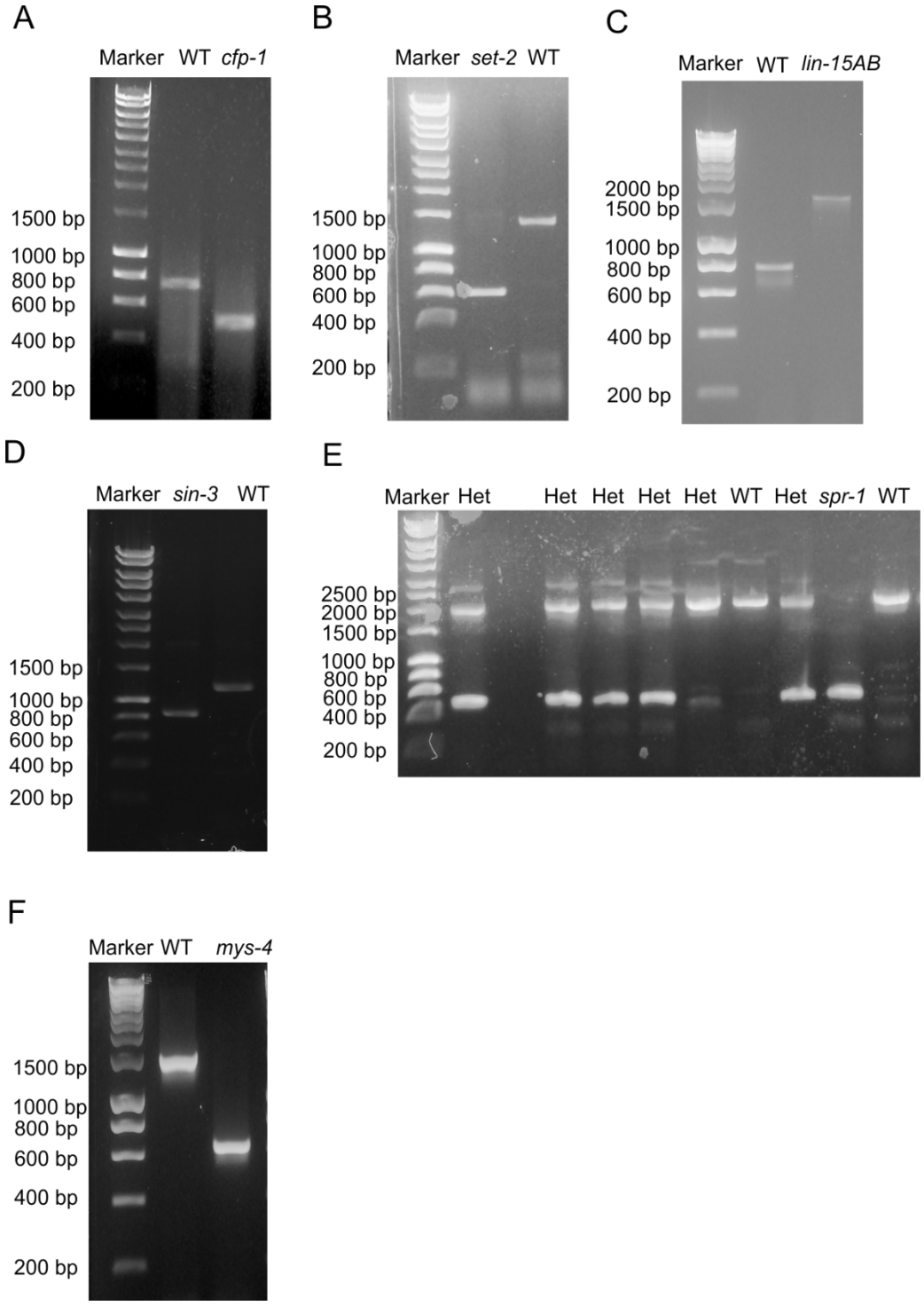
Zhu, B., Mandal, S. S., Pham, A. D., Zheng, Y., Erdjument-Bromage, H., Batra, S. K., Tempst, P. & Reinberg, D. 2005. The human PAF complex coordinates transcription with events downstream of RNA synthesis. *Genes & Development,* 19(14), pp.1668-1673.

