Unraveling the Role of a *Drosophila* Zinc Finger Transcription Factor in Epigenetic Cellular Memory

Department of Life Sciences

Syed Babar Ali School of Science & Engineering

Lahore University of Management Sciences

Submitted in Partial Fulfilment of the Requirements

for the Degree of Bachelor of Science

Muhammad Abdullah Jauhar

2022

Advisor:

Prof. Dr. Muhammad Tariq

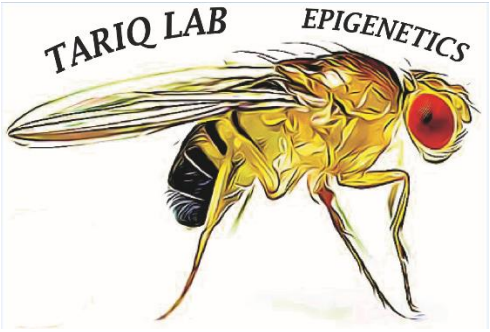
Supervisor: Prof. Dr. Muhammad Tariq

Department of Life Sciences

Syed Babar Ali School of Science & Engineering

Lahore University of Management Sciences

Date of defense: 29th April 2022



LUMS
A Not-for-Profit University

Lahore University of Management Sciences

Department of Life Sciences

Certificate of Approval

I hereby recommend that the thesis prepared under my supervision by:

Muhammad Abdullah Jauhar

---

on title

Unraveling the role of a Drosophila Zinc Finger Transcription Factor in Epigenetic Cellular Memory

---

be accepted in partial fulfillment of the requirements for the BS. degree.

Supervisor (Chairperson of Defense Committee)

Prof. Dr. Muhammad Tariq

Name *Signature* *Date:*

Recommendation of Thesis Defense Committee:

(At least one more member is required to sign as per rule)

Name *Signature* *Date:*

Declaration of originality

I hereby declare that the work accomplished in this thesis is the result of my research carried out in the Epigenetic Lab of the Department of Biology, Lahore University of Management Sciences, Pakistan from February 2019 to May 2022. This thesis has neither been published previously nor does it contain any material from the published resources that can be considered a violation of international copyright law. The findings and conclusions came from my research with the complete discussion under the supervision of Dr. Muhammad Tariq. I, further declare that no part of this work has been presented for any other degree anywhere.

Signature: _____

Name: _____

Date: _____

Acknowledgments

First and foremost, I would like to thank my parents and my family. I dedicate this work to them. It was their dreams & zest that made me who I am today, and words can never do just to the extent to which I am forever indebted to them.

Similarly, nothing would have been possible without Prof. Muhammad Tariq. I am forever grateful to him for allowing me to join his lab during my freshman over three years ago. He gave me the opportunity to learn the purest form of science, to develop the wings of fearlessness, and also taught me the skills to pass on the baton to incoming generations of scientists. Science blooms in his lab due to his presence!

I am extremely thankful to the members of the epigenetic family for being there for me throughout these years. Special thanks to Najma Shaheen, Jawad Akhtar & Hina Ahmad for generously investing a great amount of time and energy towards my experiments and the never-ending discussion on scientific ideas. Many thanks to Khalida Mazhar, Mahnoor Hussain Bakhtiari, Muhammad Ahmad, Zain Umer, Muhammad Haider Farooq Khan & Ammad Shaukat for being amazing lab mates and making this journey extraordinary.

A big thanks to the people ensuring the integrity of the nuts and bolts of our caravan; Iqra Manzoor, Abrar Hussain Shahid, Robin Javed, Johnson Bhatti & Sagheer Ahmad. They are the true heroes behind the discoveries that we make in the labs. I can never thank them enough for going out of the way to assist my research.

I would like to thank my dear friends, Saifullah Shakeel, Savera Manzar, Saad Imran Rana & Muhammad Ibrahim Sani, who made every moment of this journey completely worth it. They were there for me in the happiest and darkest of times. Their presence brought me joy whenever it was most needed, and I was very fortunate to have them in my life.

Last but not the least, I want to thank the people working under my supervision; Sadaf Yasmin, Hassan Mujtaba, Salman Tuasene Khawaja, Hooraina Hassan Siddiqui, Rafay Ali Khan, Sheikh Zohaib Ali, Amna Noor, Shahmeen Munir, Waqar Safdar, Moeez Dar & Mahnoor Sakina Naqvi. Their presence taught me many things and I would have never learned that if it wasn't for them.

Nothing would have been possible without these people, and I am very grateful to all of them.

Muhammad Abdullah Jauhar

2022-10-0032

BS Biology Class of 2022

In the loving memory of Thomas Hunt Morgan, Max Delbruck, Barbara McClintock & Seymour Benzer

May long live their Science!

خلاصہ:

حیوانات کی ان اقسام میں جو اپنی پیچیدگی میں سادہ حیاتیات پر غلبہ رکھتیں ہیں، خلئیے کی تقدیر کے تعین سے جڑے امتیازی جین اکسپریشن کے فیصلوں کی برقراری پولی کامب اور ٹرائی تھورائکس پروٹینز کے خاندانوں کے انحصار میں ہے۔ جہاں دہائیوں کی تحقیق سے ہم پولی کامب گروپ کے متعلق غیرمعمولی حقائق سے آشنا ہیں، وہیں ٹرائی تھورائکس گروپ کے بارے میں موجود معلومات انتہائی محدود ہے جبکہ ان دونوں کا مناسب توازن حیاتیات کی درست تشکیل نو کیلئے ناگزیر ہے۔ اس محدود دائرہ کار کو وسیع کرنے کے حدف سے ہماری سائنسی لیباٹری نے ایک رورس جیجیٹک سکرین منعقد کی جس میں دو سو سے زائد نئےممکنہ ٹرائی تھورائکس گروپ کے ارکان کی نشاندہی کی گئی۔ اس فہرست میں ایک زنک فنگر ٹرانسکرپشن فیکٹر ، سی جی اے بی سی، اپنے انفرادی خصائص کی بنا پر توجہ کا مرکوز رہا۔ اس کی سی ٹو ایچ ٹو ڈومینز جو چھوٹے اوزاروں کی طرح مخصوص ڈی این اے کو پڑھنے کی صلاحیت رکھتیں ہیں، ٹرائی ،تھورائکس کو ان کے کروماٹن کارخانوں پر بھرتی کرنے کی صلاحیت رکھتیں ہیں۔ کئی دہائوں کی تحقیق کے باوجود بالائی جنیٹکس کے اس پہلو پر بددستور گہرے بادل چھائے ہوئے ہیں اور سائنسدان منزل کی تلاش میں کوئی بڑی فتح حاصل کرنے میں ناکام رہے ہیں۔ اس مقالے میں مالیکیولر کلونگ انجام دی گئی جس میں سی جی اے بی سی کو مختلف ایپیٹوپس کے ساتھ وابستہ کیا گیا۔ فلیگ ٹیگ شدہ سی جی اے بی سی کی مدد سے کروماٹن عمئنوپرائسیپیٹیشن انجام دی گئی جس میں حیران کن نتائج سامنے آئے۔ سی جی اے بی سی ان تمام ہومیوٹک اور غیرہومیوٹک کارخانوں پر موجود تھی جن پرگذشتہ رسالوں میں پولی کامب اور ٹرائی تھوراعکس کی موجودگی کی شناخت کی جاچکی ہے اور جنہیں ہم پولی کامب ریسپونس علیمینٹنس کہتے ہیں۔ اس کے ساتھ ساتھ سی جی اے بی سی کی عدم موجودگی ٹرائی تھوراعکس کے لگائے گئے کروماٹن کے تیسرے ہسٹون کے ستائسوئں لاعزین کی ایسیٹالیشن کے عالمی درجوں میں حیران کن حد تک کمی کا باعث بنتی ہے۔ یہ نشان ٹرائی تھوراعکس کے باعث جین کی ایکٹیویشن کی بنیاد ہے۔ ان دریافتوں سے ایک بات تو عیاں ہے کے سی جی اے بی سی ٹرائی تھوراعکس کے ساتھ باہمی تعاون میں فرائض انجام دیتاہے۔ چونکہ یہ ماں کی جانب سے بچے کو دیا جاتا ہے،اسلئے اس میں انفرادی کابلیت موجود ہے جو ابتدا میں آنے والے پیغامات اور ٹرائی تھوراعکس کے درمیان پلُ کا کردار ادا کرسکتا ہے۔ مستقبل میں ان حقائق سے پردا اٹھانا ناگزیر ہوگا جن کی بنا پر سی جی اے بی سی اپنے فرائض کی انجام دہی کر رہاہے۔

Abstract:

In higher eukaryotes, differential gene expression patterns linked to cell fate determination are maintained by the Polycomb & Trithorax group of proteins (PcG & TrxG). While members of PcG have been well studied, the TrxG remains largely uncharacterized. In the quest for novel regulators of the Trithorax family, a reverse genetic screen identified a Zinc finger protein, CGABC, as a candidate TrxG regulator of the epigenetic cellular memory. The presence of sequence-specific C2H2 zinc finger domains makes this protein an excellent candidate for the recruitment of TrxG to their target loci, a field that is poorly understood to date. In this dissertation, molecular cloning of full-length CGABC ORF was performed, and it was placed in frame with different epitope tags. Chromatin immunoprecipitation of FLAG-tagged CGABC revealed its association with the Polycomb Response Elements (PREs) and promoter regions of homeotic and non-homeotic targets of PcG/TrxG. Depletion of CGABC leads to a significant inhibition of the global H3K27ac levels, which is the hallmark of TrxG-mediated gene activation. Since CGABC is maternally deposited, it has the unique potential of linking early pioneer factors with the recruitment of TrxG complexes to their target loci in a cell type-specific manner. The future aim is to understand the mechanism through which CGABC supports TrxG in establishing a transcriptionally active state of genes.

Table of contents

1. Introduction:

Multicellular organisms have a remarkable ability to give rise to different cell types stemming from a single invariable genome (Blackledge, Rose, and Klose 2015). Humans contain over 200 different cell types, all with the same genetic composition. Yet these cell types are significantly different in terms of structure and functionality. Cell fate determination is a critical aspect of development where the embryo gives rise to these distinct cell types. Studies of the last few years have suggested that cell type-specific Transcription Factors (TFs) have the ability to influence cell fates. These transcription factors are involved in the initiation of differential gene expression that is required for cell fate determination and commitment to specific cell lineage and consequently leaving the multipotent state (Kuroda et al. 2020). Transcription factors are effector molecules that respond to an incoming signal by acting either as gene repressors or activators. Once the initial signal has gone, these gene expression profiles must be preserved in a locus-specific manner, which is crucial for maintaining cell fates. This maintenance of transcriptional states of gene expression associated with certain cell fate is referred to as epigenetic cellular memory. In metazoans, evolutionary conserved Polycomb (PcG) and Trithorax (TrxG) groups of proteins contribute to epigenetic cellular memory by maintaining the repressed and active gene expression states, respectively (Schuettengruber et al. 2017).

While the mechanism through which these gene regulatory decisions are made and later maintained remains largely unknown, several covalent chromatin modifications have been known to influence gene expression patterns and play a pivotal role in the maintenance of epigenetic cell memory. It has been long known that the promoter regions are methylated at cytosine residues (5mC) in a CG dinucleotide context. It is known that 5mC contributes to the repression of genes when promoter regions of a gene are methylated and 5mC contributes to the differential gene expression patterns during development. Methylation of DNA was also the first heritable epigenetic modification discovered as the methylation patterns in the CG dinucleotide context are preserved upon the replication of DNA by the maintenance methyltransferase enzymes (Sun and Zhu 2022; Ling and Ronn 2019; Bird and Wolffe 1999; Deaton and Bird 2011).

Eukaryotic genomes are sophistically packaged into nucleosomes that form higher-order chromatin structures. The nucleosomal structure comprising the histone octamer provides an excellent opportunity for another level of epigenetic gene regulation via covalent modifications on specific histone residues (Kornberg and Lorch 2020; 1999; Gibney and Nolan 2010). Characterization of several histone modifications has led to mechanistic insight regarding their effect on gene expression states. Several complexes have been purified with abilities to add post-translational modifications to the histones, including but not limited to phosphorylation, acetylation, methylation, ubiquitylation, sumoylation, etc. These heritable modifications are crucial for the maintenance of cell type-specific gene expression during development (Jenuwein and Allis 2001; Grewal and Moazed 2003; Delcuve, Rastegar, and Davie 2009).

Almost 70 years ago, genetic screens in *Drosophila* identified a group of genes responsible for the regulation of the homeotic genes. Later it was identified that PcG is involved in the maintenance of repression of developmental genes while TrxG, antagonizing PcG, is involved in maintaining active gene states. The PcG/TrxG antagonism is crucial for the maintenance of differential gene expression throughout development (Lewis 1978; Morata, Sánchez-Herrero, and Casanova 1986; Bender, Turner, and Kaufman 1987; Kennison 1995).

PcG/TrxG maintains tissue-specific gene expression patterns by locally modifying and altering the chromatin structure. Several PcG/TrxG members are readers and writers of histone covalent modifications. Enhancer of zester [E(z)], a core PcG member, catalyze tri-methylation of Lysine 27 of Histone H3 (H3K27me3) (Cao et al. 2002; Czermin et al. 2002); a covalent mark that has been associated with gene PcG mediated repression. Similarly, dRING protein mono-ubiquitilates Lysine 119 of Histone H2A (H2AK119ub) to promote gene repression (H. Wang et al. 2004). Together these two modifications are a hallmark of PcG-mediated transcriptional gene silencing. On the other hand, TrxG complexes are associated with the H3K4me3 (Roguev et al. 2001), H3K36me2 (Tanaka et al. 2007), and H3K27ac (Tie et al. 2009). These histone modifications are the hallmarks of TrxG-mediated gene activation (Schuettengruber et al. 2017).

