

Planarian flatworms as a new model system for understanding epigenetic regulation of stem cell pluripotency and differentiation

Anish Dattani*, Divya Sridhar and Aziz Aboobaker*

Department of Zoology, South Parks Road, University of Oxford, OX1 3PS, UK

* corresponding authors

Keywords:

Neoblast

Regeneration

ChIP-seq

DNA methylation

Set domain-containing genes

NuRD complex

Polycomb repressive complex

PIWI-piRNA complex

Abstract:

Planarian flatworms possess pluripotent stem cells (neoblasts) that are able to differentiate into all cell types that constitute the adult body plan. Consequently, planarians possess remarkable regenerative capabilities. Transcriptomic studies have revealed that gene expression is coordinated to maintain neoblast pluripotency, and ensure correct lineage specification during differentiation. But as yet they have not revealed how this regulation of expression is controlled. In this review, we propose that planarians represent a unique and effective system to study the epigenetic regulation of these processes in an *in vivo* context. We consolidate evidence suggesting that although DNA methylation is likely present in some flatworm lineages, it does not regulate neoblast function in *Schmidtea mediterranea*. A number of phenotypic studies have documented the role of histone modification and chromatin remodelling complexes in regulating distinct neoblast processes, and we focus on four important examples of planarian epigenetic regulators: Nucleosome Remodeling Deacetylase (NuRD) complex, Polycomb Repressive Complex (PRC), the SET1/MLL methyltransferases, and the nuclear PIWI/piRNA complex. Given the recent advent of ChIP-seq in planarians, we propose future avenues of research that will identify the genomic targets of these complexes allowing for a clearer picture of how neoblast processes are coordinated at the epigenetic level. These insights into neoblast biology

may be directly relevant to mammalian stem cells and disease. The unique biology of planarians will also allow us to investigate how extracellular signals feed into epigenetic regulatory networks to govern concerted neoblast responses during regenerative polarity, tissue patterning, and remodelling.

Introduction:

During development in mammals, the precise coordination of transcription ensures the transition of early embryonic pluripotent stem cells and their progeny into a vast array of cell types that constitute the entire adult organism. The control and maintenance of gene expression during ongoing lineage commitment and differentiation is also dependent on pervasive epigenetic control – the heritable modulation of gene activity that is independent of the underlying DNA sequence. Epigenetic modifications can be broadly classified into 3 groups: (1) DNA methylation, (2) histone modification and nucleosome positioning (3) small RNA mediated transgenerational inheritance (not discussed further in this paper, but reviewed in [1,2]).

The methylation of cytosine at the C5 position in CpG di-nucleotide islands, usually located upstream of gene promoters, acts as a beacon to attract epigenetic machinery involved in regulating gene repression or often through chromatin remodelling to robustly block transcriptional machinery from accessing these sites altogether [3]. The DNA methylation process has been mainly studied in vertebrates and plants owing to a ubiquitous presence in these taxa. DNA methylation has a patchier distribution among invertebrates, for example being absent altogether in *Caenorhabditis elegans* and still contentious in *Drosophila melanogaster* [4–7]. Relatively little is known about whether and how it is involved in regulating gene expression across the breadth of the Animal Kingdom or how this has evolved.

The presence of covalent modification of the four nucleosome core histone proteins (H2A, H2B, H3, and H4), around which DNA is wrapped, is in contrast to DNA methylation ubiquitous across eukaryotes [8]. These nucleosomes are organised in a higher order chromatin structure through Histone 1 (H1) linker proteins between nucleosomes, and each genetic locus of expression can have its own unique structure and set of modifications that can vary over time between cell states and types, optimised to the transcriptional need for that gene. Modifications to histone proteins include acetylation, methylation, phosphorylation and ubiquitination. These either directly affect the chromatin structure by affecting DNA/nucleosome interaction or act by attracting effector complexes that contain modification-specific binding domains. These complexes act on chromatin in various ways to influence whether genes are upregulated, downregulated or silenced [9–11]. One general feature of the epigenome, that can be

exploited experimentally by DNA sequencing library technology, is that open chromatin correlates with actively transcribed genes and active enhancers, and in some ways can be seen as the interface between transcriptional and epigenetic regulation. Thus, an important role of some histone modifications is to recruit chromatin remodelling complexes that reposition nucleosomes, expel nucleosomes entirely, or exchange histone variants, thereby affecting the accessibility of the transcriptional machinery to gene promoters and regulatory enhancers involved in providing temporal and spatial specificity.