Zinzen, R. P., Senger, K., Levine, M. & Papatsenko, D. 2006. Computational models for neurogenic gene expression in the Drosophila embryo. *Current Biology,* 16(13), pp.1358-1365.

Zupkovitz, G., Tischler, J., Posch, M., Sadzak, I., Ramsauer, K., Egger, G., Grausenburger, R., Schweifer, N., Chiocca, S., Decker, T. & Seiser, C. 2006. Negative and Positive Regulation of Gene Expression by Mouse Histone Deacetylase 1. *Molecular and Cellular Biology,* 26(21), pp.7913-7928.

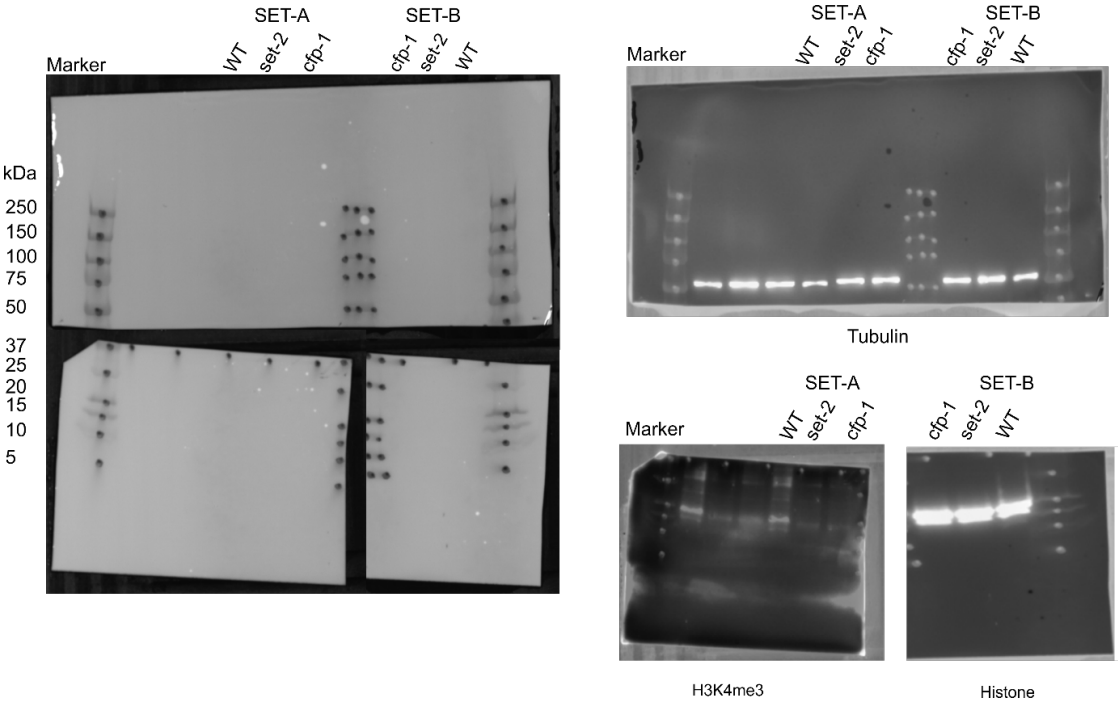
Appendix

9.1 Appendix: Gel images depicting genotype of mutants



(A) Gel image of genotyping of wild type animals and *cfp-1(tm6369)* mutants, band size of the wild-type is 708 bp and the *cfp-1(tm6369)* mutant is 454 bp. (B) Gel image of wild-type animals and *set-2(bn129)* mutants, band size of the wild-type is 1340 bp and the *set-2(bn129)* mutant is 592 bp. (C) Genotyping of wild-type animals and *lin-15AB(n765)* mutants. Wild-type band size is 750 bp. The *lin-15AB(n765)* mutant contains an insertion, thus band size in the mutant is 1.7 Kb. (D) Gel image of wild-type animals and *sin-3(tm1276)* mutants, band size of wild-type animals is 1153 bp and the *sin-3(tm1276)* mutant is 838 bp. (E) Genotyping image of wild-type animals and *spr-1(ok2144)* mutants, band size of the wild-type is 1921 bp and the *spr-1(ok2144)* mutant is 455 bp. Please note the gel shown is genotyping of F1 from the *spr-1(ok2144)* mutant mother. Two bands (one wild-type and one mutant) in the same lane suggests that the animal is heterozygous (Het). (F) Gel image of wild-type animals and *mys-4(tm3161)* mutants, size of wild type band is 1442 bp and *mys-4(tm3161)* band is 621 bp.

9.2 Appendix: Western blot analysis



To measure the levels of H3K4me3 in WT, *cfp-1(tm6369)* and *set-2(bn129)* mutants, western blot analysis was carried out. The cell extracts of WT, the *cfp-1(tm6369)* and