As PcG/TrxG system is crucial for the maintenance of the epigenetic cellular memory, defects in their function have been shown to result in various carcinomas (Mills 2010). Hence it is essential to decipher the precise mechanism through which PcG/TrxG regulates transcription, a process that remains largely unknown.

While we have a sufficient understanding of how PcG/TrxG are modifying the chromatin, how they are recruited to their target site still remains a mystery. Although the key players of both PcG/TrxG are highly conserved during the stages of development, the elements that they occupy on the chromatin can diverge significantly. Polycomb response elements (PREs) were discovered in *Drosophila melanogaster* as the DNA regulatory elements that recruit PcG to promote gene repression during the stages of development. However, these sequences are not fully conserved in mammals including humans. Tissue-specific gene regulation requires locus-specific target recognition by these complexes. For PcG members, transcription factors such as Polyhomeotic (Pho), Polyhomeotic like (Phol), and Pipsqueak (Psq), have been shown to recruit PcG to their targets (Schuettengruber et al. 2017). For TrxG recruitment to the chromatin, very little evidence is available. Zeste was shown to recruit Brahma, the fly homolog of ATP-dependent chromatin remodelers, SWI2/SNF2, at the cellular memory module of the Fab7 region that regulates the expression of the homeotic gene Abdomen-B (AbdB) (Déjardin and Cavalli 2004). *Drosophila* GAGA factor, the product of the Trithorax-like (Trl) gene colocalizes with TrxG complexes on several PREs but its role in their recruitment remains unclear as it is not a part of most TrxG complexes (Strutt, Cavalli, and Paro 1997). It suggests that future studies need to focus on TrxG recruiting proteins in greater detail. Together, the absence of sequence-specific DNA domains in the core PcG/TrxG complexes yet their precise accuracy in regulating their target loci in a locus-specific manner raises the question of what exactly guides these complexes to their target loci. It has been proposed that DNA binding TFs and non-coding RNAs might bridge this gap in our understanding of PcG/TrxG recruitment (Schuettengruber et al. 2017).

Zinc finger proteins are one of the most abundant classes of proteins having a diverse range of functions. Their unique ability to interact with DNA, RNA, and proteins makes them an important regulator of various processes in the cell including transcriptional regulation, signal transduction, protein degradation, and several other key processes (Yusuf et al. 2021; Cassandri et al. 2017).

Over the decades of research on the PcG/TrxG system as the molecules making up the cellular memory; advances in TrxG regulation are lagging behind PcG. A possible reason for this is that PcG complexes are largely conserved while TrxG complexes are very diverse. The diverse nature of TrxG complexes makes their characterization a challenging job.

Our lab recently performed a genome-wide RNAi screen in cells using a luciferase-based reporter to discover novel TrxG factors (Umer et al. 2019). We identified over 200 candidate TrxG genes, clustering in a variety of cellular proteins, including TFs, RNA-binding proteins, cell cycle regulators, nuclear transport proteins, etc. Among the list of TrxG candidates was CGABC, a C2H2 zinc finger transcription factor. CGABC is maternally deposited in the fly embryo and is one of the early genes to be expressed. As it contains zinc finger domains which are known specific DNA binding domains, we wanted to decipher its possible role in the recruitment of TrxG for gene activation. Analysis through bioinformatics predicted that CGABC is a sequence-specific RNA Polymerase II mediated transcriptional activator (FlyBase Gene Report). Swenson et al. carried out a genome-wide RNAi screen for regulators of the Heterochromatin Protein 1 (HP1a) in *Drosophila* and discovered CGABC as one of the top hits in that screen, suggesting it's a strong positive regulator of HP1a. This is surprising because HP1a is strongly linked with gene repression and the formation of constitutive heterochromatin. We showed in Umer et al. that CGABC is a candidate TrxG gene which is associated with gene activation. Apparently, these are two conflicting findings, however, with the recent insights into the components of the constitutive heterochromatin, several members have been shown to associate with gene activation as well. Similarly, CGABC might provide a novel link between epigenetic cellular memory and maintenance of heterochromatin.

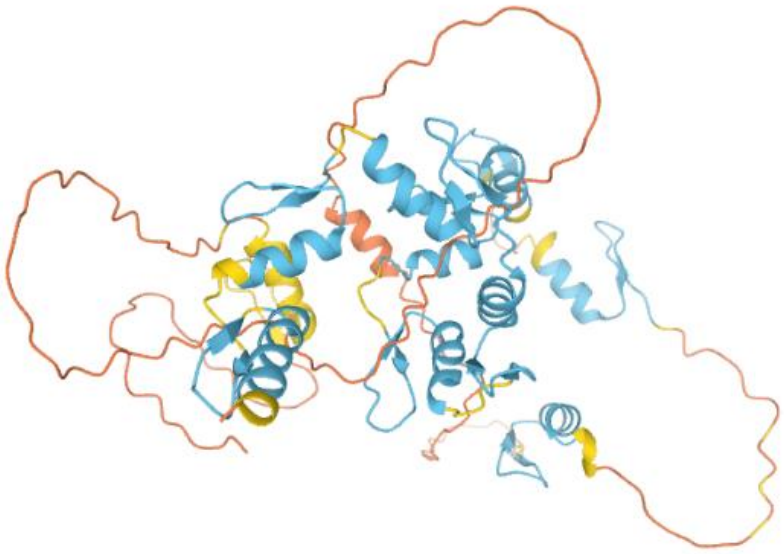


Fig1*:* CGABC structure predicted by alphafold

1.1 Aims & Objectives

The molecular and biochemical activity of CGABC remains elusive. In this dissertation, I aim to experimentally validate the prediction regarding its role in TrxG-mediated transcriptional gene regulation. In particular, this thesis aims to achieve the following specific objectives:

- Molecularly characterize CGABC to show its association with TrxG
- Validate if CGABC functions in the heterochromatin alongside HP1a besides its role in the maintenance of gene activation by TrxG

2. Literature Review:

2.1 Maintaining cell fates

Multicellular eukaryotes begin their lives as a single cell that goes through a sequential development plan of cellular specifications and growth. During early embryonic development, events like body axis and pattern formation initiate the process of differentiation that leads to the formation of diverse cell types that eventually contribute to organogenesis. The identity of a cell is largely dependent on its differential gene expression profile. Given that DNA replication and progression of the cell cycle removes the majority of the chromatin modifiers and associated proteins, it becomes crucial to restore chromatin structure and function once the cell division has completed. It is made possible through the inheritance of chromatin states during DNA replication and cell division that cells retain their 'memory' of the differential gene expression patterns after each cell cycle. To effectively achieve this task, sophisticated mechanisms evolved to maintain cell fates once the cell fates are determined so that cell type-specific gene expression patterns are faithfully propagated in all cell lineages.

In *Drosophila*, the earliest stages of development are driven by the action of maternal effect genes. These maternal effect genes are produced in the ovary of the mother and their transcripts are physically deposited on the embryo's cytoskeleton by the follicle cells. The spatial localization of these genes is crucial for the subsequent development after fertilization (Schüpbach and Wieschaus 1986; Schupbach and Wieschaus 1986; Gavis and Lehmann 1992). The antagonizing effects of two of the maternal genes, bicoid, and nanos, are essential for the establishment of anterior-posterior body axes. As the fly embryo grows as a syncytium, the products of these maternal genes are able to diffuse freely and act in a concentration-dependent manner till the blastoderm stage when the embryo eventually becomes cellularized (C. Wang and Lehmann 1991; G. Struhl 1989). The anterior effect genes, bicoid, and hunchback are involved in the patterns formation of anterior regions while their posterior counterparts, nanos and causal, are involved in the formation of posterior regions (Mlodzik and Gehring 1987; Stathopoulos and Newcomb 2020; Nusslein-Volhard 1991; Chille et al. 2021; Schüpbach and Wieschaus 1986).

Since maternal effect genes are transient in nature, the maternal effect genes activate gap genes, another class of transcription factors, which are the first zygotic genes to get activated. They get their name from the phenotype of large gaps observed in the embryo patterning upon their mutation (Nüsslein-volhard and Wieschaus 1980). Interactions between gap genes result in the activation of pair-rule genes that divides the embryo into 14 distinct parasegments. These pair-rule genes pass the baton to the segment polarity genes which further define the parasegment boundaries (Ingham, Baker, and Martinez-Arias 1988). Once the parasegments have been defined, the stage is set for the homeotic genes to come in and define the developmental identity of each parasegment. Once a homeotic gene has been switched on in a spatio-temporal manner, it will remain active or repressed in a specific segment for the entire life of the organism (Gary Struhl and White 1985). Maintenance of gene expression patterns of homeotic genes is crucial for development because misregulation of homeotic genes results in severe morphological defects and in severe cases leads to the death of growing embryos. Importantly, the expression of homeotic genes is maintained in specific body segments by PcG/TrxG system. This 'memory' of keeping a homeotic gene in the "on" or "off" state is what was first defined as epigenetic cellular memory and hence led to the establishment of the PcG/TrxG paradigm (Jackle et al. 1986).

Ed Lewis and his wife Pamela Lewis identified a mutation that results in homeotic transformation specific to the Bithorax complex. This locus which they referred to as "Polycomb" was hypothesized to be regulating the master regulators of the bithorax complex (Lewis 1978). Later studies identified more mutations showing a similar phenotype as Polycomb mutants and collectively they were called Polycomb Group genes (PcG). Importantly, mutations in PcG genes did not disrupt the early parasegmental expression of homeotic genes suggesting that they were not the initial activators of these genes but rather were maintaining the expression during the later stages of development (G Struhl 1985).

Certain loss of function mutations of the Bithorax complex in *Drosophila* were initially named as Trithorax. Later a whole group of mutations was discovered that suppressed the extra sex combs phenotype by Pc mutants and these suppressors of PcG were grouped as TrxG genes. This revealed that TrxG antagonizes PcG function by acting as anti-suppressors to ensure faithful expression of

their target genes which happens to be the targets of PcG as well (Kingston and Tamkun 2014; Schuettengruber et al. 2007; Kassis, Kennison, and Tamkun 2017).

2.2 Biochemical characterization of PcG

PcG consists of two largely conserved complexes; Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2) (Schuettengruber et al. 2017; Ringrose and Paro 2004; Grimaud, Nègre, and Cavalli 2006).

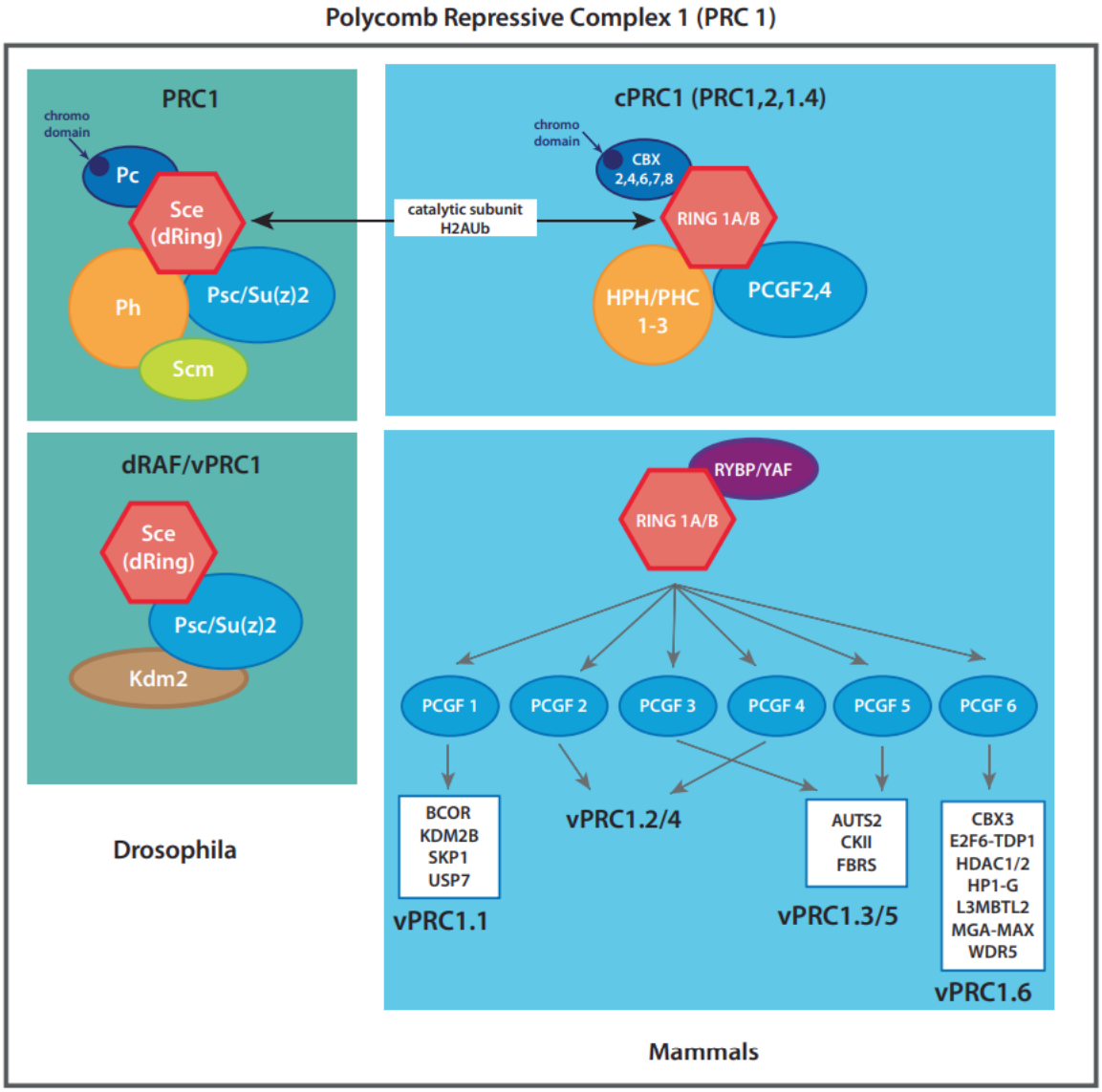


Fig2: PRC1 complexes in *Drosophila* & Mammals (Paro et al 2020)

Polycomb, Posterior sex combs, Sex comb extra & Polyhomeotic were shown to be the key members of the PRC1 (Schwartz and Pirrotta 2007). PRC1 utilizes the chromodomain of Pc to bind to H3K27me3 to block transcription and chromatin remodeling. Sex comb extra, also known as dRING, is the catalytic component of PRC1 with E3 ubiquitin ligase activity that monoubiquitylates H2AK118 (Kuroda et al. 2020).