The bulk of our knowledge on the epigenetic regulation of stem cell pluripotency and differentiation comes from work using Embryonic Stem Cells (ESCs). ESCs are present in the inner cell mass (ICM) of the blastocyst stage embryo and can produce progenitors that contribute to any type of adult tissue. In culture they retain an indefinite capacity for self-renewal when differentiation is inhibited by a variety of media conditions that mimic aspects of the microenvironment of the ICM. These conditions include culture with cytokine leukaemia inhibitory factor in the presence of serum (serum/LIF) [12,13], in serum-free medium with 2 small molecule inhibitors (2i/LIF)[14], or with knockout serum replacement (KOSR/LIF)[15]. Different culture conditions give rise to different ESCs, 2i/LIF-grown ESCs are unrestricted and highly plastic and reflective of the early blastocyst, whereas serum-grown ESCs are more heterogenous and restricted in their cell potential reflective of the late blastocyst [15,16]. However, it has not been demonstrated whether the transcriptomic and epigenetic states of 2i/LIF cultured ESCs are stable over long-term culture. A recent study has shown that 2i/LIF ESCs lose DNA methylation at imprinted loci, which leads to an impaired developmental potential and karyotypic abnormality [17,18]. The other burgeoning stem cell study system are induced pluripotent stem cells (iPSCs), that are been reprogrammed to a pluripotent state from somatic cells usually by over expression of the key Yamanaka transcription factors [19]. This leads to broad epigenetic changes that are heterogenous between reprogramming events and often somatic epigenetic remodelling is incomplete [20,21]. Both ESCs and iPSCs represent *in vitro* study systems and it is likely that significant regulatory differences exist between these systems and stem cells in their *in vivo* contexts. Given the importance of understanding pluripotency for biomedical research it is very surprising that other animals models where pluripotent adult somatic cells are present are still relatively poorly supported. These models are simpler, accessible and can allow study of pluripotent cells *in vivo*. At the very least it is likely that a comparative

study of epigenetic control mechanisms will be broadly informative, regardless if mechanisms are conserved or divergent.

Here we propose that the pluripotent adult stem cells, called neoblasts (NBs), of the regenerative planarian flatworm *Schmidtea mediterranea* can be used to study the epigenetic regulation of stem cells in an *in vivo* context, representing a useful non-mammalian system. Early experiments in this model system have shown that knockdown of orthologs of mammalian epigenetic regulators by RNA interference (RNAi) can lead to different stem cell defects and errors in lineage commitment of stem cell progeny, culminating in a loss of regenerative capacity. The recent advent of Chromatin Immunoprecipitation Sequencing (ChIP-seq) in planarians will allow for the investigation of these defects in greater detail [22–24], enable assessment of the conservation of epigenetic programs and potentially identify important functions and targets of epigenetic complexes that may have been either overlooked or difficult to study in mammalian ESC or iPSC culture based systems. In this review, we consolidate the planarian studies of epigenetic regulators so far. We describe and synthesize our understanding of the phenotypic defects of RNAi of genes involved in epigenetic complexes and propose avenues of exploration to understand how planarian NBs respond to extracellular signals to coordinate differentiation under homeostatic and regenerative conditions. We point the reader to a number of reviews that provide detailed introductions to the planarian system as a regenerative and stem cell model, which we do not revisit in detail here [25–27].