set-2(bn129) mutants were loaded twice on the same gel (designated “set A” and “set B”). After electrophoresis proteins were transferred from the gel to a nitrocellulose membrane. The membrane was cut horizontally and vertically as indicated, to be able to probe with an anti-tubulin antibody (top) and antibodies against Histone 3 (right bottom) and H3K4me3 (left bottom). This was necessary, because H3K4me3 and Histone 3 have approximately the same molecular weight (15 kDa). H3K4me3 levels in both COMPASS mutants was dramatically reduced compared to wild-type. Please note that the first three lanes (not labelled) contain a different set of samples not relevant for this analysis.

9.3 Appendix: CT values of *gpdh-1*, *pmp-3* gene and *tba-1* genes

Average CT value ± SEM of *gpdh-1* transcripts in wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants at control and 150 mM salt concentration. The table is the average of 4 biological replicates.

	gpdh-1	***pmp-3***	***tba-1***
Wild-type (control)	29.58 ± 0.20	22.02 ± 0.54	19.12 ± 0.35
Wild-type (150mM)	27.47 ± 0.05	22.48 ± 0.55	19.70 ± 0.58
***cfp-1* (control)**	28.80 ± 0.80	22.48 ± 1.04	19.68 ± 0.74
***cfp-1* (150 mM)**	26.56 ± 1.18	23.10 ± 1.61	20.36 ± 1.63
***set-2* (control)**	29.57 ± 0.64	22.26 ± 0.22	20.08 ± 0.06
***set-2* (150 mM)**	26.08 ± 0.19	22.09 ± 0.58	19.63 ± 0.43

9.4 Appendix: CT values of heat inducible and housekeeping genes

Table showing the average CT value ± SEM of *F44E5.4, C12C8.1* and *hsp-16.2* transcripts in wild-type, *cfp-1(tm6369)* and *set-2(bn129)* mutants at the control and after heat shock. The table is the average of 3 biological replicates

	hsp16.2	***pmp-3***	***tba-1***	***F44E5.4***	***C12C8.1***
Wild-type (control)	24.28 ± 0.35	21.62 ± 0.28	19.04 ± 0.12	26.33 ± 0.13	28.79 ± 0.12
Wild-type (heat shock)	19.73 ± 0.14	22.74 ± 0.26	20.58 ± 0.26	18.06 ± 0.34	22.84 ± 0.27
***set-2* (control)**	23.70 ± 0.07	21.87 ± 0.07	19.48 ± 0.18	25.33 ± 0.18	27.29 ± 0.36
***set-2* (heat shock)**	17.98 ± 0.26	23.21 ± 0.13	20.70 ± 0.04	16.73 ± 0.08	20.50 ± 0.13
***cfp-1* (control)**	24.13 ± 0.26	22.94 ± 0.09	19.80 ± 0.62	26.13 ± 0.36	28.54 ± 0.10
***cfp-1* (heat shock)**	17.06 ± 0.15	22.67 ± 0.45	19.56 ± 0.3	14.94 ± 0.15	19.48 ± 0.12