PRC2 also comprises of four core members including E(z), Suppressor of zeste 12 [Su(z)12], a chromatin assembly factor p55, and Extra sex combs (Esc). The histone methyltransferase, E(z), of PRC2 tri-methylates H3K27 which is a hallmark of PcG-mediated gene repression (Muller et al., 2002). The other components of PRC2 have been shown to enhance E(z)'s methyl transferase activity (Czermin et al. 2002; Vidal 2019; Fuchsberger et al. 2016).

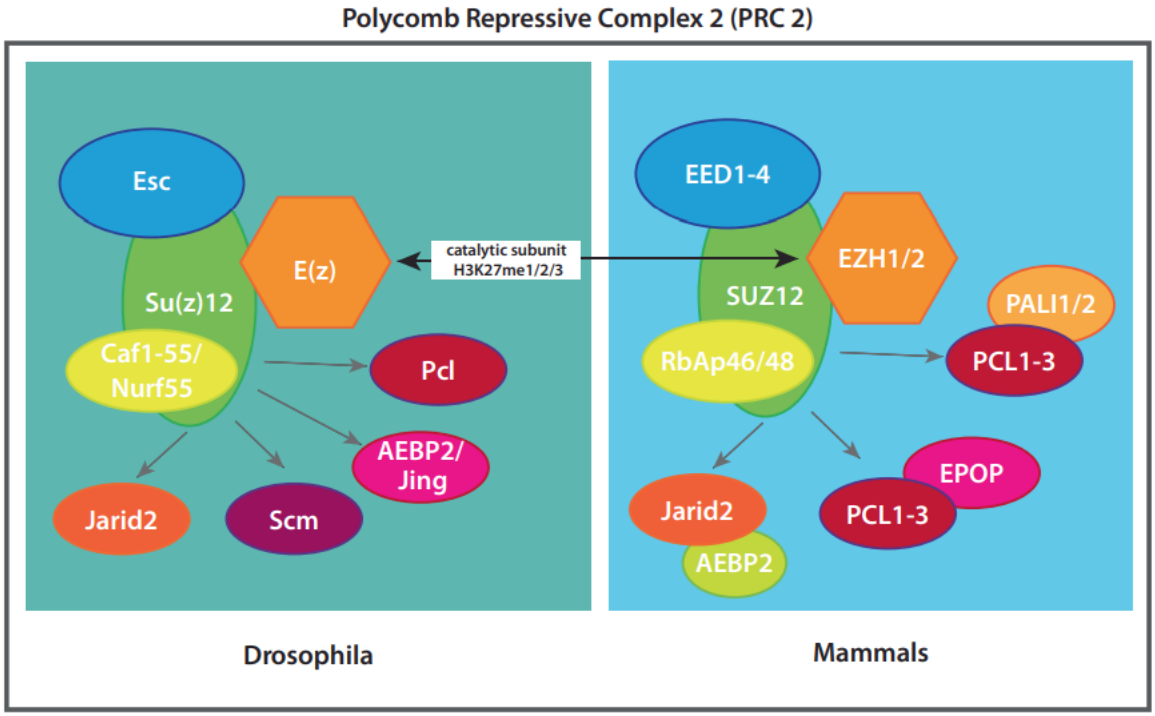


Fig3: PRC2 in *Drosophila* & Mammals (Paro et al 2020)

2.3 Biochemical characterization of TrxG

On the other hand, the Trithorax family was found to be very diverse as compared to the Polycomb. A large and divergent set of proteins involved in several histone post-translational modifications (PTMs), chromatin remodeling, and transcriptional coactivation by recruiting RNA Pol II

comprised the TrxG. Evolutionary conserved dCOMPASS complex has been shown to deposit H3K4me3 via its Trx counter whereas the TAC1 complex deposits H3K27ac via its dCBP counterpart (Miller et al. 2001; Petruk et al. 2001; Roguev et al. 2001). SWI/SNF complexes and several ATP-dependent chromatin remodeler's activity is dependent on the TrxG members (Tie et al. 2009; Kingston and Tamkun 2014).

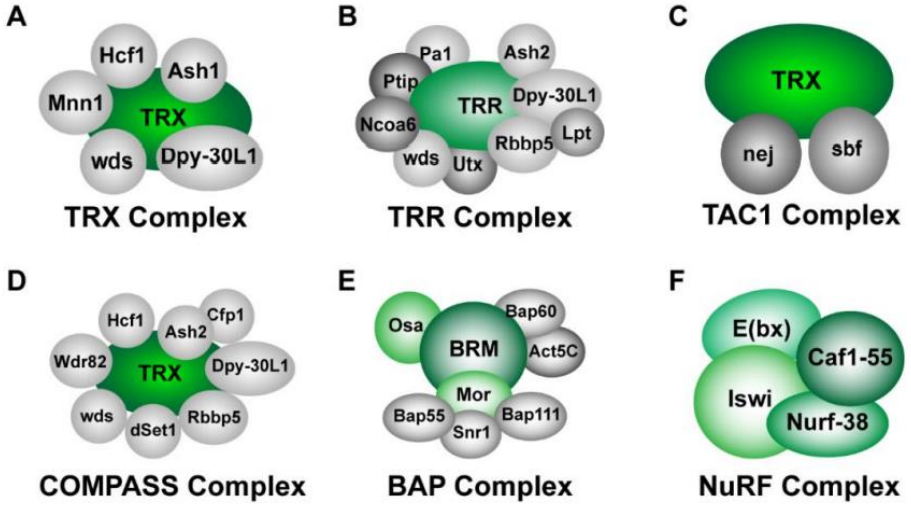


Fig4: Trithorax complexes (Kassis, Kennison, and Tamkun 2017)

PcG/TrxG members are present in the cell at all times but they only specifically regulate their target loci when they are required. This paradox becomes even more interesting with the fact that only a handful of members of these groups have the ability to identify and bind specific DNA sequences. Importantly, specific *cis*-acting DNA elements known as PREs were identified in *Drosophila* as specific regulatory elements that recruit PcG to promote silenced gene states throughout the stages of development. As mentioned earlier, DNA binding members: Pho, PhoI, Trl, etc contain binding sites for these PREs. Little information is available for specific recruitment of Trithorax group complexes to their target loci. It is suggested that TrxG recruitment may involve an even more diverse set of TFs than PcG recruitment. Recently, a pioneer factor FOXA1 was shown to recruit MLL3 to direct H3K4me1/2 at the enhancer regions. From the insights PcG/TrxG field has gained in the last 70 years, one thing gets absolutely clear; we are still not there yet to explain the specificity of these complexes. Transcription Factors are a strong candidate that provides the missing blocks of this jigsaw.

CGABC is a maternally deposited transcription factor, predicted to contain 10 C2H2 zinc finger domains. C2H2 zinc fingers are the largest class of eukaryotic transcription factors but to date, it has largely remained uncharacterized (Fedotova et al. 2017). A zinc ion forms a coordinated covalent bond between two cysteine and two histidine residues from a beta-hairpin and an alpha helix respectively. C2H2 domain has been reported to bind with DNA and RNA, and mediate protein-protein interactions (Gamsjaeger et al. 2007; Brayer and Segal 2008). Members of PcG/TrxG rely on zinc finger proteins for their DNA binding. GAGA factor, Trithorax-like, contains highly conserved C2H2 zinc finger domains. PRC2 core members Su(z)12 and Jing also contain zinc finger domains. Importantly, Pleiohomeotic contains C2H2 zinc finger domains shown to recruit it to the PREs (Schwartz and Pirrotta 2007; Schuettengruber et al. 2017). Hence, CGABC becomes an exciting target for the recruitment of TrxG in a cell type-specific manner.

3. Materials and Methods

3.1 RNA Isolation

Cells from a fully confluent T25 flask were harvested in a 2ml microfuge tube at 4000 rpm for 5 minutes at 4ºC. The cell pellet was washed with 1X PBS to remove the media followed by centrifugation at 4000 rpm for 5 minutes at 4ºC. Cells were resuspended in 1ml of TRIzol reagent (Invitrogen) on ice for lysis. Once the solution was homogenous, the tube was incubated at room temperature for 3.5 minutes. 200µl of chloroform was added to the mixture followed by vortexing for 15 seconds or until the color becomes bright pink. The tube was then incubated at room temperature for 2 minutes followed by centrifugation at 13000 rpm for 15 minutes at 4ºC (layers separate well at 4ºC). The tube was gently removed from the centrifuge and the upper aqueous layer (containing nucleic acids) was carefully removed without disrupting the below organic protein layers. The supernatant was moved to a fresh tube and 1ml of isopropanol was added to precipitate RNA. The mixture was thoroughly mixed and incubated at room temperature for 10 minutes followed by overnight incubation at -80ºC. The next day, the solution was spun at max speed for 30 minutes at 4ºC to pellet down RNA. The supernatant was carefully discarded, and the pellet was washed with 1ml 75% ethanol made in Diethyl pyrocarbonate (DEPC) water followed by centrifugation at max speed for 5 minutes at 4ºC. The supernatant was removed carefully, and the pellet was air dried for 10 minutes. Finally, the dried pellet was dissolved in 20µl of DEPC water by gentle flicking until the pellet disappears. RNA was stored at -20ºC.

3.2 DNAse treatment

9ug of RNA was treated with TURBO DNase (Invitrogen) according to the manufacturer's instructions to remove any possible genomic DNA contamination. The DNAse treated RNA (DTR) was stored at -20ºC and used for subsequent cDNA synthesis.

3.3 cDNA synthesis

2µg of DTR was used for cDNA synthesis using Superscript III first strand synthesis according to the manufacturer's protocol (Invitrogen).

3.4 dsRNA synthesis

Template DNA for dsRNA synthesis was amplified using specific primers fused to the T7 promoter sequence as described in (Swenson et al. 2016). Megascript T7 transcription kit (Ambion) was used to set up the *in vitro* transcription (IVT) reaction using the purified amplified DNA (2µg). The dsRNA pellet was dissolved in DEPC water with a final concentration of 1µg/ul and was stored in 20µl aliquots at -80ºC.

3.5 Molecular Cloning:

Gateway cloning was used to generate constructs expressing epitope-tagged *CGABC* ORF. Primers were designed to amplify *CGABC* ORF from the cDNA of *Drosophila S2* cells. The PCR product was eluted using PCR purification kits according to the manufacturer's instructions (Invitrogen). To generate the entry clone, the purified PCR product was ligated with the linearized *pENTR DTOPO* vector (Thermo Scientific) in the ratio of 2:1. The ligation mixture was incubated at room temperature for 30 minutes. 2ul of this mixture was then transformed into a mixture of *E. coli* competent cells. Transformants were selected on Kanamycin plates (50ug/ul) followed by plasmid isolation. Each clone was digested through restriction digestion and confirmed by sequencing (Macrogen).

For the cloning of destination vectors, LR clonase reactions (Thermo Scientific) were set up using the sequencing confirmed *pENTR-CGABC* along with the respective destination vectors: *pAG, pPM,* and *pMTHF* (Drosophila Gateway Vector Collection). The transformants were selected on Ampicillin plates (100ug/ul) followed by plasmid isolation and restriction digestion with appropriate enzymes.

3.6 RNAi-mediated knockdown in *Drosophila* Dmel2 cells

1 million cells were seeded in the Express V medium containing 20µg of dsRNA against CGABC for 5 days at 25ºC followed by RNA isolation and cDNA synthesis as described above.

3.7 Transient Transfection & Immunofluorescence:

Drosophila S2 cells were transfected with *pAG-CGABC* using the effectene transfection reagent according to the manufacturer's instructions (Qiagen). The media was discarded after 24 hours to

avoid cytotoxicity and a fresh medium was added. Cells were harvested after 72hrs and transferred to the coverslip. Cells were fixed using 4% paraformaldehyde for 30 min, followed by incubation with DAPI for 30 minutes. The coverslips were mounted with flouromont (Thermo Scientific) and placed on glass slides for confocal microscopy.