2. The case for Planarian flatworms as an *in vivo* model system to study stem cell epigenetics

2.1 *Transcriptional profiling of planarian somatic neoblasts reveals similarity with ESCs*

Planarian flatworms represent a good system for characterizing the epigenome of stem cells. These animals contain a population of self-renewing pluripotent adult stem cells, called neoblasts (NBs), which divide to replace any cell type during homeostasis or following injury [25,26,27]. Importantly, there is some evidence that planarian NBs have an underlying pluripotency program that is conserved with mammalian ESCs as well as the pluri- and multipotent adult stem cells and ESCs of other animals. Independent studies and methods have uncovered that genes with enriched expression in NBs are

enriched with genes known to be expressed in ESCs and involved in the balance between self-renewal and differentiation such as regulators and targets of Oct4, RNA splicing factors, epigenetic modifiers, and RNA binding proteins [28–30]. The planarian NB transcriptome broadly also reflects that of the multipotent stem cells of Hydra and totipotent archeocytes of the demosponge *Ephydatia fluviatilis*, together suggesting the existence of an ancestral stem cell expression repertoire, rich in RNA regulatory actors and poor in transcription factors [31,32]. While more work is required to assess the extent and nature of this conservation, these early findings lend credence to the use of planarian NBs as a model system for stem cell epigenetic studies, as discoveries in planarians may be directly relevant to mammalian ESC biology.

2.2 Planarians provide the opportunity to study the role of epigenetic mechanisms in controlling stem cells in vivo

One main advantage planarians have over ESCs is that they represent an *in vivo* system whereby the epigenetic response to extracellular signals can be explicitly tested. Planarian studies have identified many key signalling pathways that regulate regenerative polarity [33–41], found where these signals originate from [40,42–44], and in some cases the transcriptional changes they control in NBs responding correctly to injury [42,45,46]. We have also made progress in understanding how the dynamic process of tissue homeostasis in planarians is controlled [47]. So far, however, we do not know the role of epigenetic mechanisms in regulating regenerative polarity, tissue patterning or tissue homeostasis as most studies to date have investigated epigenetics only in the context of stem cell maintenance and differentiation. However, with the application of epigenomic techniques, such as ChIP-seq and in the future Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq), these processes can be studied in the context of the whole regenerative response. We can ask how the epigenome of NBs responds to different signals and conditions and how this then impacts on changes in gene expression that allow regeneration of the correct structures. For example, we know that NBs respond very differently to anterior and posterior facing wounds that produce different positional signals [48]. They also behave differently in response to starvation conditions resulting in regulated de-growth [49,50]. These represent fundamental *in vivo* stem cell responses for which

epigenetic responses are not yet described. By combining specific regenerative or environmental scenarios with RNAi of key signalling pathways (e.g. Wnt [36,37,40,51], Hh [37,38], TOR [52], JNK [53,54]) it should be possible to assess the importance of epigenetic mechanisms in the NB response to these conditions and control signals. In summary, while all current examples of using ChIP-seq for histone marks in planarians have been in the context of studying enzymes that are mediators of histone marks, these studies have established the basis for a much broader investigation in all the exciting experimental paradigms offered by planarians.

3. Loss of DNA methylation in invertebrates and flatworm lineages

3.1 The patchy distribution of DNA methylation in invertebrates.

DNA methylation, the transfer of a methyl group to the cytosine ring of DNA (5mC), typically occurs in the context of CpG dinucleotides, and is responsible for the silencing of the underlying DNA segment. DNA methylation-based silencing is responsible for both long-term transposable element (TE) silencing, preventing these selfish genetic elements from disrupting genomic integrity, and genomic imprinting, which allows for monoallelic expression of a subset genes dependent on parental origin [55]. DNMT1 and DNMT3 are the two generally accepted families of DNA methyltransferases ancestral to both animals and plants [56,57]. DNMT2 has been dismissed as a misnomer as it has a role in tRNA methylation [58,59], although earlier studies suggested it had a role in retro-element silencing [60,61]. 5mC is a substrate for methyl-binding domain containing proteins (MBDs) that attract nucleosome remodelling and histone modification complexes to the DNA segment. The ancestral group of MBD genes in animals is MBD2/3 (or less commonly, but correctly referred to as MBD1/2/3) and MBD4/MeCP2 which following 2 rounds of whole genome duplication (2R) resulted in the paralogs MB1, MBD2, MBD3, MBD4 and MeCP2 in vertebrates [62,63].