3.8 Stable Transfection

Drosophila S2 cells were transfected with *pMTHF-CGABC* using effectene transfection reagent (Qiagen) according to the manufacturer's instructions. Transfected cells were selected against Hygromycin B (250 ug/ul) for 10 passages. Once all the cells in the control died, the transfected cells were transferred to a new flask and the selection pressure with antibiotic was stopped. Finally cloning of CGABC in frame with FLAG, HIS tag was confirmed via western blot with anti-FLAG antibody of whole cell lysate prepared 48hrs after induction with 500μM $CuSO_4$.

3.9 Western Blotting:

Cells were harvested at 2000rpm for 5 minutes at 4ºC. The cell pellet was washed with 1ml of ice-chilled PBS. Cells were again harvested at 2000rpm for 5 minutes at 4ºC. The pellet was treated with 1ul of *Benzonase* (Millipore). The pellet was incubated on ice for 10 minutes. 200ul of 2X SDS loading buffer was used to lyse the cell pellet. The mixture was heated at 95C for 10 minutes followed by incubation on ice for 1 minute. The mixture was centrifuged at 10000rpm for 5 minutes. Samples were resolved on 15% SDS-PAGE along with a molecular marker (Seeblue plus2). The resolved proteins were transferred to a nitrocellulose membrane using a wet transfer system at 4ºC for two hours at 100V. The transfer of proteins to the membrane was confirmed by *Ponceau* staining. The membrane was then washed with 1X PBST and blocked for 45 minutes at room temperature with 5% skimmed milk prepared in 1X PBST. Anti-FLAG antibody (1:1000) was prepared in the same blocking solution and the membrane was incubated overnight at 4ºC. Next morning, the blot was washed thrice with 1X PBST. Finally, anti-mouse HRP conjugated secondary antibody (Abcam) dilutions were prepared (1:2000) in the same blocking solution. The blot was incubated with the secondary antibody mixture for 2hrs at room temperature with gentle shaking. The blot was then washed with 1X PBST twice and once with 1X PBS. Blots were developed using ECL reagents (GE Healthcare) and were analyzed using BioRad ChemiDoc. Anti-tubulin was used as loading control (1:1000) followed by the same procedure as described above.

3.10 ChIP from *pMTHF-CGABC* cells

Cells harvested from a fully confluent T25 flask were equally divided into 3 T75 flasks and were allowed to grow for a day. Next day, they were induced with 500µM CuSO4 and were allowed to grow for another 48 hrs. The cells were gently scrapped off and pooled into 250ml autoclaved flask. 200µl of cells were removed for cell counting using a hemocytometer. A magnetic stirrer was placed in the flask and the cells were cross-linked using 1% formaldehyde as fixative solution for strictly 10 minutes at room temperature. Glycine was added to the mixture to a final concentration of 125mM to stop the fixation. Cells were incubated on ice for 10 minutes. The fixed cells were centrifuged at 2000 rpm for 5 minutes at 4ºC followed by washing with 1X PBS to remove the fixative solution. These cells were lysed in Solution A including protease inhibitors followed by incubation on ice for 5 minutes. The cells were again spun at 2000 rpm for 5 minutes and subsequently washed twice with solution B. Cells were incubated on ice for 5 minutes. Nest, the supernatant was discarded and cells were resuspended in sonication buffer in such a way that every 330µl contained 30 million cells. These aliquots were sonicated to shear the chromatin using bioruptor (Diagenode) for 25min (30 sec on/30 sec off). The lysate was immediately centrifuged at max speed for 15 minutes at 4ºC and the supernatant was collected into a fresh microfuge tube.

25µl protein A and 25µl protein B DYNA beads (Novex) were washed thrice for 5 minutes using 1X RIPA buffer along with the protease inhibitors. For this, 1ml of 1X RIPA was added to the beads and they were incubated at 4ºC on a rotator fixed at 20 rpm. The supernatant was discarded each time the beads were washed, after the final wash they were resuspended in 50µl of 1X RIPA buffer. Equal volume of 2X RIPA buffer was added to the chromatin and the mixture was added to 50µl of DYNA beads for pre-clearing. The tube was set on rotation at 20rpm for 2 hours at 4ºC after which the cleared chromatin was transferred to a fresh tube. 10% of the pre-cleared chromatin was separated for input. The rest was incubated with 5µl of anti-FLAG for overnight at 4ºC (20rpm). Next day 50µl of DYNA beads (Novex) were added to the tube to pull down the antibody-bound chromatin for 4 hours at 20rpm at 4ºC. After which the beads were separated using DYNA mag. The beads were washed thrice with 1X RIPA, once with 1X LiCl buffer, and lastly once with 1X TE for 5 minutes each at 4ºC. Elution buffer was prepared fresh and 250µl was used to elute the chromatin from the beads. This was carried out at 65ºC for 15 minutes on the heat block with

gentle mixing. Reverse cross-linking was done using 20µl of 5M NaCl at 65ºC overnight incubation on the heat block. 10µl of 0.5M EDTA (pH 8), 20µl 1M Tris (pH 6.5), and 2µl of proteinase K (10ug/ul) was added to the reverse cross-linked chromatin and the mixture was incubated at 45ºC for 2 hours with gentle mixing on the heat block.

The reverse crossed-linked chromatin was subjected to phenol-chloroform extraction and DNA was purified from the mixture by using 1 volume of phenol, chloroform, and isoamyl alcohol (25:24:1) followed by centrifugation at 13000 rpm for 5 minutes at 4ºC. The aqueous layer containing DNA was transferred to a fresh microfuge tube and the organic layer was extracted once again with 1 volume chloroform. The mixture was centrifuged at 13000 rpm for 5 minutes at 4ºC and the aqueous layer was transferred to the same microfuge tube. To precipitate the DNA, 1µl glycogen, $1/10^{th}$ volume of 3M Sodium Acetate (pH 5.2), and 2.5 volumes of absolute ethanol were added to the aqueous layer and it was incubated at -80ºC overnight. The next day the tube was centrifuged for 30 minutes at max speed 4ºC to pellet down the purified DNA. The pellet was washed with 70% ethanol, air dried for 10 minutes, and dissolved in PCR water.

qPCR was set up using specific primers as described in (Umer et al. 2019; Shaheen et al. 2021) and enrichment levels were calculated using the ΔΔCT method as described in (Schmittgen and Livak 2008).

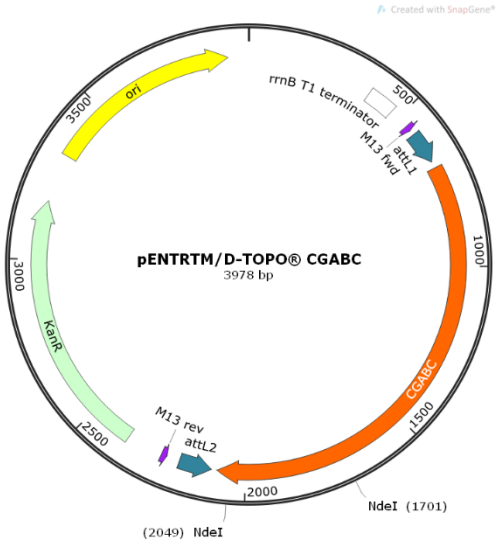
4. Results & Discussion

Since *CGABC* is a novel gene and totally uncharacterized, it was aimed to start with the molecular cloning of *CGABC* ORF in different vectors to generate epitope-tagged CGABC under the control of specific promoters.

4.1 Molecular cloning of *CGABC* in *pENTR™/D-TOPO* and confirmation through restriction digestion.

The full-length *CGABC* ORF was PCR amplified from cDNA prepared from *Drosophila* S2 cells (Appendix 2), and it was cloned in Gateway *pENTR/D-TOPO* gateway entry vector (Fig 5A and see methods). Plasmid DNA isolated from bacterial transformants was digested with *NdeI* for confirmation of recombinant clone. After digestion with *NdeI*, the release of 348bp of the expected fragment confirmed the cloning of the CGABC full-length open reading frame in the entry vector (Fig5B and see methods). The recombinant entry clone for CGABC was confirmed by sequencing and no mutations were detected.

A



B

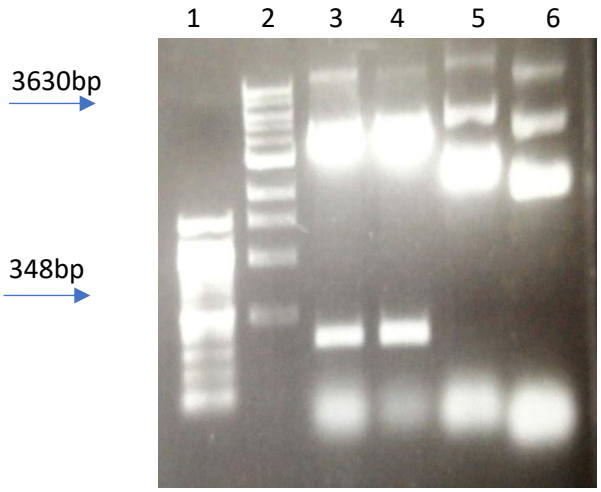


Fig5: Molecular cloning of full-length CGABC in Gateway entry vector. (A) Schematic diagram of Gateway entry vector carrying *CGABC* cloned in *pENTR/D-TOPO*. Restriction sites for *NdeI* enzyme used for confirmation of recombinant plasmid are highlighted. (B) Restriction digestion of recombinant plasmid shown in A with NdeI show expected fragments seen at 348bp & 3630bp which confirm cloning of CGABC in entry vector. Lanes 3, 4 show digested plasmids whereas lanes 5, 6 contain uncut plasmids in lanes 3 and 4 respectively; L1 100 bp ladder; L2 1kb ladder.

Molecular cloning of *CGABC* in *pAG-CGABC, pPM-CGABC* vector for epitope tagging

The CGABC ORF cloned in the entry vector was used to subclone in different gateway destination vectors using LR clonase reaction to recombine CGABC ORF in frame with coding regions of different epitope tags. For example, destination vector pAG which carries EGFP tag was used for N-term tagging of CGABC under actin promoter, and pPM vector was used for Myc tagging of CGABC under UAS inducible promoter (Figure 6 A, B). Plasmid from bacterial transformants which carry potential *pAG-CGABC* clone was confirmed using restriction digestion with *BamH1* which resulted in two expected fragments of 4555bp and 3075bp (Figure 6C). This confirmed cloning of CGABC in frame with the coding region of EGFP at N-term of CGABC and this EGFP-CGABC is under actin promoter. Moreover, cloning of *CGABC* ORF in frame with the coding region of Myc tag under UAS promoter in *pPM* vector was confirmed using restriction digestion with *Xho1* which resulted in the release of expected 9516bp & 2615bp fragments (Figure 6C).

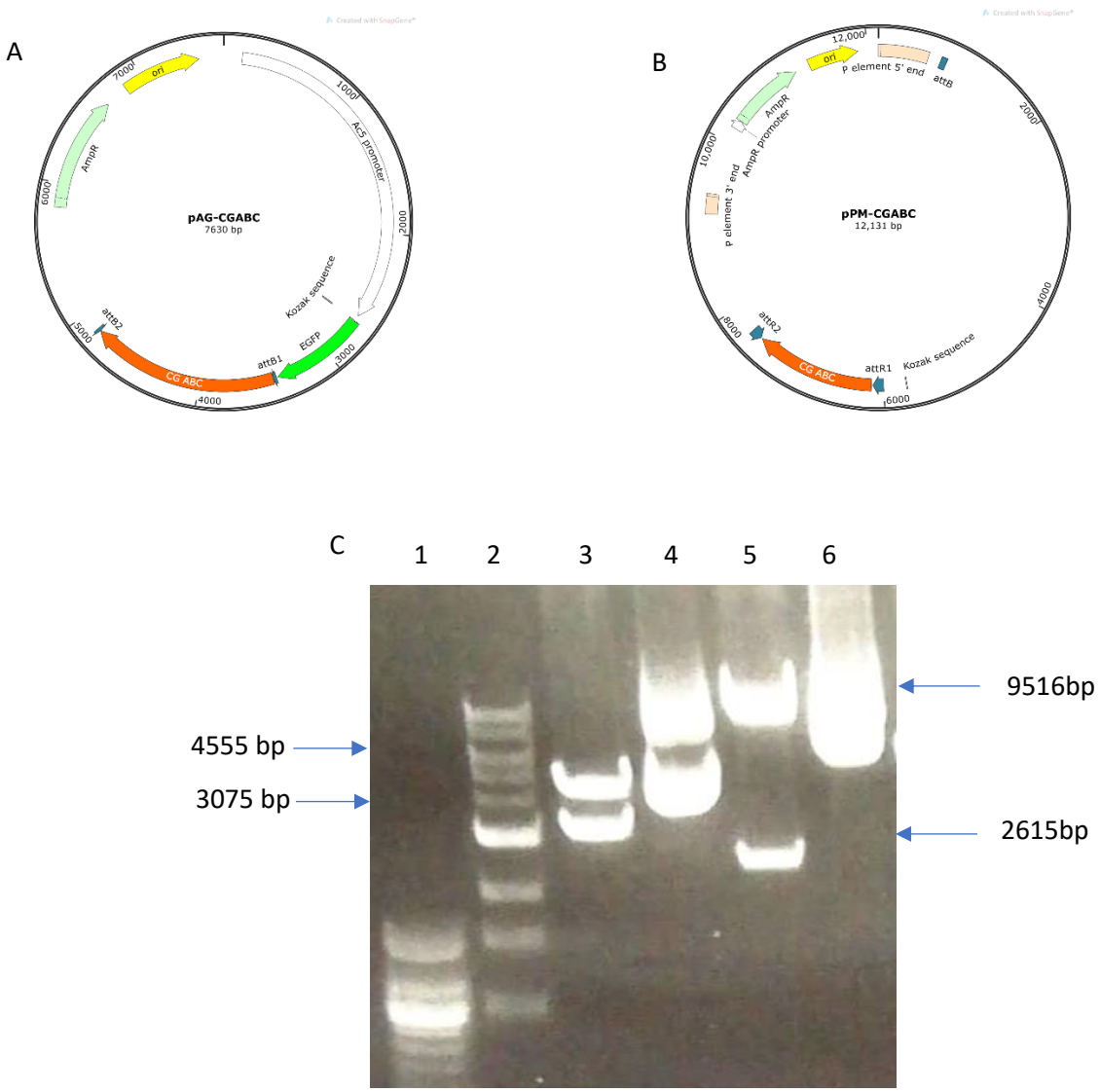


Fig6: Molecular cloning of full-length CGABC in Gateway destination vector. (A) Schematic diagram of Gateway entry vector carrying CGABC cloned in *pAGW* under Act5c promoter and N-term EGFP tag. (B) Schematic diagram of Gateway entry vector carrying CGABC cloned in *pPMW* under UAS promoter and N term Myc tag. (C) Restriction digestion of recombinant plasmids in A & B with *BamHI* and *XhoI*, respectively, confirms the cloning CGABC in the destination vectors. Lane 3 shows *pAGCGABC* digested with *BamHI* giving fragments at 4555bp & 3075bp. Lane 5 shows *pPMCGABC* digested with *XhoI* giving fragments at 9516bp & 2615bp. Lane 4 & 6 show uncut plasmids in lane 3 & 5 respectively. Lane 1: 100bp ladder Lane 2: 1kb ladder.

4.2 Molecular cloning of *CGABC* in *pMTHF* and confirmation through restriction digestion.

The *CGABC* ORF cloned in the entry vector was also used to subclone in *pMTHFW* using LR clonase reaction. *pMTHFW* carried FLAG & His tags was used for the N-term tagging of CGABC under the metallothionein promoter (Figure 7A). Plasmids from bacterial transformants that carried the *pMTHF-CGABC* clone were digested with *HindIII* that gave two bands at 7343bp and 255bp which confirmed the presence of CGABC in frame with N-term FLAG & His tag. This FLAG-His-CGABC is under the control of metallothionein promoter that responds to heavy metals such as Copper and Cadmium (Figure 7B).

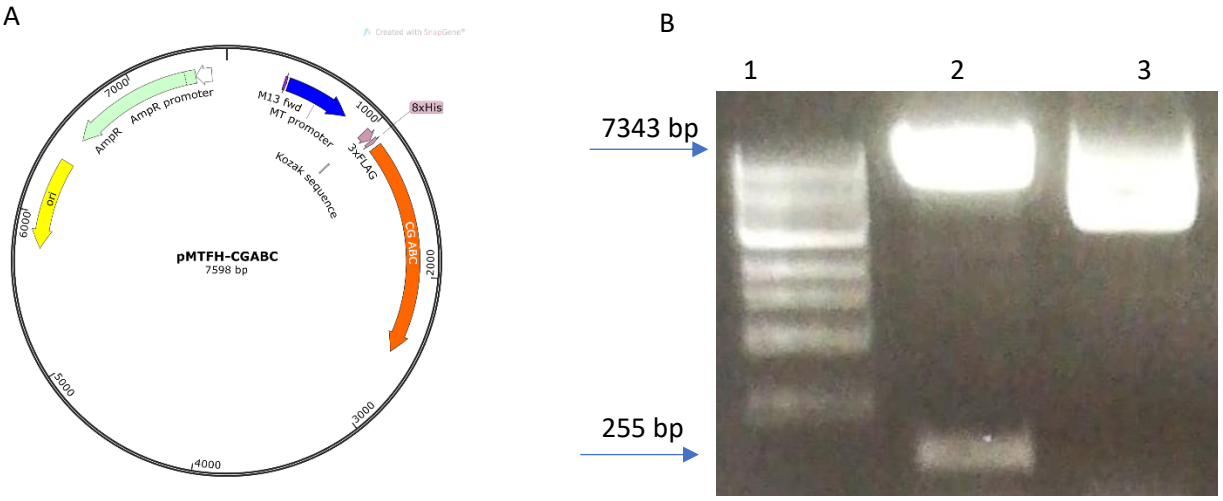


Fig6: Molecular cloning of full-length CGABC in Gateway destination vector pMTHFW. (A) Schematic diagram of Gateway destination vector carrying *CGABC* cloned in *pMTHFW*. (B) Restriction digestion of recombinant plasmid shown in A with *HindIII* show expected fragments seen at 7343bp & 255bp which confirm cloning of CGABC in *pMTHFW*. Lane 2 shows digested plasmid whereas lanes 3 contains uncut plasmid; Lane 1 is 1kb ladder.

4.3 Molecular cloning of *CGABC* in *pET21a(+)* and confirmation through restriction digestion

All the CGABC constructs described above will results in overexpression which is not optimal for the characterization of CGABC, however, all these constructs will be extremely beneficial in generating initial functional data. Since CGABC is a transcription factor that is expressed in very

little amounts, hence it is crucial to characterize it at the endogenous expression levels. To this end, it is imminent to generate antibodies against CGABC. The domainless region from 36th to 179th amino acid of the full-length CGABC ORF was chosen as the antigenic region. The region was enriched with hydrophilic amino acids which made the region highly immunogenic. The absence of domains ensured that the polyclonal antibody will specifically detect CGABC. The antigenic region of CGABC was PCR amplified from the cloned Gateway entry vector *pENTR-CGABC* using *EcoR1* and *HindIII* in the forward and reverse primers respectively. *pET21a(+)* and the PCR amplified product was digested with *EcoRI* & *HindIII* and ligation was set up using T4 DNA Ligase (Invitrogen). Plasmids isolated from the bacterial transformants containing the potential *pET21a(+)-AntigenicCGABC* construct were confirmed via restriction digestion. Upon double digestion with *EcoRI* & *HindIII*, the fragments obtained at 435bp and 5430bp confirmed the presence of the antigenic region of CGABC in the pET21a(+) vector in frame with the His tag at the C term (Figure 7 A & B). This antigenicCGABC-His construct is under the control of the T7 promoter. The cloned construct was further confirmed via sanger sequencing and the antigenic region of CGABC was found to be in frame with the His tag.

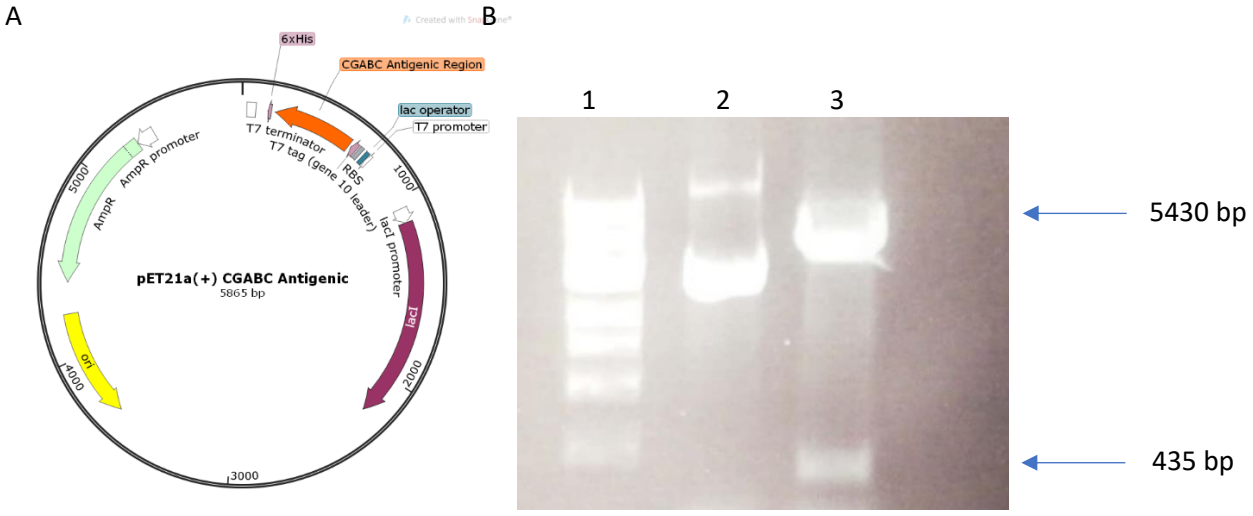
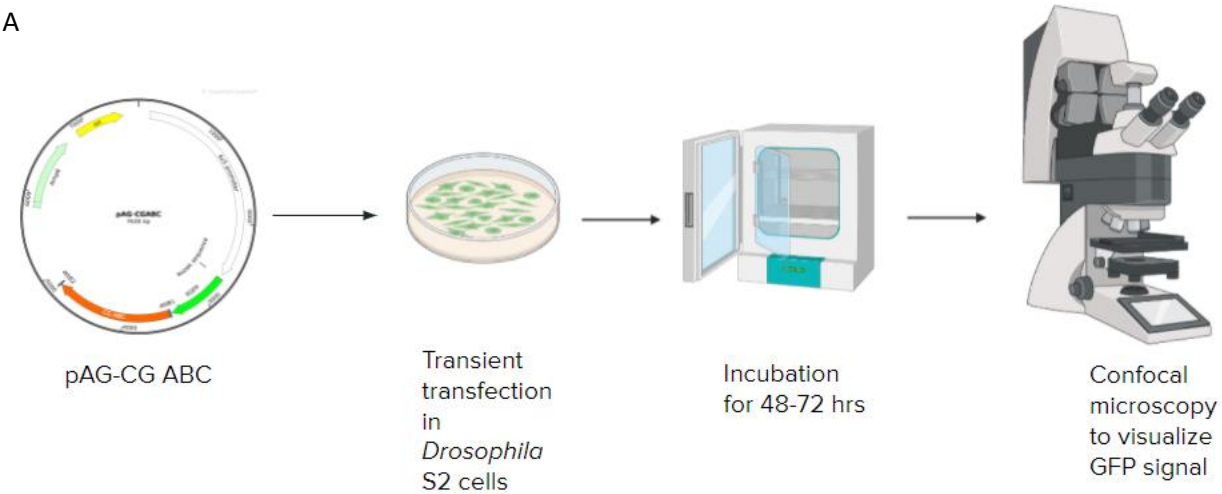


Fig6: Molecular cloning of antigenic CGABC in pET21a(+) expression vector. (A) Schematic diagram of pET21a(+) expression vector carrying antigenic region of CGABC with C-term His tag under the T7 promoter. (B) Restriction digestion of recombinant plasmid shown in A with *EcoRI* and *HindIII* show expected fragments seen at 5430bp & 435bp which confirm cloning of CGABC in pET21a(+). Lane 3 shows digested plasmid whereas lane 2 contains uncut plasmid; Lane 1 is 1kb bp ladder.

4.4 Subcellular localization of *CGABC* using *pAGCGABC*

Since PcG and trxG factors primarily act at the level of chromatin and are localized in the nucleus, it is crucial to determine where CGABC also localizes in the nucleus. To this end, *Drosophila S2* cells were transiently transfected with *pAG-CGABC* expressing EGFP tagged full-length CGABC under constitutively active actin promoter. After 48hrs hours of transfection, EGFP-tagged CGABC was visualized in transfected cells using confocal microscope. It was observed that EGFP-CGABC was localized inside the nucleus. However, it was observed that EGFP-CGABC does not entirely overlap with DAPI rather a strong focal expression was seen inside the nuclear membrane which is most likely the nucleolus as reported by (Swenson et al. 2016) as well. Since CGABC was over-expressed using actin promoter, this may not be the accurate localization of CGABC as overexpression of nuclear proteins often results in their nucleolar localization (Musinova et al. 2011). As CGABC is a transcription factor and they are expressed at very low levels in the cell, so significantly increasing its expression may have resulted in its nucleolar localization. However, EGFP-CGABC signal in the nucleus clearly indicates that molecular cloning of CGABC in frame with EGFP and nucleolar staining also indicates CGABC localizes in the nucleus.

A



B

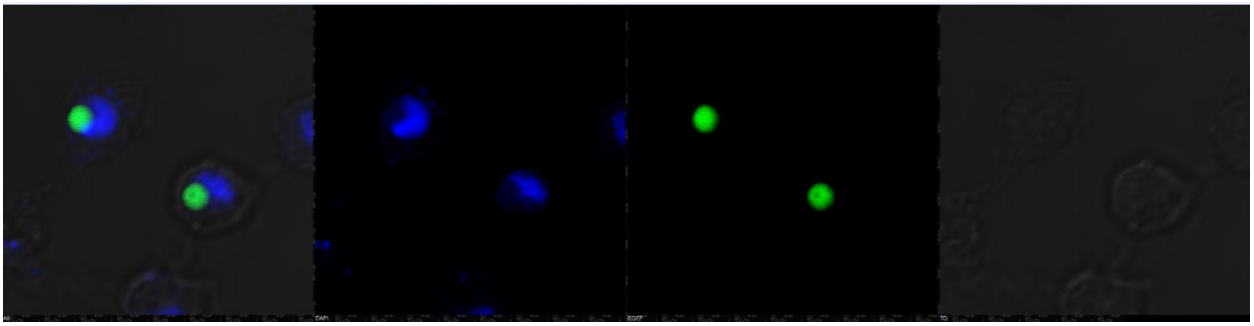


Fig8: CGABC shows nuclear localization: (A) Schematic workflow for the determination of subcellular localization of CGABC. (B) After 48hrs of transfection, the cells were visualized under confocal microscope at 20x magnification. EGFP (green) signal represents CGABC localized in the nucleolus which partially overlaps with DAPI (blue) signal.