The overall pattern of DNA methylation is variable between organisms. Vertebrates, and in particular mammalian genomes, are heavily methylated at most CpG sites (i.e. global DNA methylation) which correlates with transcriptional silencing. Only a small number of CpGs are lowly methylated, localized to short genomic regions, and usually located in proximity to transcribed gene promoters and

enhancers. These hypomethylated regions are typically characterized by a high GC content and are referred to as CpG islands [64,65]. Conversely, invertebrate genomes tend to be sparsely methylated and in many cases DNA methylation and the machinery for producing this mark are absent from the genomes entirely (**Figure 1**) [4]. Loss of methylation has happened independently within many groups, including nematodes, arthropods and flatworms, which have members with and without DNA methylation and/or DNA methyltransferases. Although the ultimate reason as to why many invertebrates have lost DNA methylation is unclear [57,66], the proximate cause is most likely associated with the mutagenic load associated with DNA methylation (such as GT mis-match from 5mC to thymine deamination [67–70] and alkylation damage [71]).

3.2 Planarians most likely lack endogenous DNA methylation as they have no cognate DNA methyltransferases

DNA methylation studies in flatworms have been controversial and contradictory. Given that NB-like cells are a conserved feature of the phylum Platyhelminthes, it is important to know whether the DNA methylation plays a role in epigenetic regulation of these cells. Early studies argued that DNA methylation was not present in the phylum Platyhelminthes on the basis of methylation-based restriction endonucleases followed by amplification of restriction fragments (Methylation Sensitive Amplified Polymorphism – MSAP) [72]. While DNMT1 and DNMT3 are absent in specific flatworm lineages, including the four classes Trematoda, Cestoda, Turbellaria, and Monogenea, it was argued that having DNMT2 and MBD2/3 orthologs could be indicative of a propensity for DNA methylation in these organisms [73]. However DNMT2 is a tRNA methyltransferase with no convincing role in DNA methylation having ever been established [59,61,74]. One study with the parasitic flatworm *Schistosoma mansoni* claimed the presence of methylated DNA in precise locations in the genome and a role for cytosine methylation in the regulation of oviposition. This work utilized 5-azacytidine (AzaC) to inhibit DNA methylation in adult mating pairs, but it should be noted this drug also inhibits RNA methylation by DNMT-2 with high efficiency and as such the phenotypic effects could reflect inhibition of this process [75,76]. Indeed, another study utilizing whole-genome bisulfite sequencing showed that the *S. mansoni* genome was not methylated and that incompletely converted cytosines following bisulfite treatment likely accounted for why DNA methylation was found in an earlier study [74]. It

therefore seems likely that the *S. mansoni* genome is not endogenously methylated, but perhaps previous positive data actually reflects methylated cytosine scavenged from the host or culture environment [77,78]. In contrast to parasitic genomes, the genome of the more basal flatworm *Macrostomum lignano* has been shown to have both DNMT1 and DNMT3 and low levels of DNA methylation [79,80].

The genome of *S. mediterranea* was definitively shown to lack cytosine-dependent methylation on the combined basis of MSAP, a lack of antibody staining against 5mC, and undetectable levels of 5mC in High Performance Liquid Mass Chromatography coupled Mass Spectrometry (HPLC-MS) [81]. Moreover, neither DNMT1 nor DNMT3 have been found in the genome of *S. mediterranea*, and like other closely related Platyhelminthes it does not contain an MBD4/MeCP2 (**Figure 1**). Additionally, the MBD2/3 protein in *S. mediterranea* does not contain the highly conserved ARG22 involved in forming hydrogen bonds with guanine in methylated CpG islands [81–83]. As a consequence of these different lines of evidence we can suggest that the function of *Smed-mbd2/3* is independent of DNA methylation and that DNA methylation is not involved in the epigenetic control of stem cells in *S. mediterranea*. However, given the presence of a complete set of machinery in *M. lignano* it is possible that DNA methylation is present and may have been involved in epigenetic regulation of stem cells in this species. Studies in *M. lignano*, another pertinent regenerative model [80], will address this.

4. Histone modifications and knockdown phenotypes in *Schmidtea mediterranea*

4.1 Connecting histone modification changes at genes with whole organismal phenotypes

The phenotypic effects of knockdown of histone epigenetic complexes can be effectively assessed during planarian regeneration, and differentiation defects can be assayed with *in situ* hybridization using a growing list of markers specific to different lineages. With the advent of ChIP-seq on planarian NBs we are now able to connect whole-organismal phenotypic effects with epigenetic changes to specific genes. The following section reviews studies that have adopted this holistic approach in studying the role of major histone and nucleosome modifying complexes in planarian stem cell biology and differentiation.

4.2 The multipronged NuRD complex and the role of MBD2/3 in 5mC-free planarians

Studies from different animals have shown that the nucleosome remodelling and deacetylase (NuRD) complex is essential for embryonic development. The NuRD protein complex is relatively unique in that it has at least three distinct enzymatic activities involved in chromatin directed gene regulation: deacetylation, ATP-dependent chromatin remodelling, and lysine-specific demethylation. A lack of DNA methylation in planarians means that genomic targets and biological effects of conserved histone modifiers and chromatin remodelers in NBs can be studied without consideration of an interplay with DNA methylation. This simplification in comparison to mammals may be particularly useful for studying the methylation independent roles of the NuRD complex.

Firstly, NuRD has histone deacetylase activity through the activities of the HDAC subunit, which is highly conserved and present in all eukaryotes. HDAC activity leads to a loss of the active H3K27ace on specific genes and provides a substrate for PCR2-mediated tri-methylation (H3K27me3) leading to transcriptional silencing [84]. NuRD targets in ESCs include developmental genes that are transcriptionally 'poised' or bivalent genes that harbour both active H3K4me3 and repressive H3K27me3 marks [84]. This suggests that HDAC/NuRD activity regulates the balance between the acetylation and methylation state of H3K27 such that genes can be released for transcription upon ESC differentiation to defined lineages. Other NuRD targets in ESCs, somewhat counterintuitively, include pluripotency associated genes. The activity of histone acetyltransferases (HATs), which promote the transcription of pluripotency genes in ESCs, is dampened by the HDAC activity of NuRD. Once ESCs differentiate, HAT activity diminishes, and pluripotency genes are silenced in differentiating cells [85]. Secondly, NuRD is also implicated in ATP-dependent chromatin remodelling as a result of the mutually-exclusive chromodomain-helicase-DNA-binding paralogous subunits CHD3 (Mi-2 α), CHD4 (Mi-2 β) and CHD5. These subunits utilize the energy released from the hydrolysis of ATP to ADP to induce nucleosome sliding, which either enables the recruitment of transcriptional complexes or suppresses transcription entirely. Early studies in *Arabidopsis thaliana* [86] and *C. elegans* [87,88] indicated that the CHD subunits are involved in the silencing of embryonic genes during differentiation. Discoveries in mammalian systems have since shown that CHD, as part of the NuRD complex, also functions in

guiding lineage-specific gene programs. For example, in mammals, the CHD3, CHD4, and CHD5 proteins regulate distinct and non-redundant aspects of gene regulation in three distinct stages of cortical differentiation [89].

NuRD also associates with the lysine-specific histone demethylase 1A (LSD1) to target the removal of active mono and di-methyl moieties from lysine 4 of histone 3 (H3K4) [90]. NuRD complexes containing LSD1 associate with the promoters of genes involved in cell growth (including TGF β signaling), survival, migration and tissue invasion. Indeed, LSD1-NuRD complexes prevented breast cancer invasion *in vitro* and metastases *in vivo*, indicating that the loss of the LSD1-NuRD complex or reduction in activity may predispose to cancer [90]. One hypothesis is that LSD-1 targets H3K4me2 removal at promoters leading to gene silencing [91]. Moreover, LSD1-NuRD complexes localize to active ESC enhancers to decommission them via removal of the H3K4me1 active mark, resulting in increased differentiation [92,93].

In addition to these three enzymatic subunits, in vertebrates, the NuRD complex also associates with two interchangeable methyl-CpG-binding domain (MBD) proteins, MBD2 and MBD3. MBD2 has the capacity to selectively recognize 5mC, whilst MBD3 has lost the ability to bind to 5mC during vertebrate evolution [94,95]. Whilst earlier studies suggested that MBD3/NuRD had a role independent of DNA methylation, MBD3 can bind to 5-hydroxymethylcytosine (5hmC) [96,97] - the first oxidative product in the demethylation of 5mC by the enzyme TET1 [98,99]. A recent study proposed that MBD2/NuRD and MBD3/NuRD bind to the same genomic loci, and suggests a model by which the two MBD proteins are interdependent and form a regulatory loop to reinforce transcriptional silencing (**Figure 2A**) [100]: (1) following the conversion of 5mC to 5hmC by TET1, (2) MBD3/NuRD binds to 5hmC loci leading to DNMT1 localization (3) enabling conversion of 5hmC back 5mC (4) leading to subsequent binding of MBD2. Occupation of MBD3/NuRD at 5hmC can be disrupted by further TET1 activity leading to CpG demethylation, which can precede transcription activator binding and gene expression. Importantly, MBD proteins function in coordinating crosstalk between DNA methylation and NuRD to produce a suppressive chromatin environment at target loci [101].