4.5 Generation of stable cells expressing *FLAG*-tagged *CGABC*

To determine the molecular function of CGABC, stable cells expressing FLAG-tagged CGABC under copper inducible promoter were generated using *Drosophila S2* cells. As an endogenous antibody was not available against CGABC, the FLAG tag would provide the basis of experimentation. The *pMTHFCGABC* was transfected in *Drosophila* S2 cells and the transfected cells were selected against hygromycin for 10 passages. Finally, the cell line was confirmed via western blot using anti-FLAG antibody after 48 hours of induction with 500µM $CuSO_4$. The total cell lysates were probed with anti-FLAG antibody. As compared to un-induced cells, used as control, a specific signal for FLAG was observed in induced cells at 53.3kDa on the western blot.

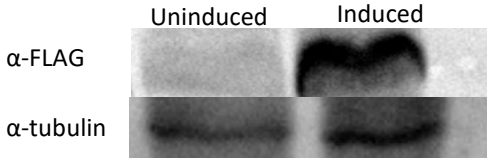


Fig9: *Drosophila* S2 cells stably expressing FLAG-CGABC was confirmed through Western blot. Whole cell lysates were prepared from stable cells after 48h of induction with 500µM CuSO4. As compared to un-induced cells, FLAG-CGABC was detected in the induced cells whereas no signal was seen in the uninduced cells. Tubulin levels were used as loading control.

4.6 CGABC associates with PcG/TrxG targets on chromatin

Following the generation of FLAG-CGABC stable cells and the nuclear localization of CGABC, it became important to determine if CGABC is interacting with the chromatin and if it colocalizes with the PcG/TrxG targets on chromatin. To this end, chromatin immunoprecipitation was performed from cells expressing FLAG-tagged CGABC after 48 hours of induction with 500μM $CuSO_4$.

ChIP-qPCR analysis showed that CGABC is enriched at PREs/promoter regions of different homeotic (Fab7, Bxd, Dfd) and non-homeotic targets (Pnt, Pnr, Psq) of PcG/TrxG. Association of CGABC with chromatin indicates that CGABC likely interacts with PcG/TrxG at chromatin and hence might be required for regulating the expression of these PcG/TrxG targets. However, enrichment of CGABC seen at intergenic regions where PcG/TrxG are not reported to bind highlights the pioneer factor-like behavior of CGABC. Pioneer transcription factors have been reported to bind to chromatin at vast regions to cause decondensation (Balsalobre and Drouin 2022). Additionally, the presence of CGABC on the intergenic regions could also be due to the overexpression of CGABC which signifies the need for anti-CGABC antibody to understand the precise role of endogenous CGABC.

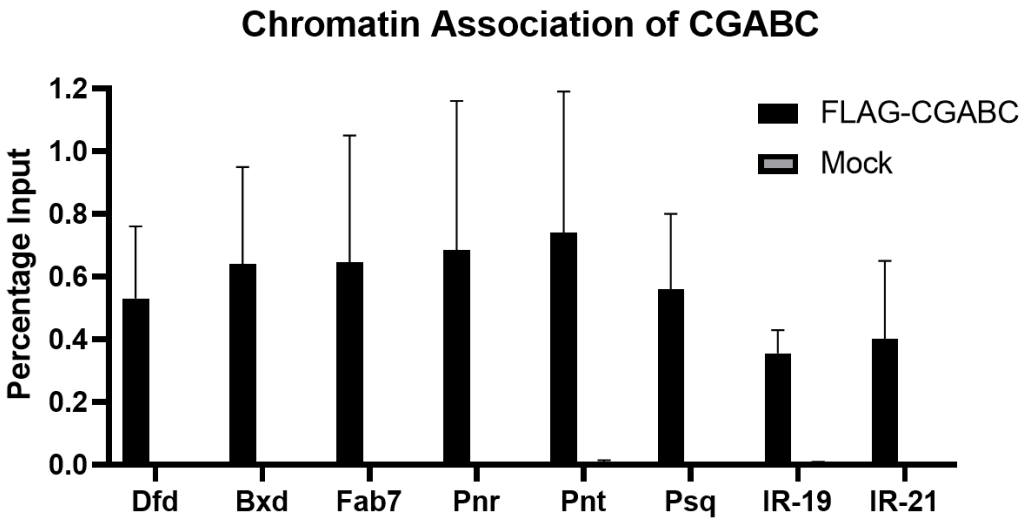


Fig10: As compared to mock (IgG) ChIP, qPCR analysis on purified DNA from anti-FLAG ChIP show strong enrichment of CGABC at several homeotic and non-homeotic target of PcG/TrxG. Purified DNA from mock and anti-FLAG ChIP was subjected to qPCR using specific primers in PREs and promoter regions of different homeotic (Dfd, bxd, Fab7) and non-homeotic genes (*pnr, pnt, psq*). No enrichment of these regions was seen in ChIP performed without antibody. IR-19 and IR-21 (Papp and Müller 2006) represent intergenic regions beyond BX-C region where PcG and TrxG proteins are normally not reported.

4.6 CGABC maintains global H3K27ac levels in Drosophila cells

Since CGABC was found to associate with PcG/TrxG targets in chromatin and CGABC was discovered as a candidate TrxG gene, it was investigated if depletion of CGABC shows any TrxG-like molecular effect at chromatin. For this, the global levels of H3K27ac which is a hallmark of TrxG mediated gene regulation were determined upon CGABC knockdown. It is pertinent to mention that H3K27ac is mutually exclusive to the H3K27me3 mark which is catalyzed by the PRC2 complex (Cao et al. 2002; Tie et al. 2009). Analysis of whole cell lysates from Dmel2 cells where CGABC was depleted using RNAi revealed a drastic reduction in H3K27ac levels as compared to control cells which were incubated with lacZ dsRNA. Strongly diminished levels of H3K27ac after perturbation of CGABC expression highlight the positive role of CGABC in TrxG mediated gene regulation. As CGABC is a zinc finger transcription factor, it can be envisaged that CGABC may mediate the recruitment of dCBP to deposit H3K27ac to promote gene activation.

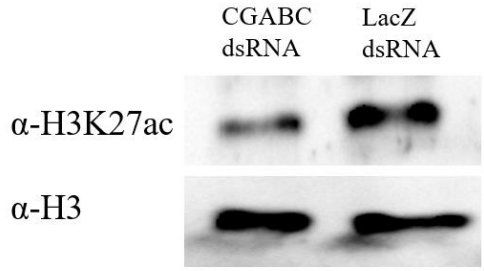


Fig11: As compared to lacZ dsRNA treated cells, drastically reduced H3K27ac levels are observed upon CGABC knockdown. *Drosophila* Dmel2 cells were treated with dsRNA against CGABC while LacZ dsRNA was used as negative control. Antibodies against H3K27ac and histone H3 were used to probe the total H3 levels which were used as loading control.

5. Future Perspectives

In this dissertation, chromatin binding of CGABC has been discovered from the overexpression lines. Moreover, CGABC is shown to positively regulate H3K27ac which is a hallmark of TrxG-mediated gene regulation. In the future, it will be interesting to:

- Determine the global gene expression patterns i.e., genes upregulated/downregulated upon CGABC knockdown in both cells and embryo
- Generate genetic mutant of CGABC and study genetic interactions with PcG/TrxG mutants
- Investigate the effect of CGABC depletion on H3K27me3, H3K9me3, and H3K4me3 marks
- Investigate the effect of CGABC depletion on Trx/Pc recruitment on chromatin
- Conduct ATAQ-seq to determine if CGABC is affecting global chromatin structure
- Determine if CGABC is Suppressor of variegation [Su(var)] or Enhancer of variegation [E(var)] by using white mortaled fly line
- Perform ChIP-seq to observe colocalization with TrxG/PcG on chromatin.

References

Balsalobre, Aurelio, and Jacques Drouin. 2022. “Pioneer Factors as Master Regulators of the Epigenome and Cell Fate.” *Nature Reviews Molecular Cell Biology 2022*, March, 1–16. https://doi.org/10.1038/s41580-022-00464-z.

Bender, M., F. R. Turner, and T. C. Kaufman. 1987. “A Developmental Genetic Analysis of the Gene Regulator of Postbithorax in Drosophila Melanogaster.” *Developmental Biology* 119 (2): 418–32. https://doi.org/10.1016/0012-1606(87)90046-7.

Bird, Adrian P., and Alan P. Wolffe. 1999. “Methylation-Induced Repression--Belts, Braces, and Chromatin.” *Cell* 99 (5): 451–54. https://doi.org/10.1016/S0092-8674(00)81532-9.

Blackledge, Neil P., Nathan R. Rose, and Robert J. Klose. 2015. “Targeting Polycomb Systems to Regulate Gene Expression: Modifications to a Complex Story.” *Nature Reviews Molecular Cell Biology* 16 (11): 643–49. https://doi.org/10.1038/nrm4067.

Brayer, Kathryn J., and David J. Segal. 2008. “Keep Your Fingers off My DNA: Protein-Protein Interactions Mediated by C2H2 Zinc Finger Domains.” *Cell Biochemistry and Biophysics* 50 (3): 111–31. https://doi.org/10.1007/S12013-008-9008-5.

Cao, Ru, Liangjun Wang, Hengbin Wang, Li Xia, Hediye Erdjument-Bromage, Paul Tempst, Richard S. Jones, and Yi Zhang. 2002. “Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing.” *Science* 298 (5595): 1039–43. https://doi.org/10.1126/SCIENCE.1076997/SUPPL_FILE/CAOR1076997.SUP.PDF.

Cassandri, Matteo, Artem Smirnov, Flavia Novelli, Consuelo Pitolli, Massimiliano Agostini, Michal Malewicz, Gerry Melino, and Giuseppe Raschellà. 2017. “Zinc-Finger Proteins in Health and Disease.” *Nature Publishing Group* 3. https://doi.org/10.1038/cddiscovery.2017.71.

Chille, Erin, Emma Strand, Mayaan Neder, Valeria Schmidt, Madeleine Sherman, Tali Mass, and Hollie Putnam. 2021. “Developmental Series of Gene Expression Clarifies Maternal MRNA Provisioning and Maternal-to-Zygotic Transition in a Reef-Building Coral.” *BMC*

Genomics 22 (1). https://doi.org/10.1186/s12864-021-08114-y.

Czermin, Birgit, Raffaella Melfi, Donna McCabe, Volker Seitz, Axel Imhof, and Vincenzo Pirrotta. 2002. "Drosophila Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity That Marks Chromosomal Polycomb Sites." *Cell* 111 (2): 185–96. https://doi.org/10.1016/S0092-8674(02)00975-3.

Deaton, Aimée M., and Adrian Bird. 2011. "CpG Islands and the Regulation of Transcription." *Genes & Development* 25 (10): 1010–22. https://doi.org/10.1101/GAD.2037511.

Déjardin, Jérôme, and Giacomo Cavalli. 2004. "Chromatin Inheritance upon Zeste-Mediated Brahma Recruitment at a Minimal Cellular Memory Module." *The EMBO Journal* 23 (4): 857. https://doi.org/10.1038/SJ.EMBOJ.7600108.

Delcuve, Geneviève P., Mojgan Rastegar, and James R. Davie. 2009. "Epigenetic Control." *Journal of Cellular Physiology* 219 (2): 243–50. https://doi.org/10.1002/JCP.21678.

Fedotova, A. A., A. N. Bonchuk, V. A. Mogila, and P. G. Georgiev. 2017. "C2H2 Zinc Finger Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors." *Acta Naturae* 9 (2): 47. https://doi.org/10.32607/20758251-2017-9-2-47-58.

"FlyBase Gene Report: Dmel\CGABC." n.d. Accessed April 18, 2022. http://flybase.org/reports/FBgn0030008.

Fuchsberger, C, J Flannick, T M Teslovich, A Mahajan, V Agarwala, K J Gaulton, C Ma, et al. 2016. "The Genetic Architecture of Type 2 Diabetes." *Nature* 536 (7614): 41–47. https://doi.org/10.1038/nature18642.

G Struhl, M Akam. 1985. "Altered Distributions of Ultrabithorax Transcripts in Extra Sex Combs Mutant Embryos of Drosophila." *EMBO J*. https://pubmed.ncbi.nlm.nih.gov/2419125/.

Gamsjaeger, Roland, Chu Kong Liew, Fionna E. Loughlin, Merlin Crossley, and Joel P. Mackay. 2007. "Sticky Fingers: Zinc-Fingers as Protein-Recognition Motifs." *Trends in Biochemical Sciences* 32 (2): 63–70. https://doi.org/10.1016/j.tibs.2006.12.007.

Gavis, Elizabeth R., and Ruth Lehmann. 1992. "Localization of Nanos RNA Controls

Embryonic Polarity." *Cell* 71 (2): 301–13. https://doi.org/10.1016/0092-8674(92)90358-J.

Gibney, E. R., and C. M. Nolan. 2010. "Epigenetics and Gene Expression." *Heredity* 105 (1): 4–13. https://doi.org/10.1038/HDY.2010.54.