Some reports suggest that MBD2/NuRD and MBD3/NuRD may function independently of CpG methylation in mammalian stem cell systems and have a role in transcriptional activation of genes and

enhancers [102–105]. However, data from these studies have since been re-analyzed and no evidence for MBD2 and MBD3 methylation-independent functions are supported [100].

Most invertebrates contain an ancestral MBD2/3 gene that following 2R resulted in one MBD2 gene and one or two copies of the MBD3 gene in vertebrates. [106]. Nuclear Magnetic Resonance (NMR) showed that the MBD2/3 protein of the sponge *Ephydatia muelleri*, a basal metazoan, can bind to methylated DNA consistent with the presence of DNA methylation in this species [107]. However, the MBD2/3 protein of *Drosophila melanogaster* lacks DNA binding activity, but continues to associate with the NuRD complex [107]. Consequently, we can posit that the ancestral MBD2/3 did bind 5mC or 5mHc and this activity has been lost secondarily in some non-methylated invertebrate species. MBD2/3 in these cases may not necessarily require DNA methylation as a genomic reference to recruit the NuRD complex to target loci, and most-likely has a DNA-methylation independent role.

Like *Drosophila*, *Schmidtea mediterranea* also has no detectable levels of endogenous cytosine methylation. RNAi of *Smed-mbd2/3* resulted in a loss of certain differentiated cell lineages (e.g. epidermis, gut and pharynx) without reducing NB number. Moreover, there was an accumulation of early epidermal NB progeny (*prog-1+*) but a reduction in late progeny (*agat-1+*). Given that *Smed-mbd2/3* mRNA is restricted to the stem cell (X1) and stem cell progeny compartments (X2), it is likely that *Smed-MBD2/3* protein influences the expression of genes involved in the terminal differentiation program. However, it remains to be addressed whether *Smed-mbd2/3* has a role as a part of the planarian NuRD complex. If MBD2/3 has an ancestral role in coordinating NuRD activity independently of methyl-binding, this mechanism may also function in mammals and would help to resolve long-standing disputes over whether MBD can function independently of CpG [97,100,102,103]. It is possible that *Smed-MBD2/3* either directly or indirectly (via binding of an unknown DNA-binding gene) associates with pluripotency related genes and/or differentiation-related genes and recruits NuRD to modulate their transcription in NBs and NB post-mitotic progeny (**Figure 2B**).

The functions of four other NuRD complex components have also been investigated by RNAi knockdown in planarians: *Smed-CHD4* [108], *Smed-HDAC1* [109–111], the nucleosome interactor RbAp48 [112,113], and the GATA-type zinc-finger domain-containing gene *p66* [114]. Similar to the *Smed-mbd2/3* RNAi phenotype, all of these genes lead to an abrogation of stem cell differentiation, but do so in differing manners. For instance, *Smed-CHD4* and *Smed-HDAC1* RNAi animals lost both early

epidermal progeny (*prog-1+*) and late epidermal progeny (AGAT-1+) cells whereas the *Smed-mbd2/3* RNAi phenotype lead to an accumulation of *prog-1* cells but a reduction in *agat-1* cells [81,110,111]. This indicates that these genes have distinct functions in the lineage differentiation process – *Smed-mbd2/3* is required for later point in differentiation, whereas *Smed-CHD4* and *Smed-HDAC-1* are involved much earlier. Conversely, *Smed-p66(RNAi)* had no effect on *prog-1* cell number, but like *Smed-mbd2/3* RNAi, lead to a decrease in AGAT-1+ cells [114]. Intriguingly, following *Smed-p66* RNAi there was also an increase in photoreceptor neurons (PRNs), but no difference in eye pigment cup cell (PCC) production, indicating that p66 acts to suppress PRN production in wild-type worms [114]. Moreover, the NB proliferation responses in *Smed-HDAC1*, *Smed-CHD4* and *Smed-RbAp48* knockdown animals are reduced, whereas *Smed-mbd2/3* and *Smed-p66* RNAi worms had normal levels of proliferation and formed blastemas [80,109-113].