Grewal, Shiv I.S., and Danesh Moazed. 2003. "Heterochromatin and Epigenetic Control of Gene Expression." *Science (New York, N.Y.)* 301 (5634): 798–802. https://doi.org/10.1126/SCIENCE.1086887.

Grimaud, Charlotte, Nicolas Nègre, and Giacomo Cavalli. 2006. "From Genetics to Epigenetics: The Tale of Polycomb Group and Trithorax Group Genes." *Chromosome Research* 14 (4): 363–75. https://doi.org/10.1007/s10577-006-1069-y.

Ingham, P. W., N. E. Baker, and A. Martinez-Arias. 1988. "Regulation of Segment Polarity Genes in the Drosophila Blastoderm by Fushi Tarazu and Even Skipped." *Nature* 331 (6151): 73–75. https://doi.org/10.1038/331073A0.

Jackle, H., E. Seifert, A. Preiss, and U. B. Rosenberg. 1986. "Probing Gene Activity in Drosophila Embryos." *Journal of Embryology and Experimental Morphology* 97 Suppl (SUPPL.): 157–68. https://doi.org/10.1242/dev.97.supplement.157.

Jenuwein, T., and C. D. Allis. 2001. "Translating the Histone Code." *Science (New York, N.Y.)* 293 (5532): 1074–80. https://doi.org/10.1126/SCIENCE.1063127.

Kassis, Judith A., James A. Kennison, and John W. Tamkun. 2017. "Polycomb and Trithorax Group Genes in Drosophila." *Genetics* 206 (4): 1699–1725. https://doi.org/10.1534/genetics.115.185116.

Kennison, James A. 1995. "The Polycomb and Trithorax Group Proteins of Drosophila: Trans-Regulators of Homeotic Gene Function." *Annual Review of Genetics* 29: 289–303. https://doi.org/10.1146/ANNUREV.GE.29.120195.001445.

Kingston, Robert E., and John W. Tamkun. 2014. "Transcriptional Regulation by Trithorax-Group Proteins." *Cold Spring Harbor Perspectives in Biology* 6 (10). https://doi.org/10.1101/CSHPERSPECT.A019349.

Kornberg, Roger D., and Yahli Lorch. 1999. "Twenty-Five Years of the Nucleosome,

Fundamental Particle of the Eukaryote Chromosome." *Cell* 98 (3): 285–94. https://doi.org/10.1016/S0092-8674(00)81958-3.

———. 2020. "Primary Role of the Nucleosome." *Molecular Cell* 79 (3): 371–75. https://doi.org/10.1016/J.MOLCEL.2020.07.020.

Kuroda, Mitzi I., Hyuckjoon Kang, Sandip De, and Judith A. Kassis. 2020. "Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming." *Annual Review of Biochemistry* 89: 235–53. https://doi.org/10.1146/annurev-biochem-120219-103641.

Lewis, E. B. 1978. "A Gene Complex Controlling Segmentation in Drosophila." *Nature* 276 (5688): 565–70. https://doi.org/10.1038/276565A0.

Ling, C, and T Ronn. 2019. "Epigenetics in Human Obesity and Type 2 Diabetes." *Cell Metab* 29 (5): 1028–44. https://doi.org/10.1016/j.cmet.2019.03.009.

Miller, Trissa, Nevan J. Krogan, Jim Dover, H. Erdjument-Bromage, Paul Tempst, Mark Johnston, Jack F. Greenblatt, and Ali Shilatifard. 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): 12902–7. https://doi.org/10.1073/PNAS.231473398.

Mills, Alea A. 2010. "Throwing the Cancer Switch: Reciprocal Roles of Polycomb and Trithorax Proteins." *Nature Reviews Cancer 2010 10:10* 10 (10): 669–82. https://doi.org/10.1038/nrc2931.

Mlodzik, Marek, and Walter J. Gehring. 1987. "Expression of the Caudal Gene in the Germ Line of Drosophila: Formation of an RNA and Protein Gradient during Early Embryogenesis." *Cell* 48 (3): 465–78. https://doi.org/10.1016/0092-8674(87)90197-8.

Morata, Ginés, Ernesto Sánchez-Herrero, and Jordi Casanova. 1986. "The Bithorax Complex of Drosophila: An Overview." *Cell Differentiation* 18 (2): 67–78. https://doi.org/10.1016/0045-6039(86)90001-1.

Musinova, Yana R., Olga M. Lisitsyna, Sergey A. Golyshev, Alexander I. Tuzhikov, Vladimir Y. Polyakov, and Eugene V. Sheval. 2011. "Nucleolar Localization/Retention Signal Is Responsible for Transient Accumulation of Histone H2B in the Nucleolus through

Electrostatic Interactions." *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1813 (1): 27–38. https://doi.org/10.1016/J.BBAMCR.2010.11.003.

Nusslein-Volhard, C. 1991. "Determination of the Embryonic Axes of Drosophila." *Development* 112 (SUPPL. 1): 1–10. https://doi.org/10.1242/dev.113.supplement_1.1.

Nüsslein-volhard, Christiane, and Eric Wieschaus. 1980. "Mutations Affecting Segment Number and Polarity in Drosophila." *Nature* 287 (5785): 795–801. https://doi.org/10.1038/287795A0.

Papp, Bernadett, and Jürg Müller. 2006. "Histone Trimethylation and the Maintenance of Transcriptional ONand OFF States by TrxG and PcG Proteins." *Genes & Development* 20 (15): 2041–54. https://doi.org/10.1101/GAD.388706.

Petruk, S., Y. Sedkov, S. Smith, S. Tillib, V. Kraevski, T. Nakamura, E. Canaani, C. M. Croce, and A. Mazo. 2001. "Trithorax and DCBP Acting in a Complex to Maintain Expression of a Homeotic Gene." *Science (New York, N.Y.)* 294 (5545): 1331–34. https://doi.org/10.1126/SCIENCE.1065683.

Rg Mü Ller, Jü, Craig M Hart, Nicole J Francis, Marcus L Vargas, Aditya Sengupta, Brigitte Wild, and Ellen L Miller. n.d. "Histone Methyltransferase Activity of a Drosophila Polycomb Group Repressor Complex Development (Francis and Kingston, 2001; Simon and Tamkun, 2002). These Two Sets of Proteins Provide Long-Term Memory of Gene Activity States after Initial on/Off." *Cell* 111.

Ringrose, Leonie, and Renato Paro. 2004. "Epigenetic Regulation of Cellular Memory by the Polycomb and Trithorax Group Proteins." *Annual Review of Genetics* 38: 413–43. https://doi.org/10.1146/annurev.genet.38.072902.091907.

Roguev, Assen, Daniel Schaft, Anna Shevchenko, W. W.M.Pim Pijnappel, Matthias Wilm, Rein Aasland, and A. Francis Stewart. 2001. "The Saccharomyces Cerevisiae Set1 Complex Includes an Ash2 Homologue and Methylates Histone 3 Lysine 4." *The EMBO Journal* 20 (24): 7137–48. https://doi.org/10.1093/EMBOJ/20.24.7137.

Schmittgen, Thomas D., and Kenneth J. Livak. 2008. "Analyzing Real-Time PCR Data by the Comparative C(T) Method." *Nature Protocols* 3 (6): 1101–8.

https://doi.org/10.1038/NPROT.2008.73.

Schuettengruber, Bernd, Henri Marc Bourbon, Luciano Di Croce, and Giacomo Cavalli. 2017. "Genome Regulation by Polycomb and Trithorax: 70 Years and Counting." *Cell* 171 (1): 34–57. https://doi.org/10.1016/j.cell.2017.08.002.

Schuettengruber, Bernd, Daniel Chourrout, Michel Vervoort, Benjamin Leblanc, and Giacomo Cavalli. 2007. "Genome Regulation by Polycomb and Trithorax Proteins." *Cell* 128 (4): 735–45. https://doi.org/10.1016/j.cell.2007.02.009.

Schupbach, Trudi, and Eric Wieschaus. 1986. "Germline Autonomy of Maternal-Effect Mutations Altering the Embryonic Body Pattern of Drosophila." *Developmental Biology* 113 (2): 443–48. https://doi.org/10.1016/0012-1606(86)90179-X.

Schüpbach, Trudi, and Eric Wieschaus. 1986. "Maternal-Effect Mutations Altering the Anterior-Posterior Pattern of the Drosophila Embryo." *Roux's Archives of Developmental Biology : The Official Organ of the EDBO* 195 (5): 302–17. https://doi.org/10.1007/BF00376063.

Schwartz, Yuri B., and Vincenzo Pirrotta. 2007. "Polycomb Silencing Mechanisms and the Management of Genomic Programmes." *Nature Reviews Genetics* 8 (1): 9–22. https://doi.org/10.1038/nrg1981.

Shaheen, Najma, Jawad Akhtar, Zain Umer, Muhammad Haider Farooq Khan, Mahnoor Hussain Bakhtiari, Murtaza Saleem, Amir Faisal, and Muhammad Tariq. 2021. "Polycomb Requires Chaperonin Containing TCP-1 Subunit 7 for Maintaining Gene Silencing in Drosophila." *Frontiers in Cell and Developmental Biology* 9 (October). https://doi.org/10.3389/FCELL.2021.727972.

Stathopoulos, Angelike, and Susan Newcomb. 2020. "Setting up for Gastrulation: D. Melanogaster." *Current Topics in Developmental Biology* 136 (January): 3–32. https://doi.org/10.1016/bs.ctdb.2019.11.004.

Struhl, G. 1989. "Morphogen Gradients and the Control of Body Pattern in Insect Embryos." *Ciba Foundation Symposium* 144. https://doi.org/10.1002/9780470513798.ch5.

Struhl, Gary, and Robert A.H. White. 1985. "Regulation of the Ultrabithorax Gene of Drosophila by Other Bithorax Complex Genes." *Cell* 43 (2 Pt 1): 507–19. https://doi.org/10.1016/0092-

8674(85)90180-1.

Strutt, Helen, Giacomo Cavalli, and Renato Paro. 1997. “Co-Localization of Polycomb Protein and GAGA Factor on Regulatory Elements Responsible for the Maintenance of Homeotic Gene Expression.” *The EMBO Journal* 16 (12): 3621–32. https://doi.org/10.1093/EMBOJ/16.12.3621.

Sun, Ruixia, and Ping Zhu. 2022. “Advances in Measuring DNA Methylation.” *Blood Science (Baltimore, Md.)* 4 (1): 8–15. https://doi.org/10.1097/BS9.0000000000000098.

Swenson, Joel M, Serafin U Colmenares, Amy R Strom, Sylvain V Costes, and Gary H Karpen. 2016. “The Composition and Organization of Drosophila Heterochromatin Are Heterogeneous and Dynamic.” *ELife* 5 (August). https://doi.org/10.7554/ELIFE.16096.

Tanaka, Yujiro, Zen ichiro Katagiri, Koji Kawahashi, Dimitris Kioussis, and Shigetaka Kitajima. 2007. “Trithorax-Group Protein ASH1 Methylates Histone H3 Lysine 36.” *Gene* 397 (1–2): 161–68. https://doi.org/10.1016/J.GENE.2007.04.027.

Tie, Feng, Rakhee Banerjee, Carl A. Stratton, Jayashree Prasad-Sinha, Vincent Stepanik, Andrei Zlobin, Manuel O. Diaz, Peter C. Scacheri, and Peter J. Harte. 2009. “CBP-Mediated Acetylation of Histone H3 Lysine 27 Antagonizes Drosophila Polycomb Silencing.” *Development* 136 (18): 3131–41. https://doi.org/10.1242/DEV.037127.

Umer, Zain, Jawad Akhtar, Muhammad Haider Farooq Khan, Najma Shaheen, Muhammad Abdul Haseeb, Khalida Mazhar, Aziz Mithani, Saima Anwar, and Muhammad Tariq. 2019. “Genome-Wide RNAi Screen in Drosophila Reveals Enok as a Novel Trithorax Group Regulator.” *Epigenetics & Chromatin* 12 (1). https://doi.org/10.1186/S13072-019-0301-X.

Vidal, Miguel. 2019. “Polycomb Assemblies Multitask to Regulate Transcription.” *Epigenomes* 3 (2). https://doi.org/10.3390/epigenomes3020012.

Wang, Charlotte, and Ruth Lehmann. 1991. “Nanos Is the Localized Posterior Determinant in Drosophila.” *Cell* 66 (4): 637–47. https://doi.org/10.1016/0092-8674(91)90110-K.

Wang, Hengbin, Liangjun Wang, Hediye Erdjument-Bromage, Miguel Vidal, Paul Tempst, Richard S. Jones, and Yi Zhang. 2004. “Role of Histone H2A Ubiquitination in Polycomb Silencing.” *Nature 2004 431:7010* 431 (7010): 873–78.

https://doi.org/10.1038/nature02985.

Yusuf, Abdurrahman Pharmacy, Murtala Bello Abubakar, Ibrahim Malami, Kasimu Ghandi Ibrahim, Bilyaminu Abubakar, Muhammad Bashir Bello, Naeem Qusty, et al. 2021. “Zinc Metalloproteins in Epigenetics and Their Crosstalk.” *Life* 11 (3): 1–20. https://doi.org/10.3390/LIFE11030186.

Appendices

Table 1

List of Primers

Primer	Sequence
CGABC Forward Primer	ATG CTC AAG TCC CTC AAG C
CGABC Reverse Primer With Stop	CTA TCC GTT GAG CTC GTC
CGABC Reverse Primer No Stop	TCC GTG AGC TCG TCG TTC
CGABC Real Time Forward	CT GCTCCACGTC CAGAAC
CGABC Real Time Reverse	CCAGGCAAATGGATTCACC
1R-CGABC- DRSC18209	TAATACGACTCACTATAGG TGGACTACAAGGAGGAAGAC
1S-CGABC- DRSC18209	TAATACGACTCACTATAGG TTCGCTGGTGTGCTTGA
2R-CGABC- DRSC24934	TAATACGACTCACTATAGG CATCTTCACCTGCCTGGAAT
2S-CGABC- DRSC24934	TAATACGACTCACTATAGG TGCAGGAACTTGTGGTGGTA

Table 2

Primers used for ChIP analysis

Primer	Sequence
bxd – s – low	GCACTTAAAACGGCCATTACGAA
bxd – s – up	GACGTGCGTAAGAGCGAGATACAG
Dfd – F	AACTCTCCGTGCGAGCGAAC
Dfd – R	ATGCTCCCTCTCAGTCGCGCT
iab7-4-low	AGCTTTTGCCACTCGTCCTGTT
iab7-4-up	AGCAGAGCTGTGCCATTGTTT
pnr_TSS_F	TCTCTTGCTCTTTCGCTCAC
pnr_TSS_R	GTTTTCCATACGCACTCACAC
pnt_TSS_F	TCATTCCAGCGATCAAGTAAAA
pnt_TSS_R	TCTTTCTCTCCGCTGCTAAGAT
psq_TSS_F	ATAAGGCGATGCCACCTAGTTA
psq_TSS_R	AATGTAGCAAAAGGTGCTCAAAG
Intergenic Region F	CCGAACATGAGACATGGA AAA
Intergenic Region R	AAAGTGCCGACAATGCAGT TA

Appendix 1

TURBO DNAse treatment

Reagents	Volume
TURBO DNAse	1 µl
10X TURBO DNAse Buffer	2 µl
RNA	Up to 9 µg
DEPC water	Up to 20 µl

The reaction was set up as above on ice and incubated at 37ºC for 50 minutes in a PCR machine. After that, 2 µl of slurry was added to inactivate the enzyme and the mixture was incubated at room temperature for 2 minutes with gentle flicking after every 15 seconds to avoid settling of the slurry. The mixture was centrifuged at 13000 rpm for 2 minutes to settle down the slurry and the supernatant was labelled as DNAse treated RNA (DTR) and was transferred to another tube.

Appendix 2

cDNA synthesis

Reagent	Volume
RNA	2 µg
Random Hexamers	2 µl
10mM dNTPs	2 µl
DEPC water	Up to 20 µl

The reaction was incubated at 65ºC in the PCR machine for 5 minutes followed by immediate incubation on ice for at least one minute. cDNA mixture was prepared as a master mix (n+1) for the below recipe.

Reagent	RT + Volume	RT – Volume
10 RT Buffer	2 µl	2 µl
25mM $MgCl_2$	4 µl	4 µl
0.1M DTT	2 µl	2 µl
RNAse OUT	1 µl	1 µl
Superscript III RT	1 µl	-

10 µl of RNA/primer mixture was added to each RT+ and RT- vial. The mixture was gently flicked and short spined. The reaction was set up in the PCR machine for the following profile:

25ºC	**10 min**
50ºC	**50 min**
85ºC	**5 min**

The quality of cDNA was validated via end point PCR using actin primers. The following reagents were used to set up the reaction:

Reagents	Volumes
PCR water	Up to 20 µl
10X PCR Buffer	2 µl
10Mm dNTPs	1 µl
Forward primer (10µM)	0.5 µl
Reverse primer (10µM)	0.5 µl
Taq Polymerase	0.25 µl
Template	50 – 100ng

The following profile was set up:

Temperature	Time	Cycles
95ºC	3 min	-
95ºC	30 sec	35 cycles
60ºC	30 sec	
72ºC	30 sec	
72ºC	10 min	-
4ºC	∞	-

The PCR amplified product were run on a 1% agarose gel made in 1X Tris Acetate EDTA (TAE) buffer. The electrophoresis was conducted at 100V for 30 minutes after which the gel was analyzed under UV Gel Doc (SYNGENE). A single band in RT+ and no band in RT- and PCR- showed successful synthesis of cDNA which was then used for subsequent experiments.

Appendix 3

dsRNA synthesis

The MEGAscript kit (Thermo Scientific) was used for the In vitro transcription (IVT) reaction. All reagents were brought to room temperature to thaw them completely, only the enzyme was placed on ice. The reaction was set up at room temperature. Important: don't use ice cold 10X reaction buffer; it will precipitate the template DNA and reaction will be unsuccessful. The reaction was set up in the following order:

Reagents	Volume
ATP	5 µl
CTP	5 µl
GTP	5 µl
UTP	5 µl
10X Reaction Buffer	5 µl
PCR Amplified Linear DNA Template	2 µg
Nuclease-free water	Up to 50 µl

Enzyme mix	5 µl

The reaction was mixed gently with the help of a pipette, followed by brief centrifugation to collect the contents of the mixture. The reaction was incubated overnight at 37ºC in a thermocycler. The next day, 1 µl of TURBO DNAse were added to the mixture to remove the template DNA and the mixture was incubated at 37ºC for 20 minutes. The reaction was stopped by adding 15 µl of Ammonium Sulphate stop solution along with 115 µl of nuclease free water followed by thorough mixing. The dsRNA was extracted by using an equal volume of phenol: chloroform followed by centrifugation at 13000 rpm for 15 minutes at 4ºC to separate the organic and aqueous layers. The aqueous layer was extracted again using an equal volume of chloroform. The final aqueous layer was transferred to a fresh tube and the RNA was precipitated by adding 1 volume of Isopropanol. The reaction mixture was incubated at -80ºC overnight followed by centrifugation at max speed for 30 minutes at 4ºC to pellet down the RNA. The supernatant was carefully removed, and the RNA pellet was washed with 70% ethanol followed by centrifugation at max speed for 5 minutes at 4ºC. The final RNA pellet was air dried for 5 minutes and dissolved in DEPC treated water in such a way that the final concentration was 1 µg/ µl. Aliquots of 20 µl were made and stored at -80ºC.

Appendix 4

Plasmid Preparation Solution (Miniprep)

P1 – Resuspension Buffer

Reagents	Concentrations
Tris-Cl	50mM
EDTA	10mM
RNAse A	100 µg / ul

The pH of the buffer was set at 8 with the help of concentrated HCl and the buffer was autoclaved. Once it had cooled down, RNAse A was added and the buffer was stored at 4ºC.

P2 – Lysis Buffer

Reagents	Concentrations
NaOH	200mM
SDS	1% w/v

This buffer should not be autoclaved and should be kept at room temperature.

P3 – Neutralization Buffer

Reagents	Concentrations
Potassium Acetate	3M

The pH of this buffer was set at 5.5 using glacial acetic acid, followed by autoclaving. After this the buffer was stored at 4ºC.

Plasmid Isolation

- Single colony from the LB agar plate containing the appropriate antibiotic was taken and inoculated into 5ml of LB broth in a test tube along with the appropriate antibiotic.
- The culture was allowed to grow overnight (16-18 hrs) at 37ºC at 200 rpm. The next morning, the bacterial cells were harvested at 8000 rpm for 5 minutes at room temperature.
- The pellet was resuspended in 300µl of P1 while keeping the tubes on ice. 300 µl of P2 were added and the tube was mixed thoroughly by inverting 4-6 times (don't vortex) after which the mixture was incubated at room temperature for 4-5 minutes. Never give lysis reaction to proceed for more than 5 minutes!
- 300 µl of P3 were added and the contents were thoroughly mixed by inverting until a homogenous mixture forms. The tube was incubated on ice for 2 minutes.

- The tube was centrifuged at 13000 for 15 minutes and the supernatant was collected in a fresh tube.
- 1 volume of isopropanol was added to the supernatant collected and mixed vigorously to precipitate the plasmid DNA.
- The precipitated DNA was pellet down by centrifugation at max speed for 30 minutes at 4ºC.
- The plasmid pellet was washed using 70% ethanol made in PCR water followed by centrifugation at max speed for 5 minutes at 4ºC.
- The purified pellet was air dried for 5 – 10 minutes and dissolved in 20 µl of PCR water and was stored at -20ºC.

Appendix 5

Western Blot Solutions

Protein SDS loading dye

Reagents	Concentration
100% Glycerol	40%
Bromophenol Blue	0.04%
SDS	6%
Tris-Cl (pH 6.8)	250mM
Milli-Q	Up to 50ml
Add 5% β-mercaptoethanol before use	

1.5M Tris pH 8.8 for resolving gel

Reagents	Amount (for V = 500ml)
Tris base	90.855g
Adjust pH to 8.8 using concentrated HCl	

1M Tris Ph 6.8 for stacking gel

Reagents	Amount (for V = 500ml)
Tris-base	60.57g
Adjust pH to 6.8 using concentrated HCl	

Resolving Gel (12%)

Reagents	Volume (20ml total)
Milli Q	6.6ml
30% Acrylamide solution	8ml
1.5M Tris (pH 8.8)	5ml
10% SDS	0.2ml
10% Ammonium Per Sulphate (APS)	0.2ml
TEMED	0.008ml
Mix by gently inverting	

Stacking Gel (5%)

Reagents	Volume (5ml total)
Milli Q	3.4ml
30% Acrylamide solution	0.83ml
1.5M Tris (pH 8.8)	0.63ml
10% SDS	0.05ml
10% Ammonium Per Sulphate (APS)	0.05ml
TEMED	0.005ml
Mix by gently inverting	

Tris Glycine Running Buffer (5X)

Reagents	Amount (1L total)
Tris-base	15.1g
Glycine	72g

SDS	5g

Transfer Buffer (10X)

Reagents	Amount (1L total)
Glycine	144g
Tris-base	30.25g

1X Transfer Buffer

Reagents	Volume (1L total)
10X Transfer Buffer	100ml
Methanol	200ml
Distilled water	700ml

1X PBST

Reagents	Volume (1L total)
1X PBS	999ml
Tween-20	1ml
Stir well using magnetic stirrer	

Blocking Buffer

Reagents	Amount (100ml total)
Skimmed milk	5g
1X PBST	95ml

Appendix 6

ChIP Solutions

10X Fixative Solution

Reagents	Amount	Final Concentraion
5M NaCl	40ml	1M
1M Tris (pH 8)	10ml	50mM
0.5M EDTA	4ml	10mM
0.5M EGTA	2ml	5mM
Make final volume to 200ml		

1X Fixative Solution

Reagents	Volume (5ml total)
10X Fixative Buffer	3.55ml
37% formaldehyde	1.45ml

2.5M Glycine Solution

Reagents	Amount (50ml total)
Glycine	9.38g
Make final volume to 50ml	

Protease Inhibitors

Reagents	Stock Concentration	Final Concentration
PMSF	100mM	1mM
Pepstatin	1µg/µl	2µg/ml
Leupeptin	1µg/µl	2µg/ml
Aprotinin	1µg/µl	2µg/ml

Solution A

Reagents	Volume (200ml total)	Final Concentration
1M Tris (pH 8)	2ml	10mM
0.5M EDTA	4ml	10mM
0.5M EGTA	200µl	0.5mM
Triton X-100	500µl	0.25%
	Make final volume to 200ml	

Solution B

Reagents	Volume (200ml total)	Final Concentration
1M Tris (pH 8)	2ml	10mM
0.5M EDTA	400µl	1mM
0.5M EGTA	200µl	0.5mM
5M NaCl	8ml	200mM
	Make final volume to 200ml	

Sonication Buffer

Reagents	Volume (200ml total)	Final Concentration
1M Tris (pH 8)	2ml	10mM
0.5M EDTA	400µl	1mM
0.5M EGTA	200µl	0.5mM
	Make final volume to 200ml	

2X RIPA Buffer

Reagents	Amount (200ml total)	Final Concentration
IM Tris (pH 8.0)	4ml	20mM
0.5M EDTA	800µl	2mM
5M NaCl	11.2ml	280mM
Triton X-100	4ml	2%
10% SDS	4ml	0.2%
10% Sodium deoxycholate	4ml	0.2%
Make final volume to 200ml. For 1X RIPA use 1:1 ratio of 2X RIPA and Autoclaved MilliQ. Add the protease inhibitors in the concentrations mentioned above.		

Lithium Chloride Buffer (1X)

Reagents	Amount (100ml total)	Final Concentration
1M Tris (pH 8.0)	1ml	10mM
Lithium Chloride	1.06g	250mM
0.5M EDTA	200µl	1mM
NP-40 (70%)	714µl	0.5%
10% Sodium deoxycholate	5ml	0.5%
Make final volume to 100ml. Add the protease inhibitors in the concentrations mentioned above.		

TE Buffer (1X)

Reagents	Volume (40ml total)	Final Concentration
1M Tris (pH 8)	400µl	10mM
0.5M EDTA	80µl	1mM
Make final volume to 40ml. Add the protease inhibitors in the concentrations mentioned above.		

Elution Buffer (1X)

Reagents	Volume (10ml total)	Final Concentration
Sodium Bicarbonate ($NaHCO_3$)	0.084g	0.1M
10% SDS	1ml	1%
Make final volume to 10ml. Add the protease inhibitors in the concentrations mentioned above.		

Sodium Acetate Solution

Reagents	Volume (50ml total)	Final Concentration
Sodium Acetate (CH_3COONa)	20.4g	3M
Make final volume to 50ml. Add the protease inhibitors in the concentrations mentioned above.		