The conflicting RNAi phenotypes for different NuRD components can be explained by most of them having roles in other complexes. For instance, both mammalian RbAp48 and HDAC-1 have been shown to be a member of the Sin3a deacetylase complex which, in a complex with *Nanog*, is involved in the activation of pluripotency factors and suppression of differentiation genes. Moreover, RbAp48 is also an important co-factor in the chromatin assembly factor (CAF-1) complex whose function is to initiate nucleosome assembly by adding histones H3 and H4 onto newly synthesized DNA [115,116]. Likewise, there is accumulating evidence that vertebrate CHD4 has functions independent of the NuRD complex [117]. **(Table 1).**

Overall, it is clear that future studies investigating the role of the NuRD complex in planarians should utilise the *Smed-mbd2/3* or *p66* phenotypes, as these subunits are specific to NuRD. Investigating the role of MBD2/3 in a DNA methylation-null organism has an important evolutionary significance, and may clarify an important DNA-methylation independent role. ChIP-seq will help to resolve whether genes are aberrantly marked by H3K27 acetylation and methylation in both NBs and NB-progeny following *Smed-mbd2/3* knockdown. Alternatively, development of a ChIP-grade SMED-MBD2/3 antibody would help to determine the targets of this protein.

4.3 SET1/MLL family of proteins - functional insights from planarian studies.

Tri-methylation of lysine 4 on histone 3 (H3K4me3) is a major conserved mark of chromatin at nucleosomes immediately downstream of transcribed genes across metazoans. In yeast, the SET domain containing 1 gene (Set1) catalyses the mono-, di-, and tri-methylation of H3K4. The SET domain is a motif of ~130 amino acids that provide histone methyltransferase activity, and the SET1 protein forms a macromolecular complex called COMPASS (complex of proteins associated with SET1) [118,119]. In *Drosophila melanogaster* there are three proteins homologous to Set1: dSET1, Trithorax (Trx), Trithorax-related (Trr) which functions in a complex with LPT (lost plant homeodomains of Trr). In mammals, there are at least six Set1-related proteins: SetD1a and SetD1b that are orthologous to *Drosophila* dSet1; MLL1 and MLL2 orthologous to *Drosophila* Trx; and MLL3 and MLL4 that are orthologous to *Drosophila* LPT/Trr, with the N-terminus of MLL3/4 corresponding to LPT and the C-terminus for Trr (**Figure 3A and 3B**) [120]. Another homolog of Set1/MLL, called MLL5, is found in *Drosophila* and mammals, but lacks histone methyltransferase activity and has diverged in sequence and structure from other SET/MLL proteins [121,122].

Expansion of the COMPASS family evolutionary time implies diversification in H3K4 methylation function. *Drosophila* dSet1 and mammalian SetD1A and SetD1B complexes mediate the bulk of genomic H3K4me di- and tri-methylation indicating an involvement in global gene activation [123–125]. Conversely, mammalian MLL1 and MLL2 are required for the methylation of a subset of developmentally important gene promoters. MLL2 is largely responsible for the methylation of H3K4 at bivalent genes in ESCs, whereas MLL1 is required for the H3K4 trimethylation of a smaller subset of genes and may be functionally redundant [126,127].

Unlike the Set1/SetD1A/SetD1B and Trx/MLL1/2 complexes, the Trr/MLL3/MLL4 complexes are likely responsible for the deposition of H3K4me1 at promoters and, in particular, enhancers. Although active gene promoters are marked by H3K4me3 closest to the TSS, H3K4me1 at TSS-proximal regions is a mark of inactive genes and correlates with MLL3/4 occupancy at these regions. H3K4me1 spatially restricts H3K4me3 interactors on active genes, resulting in a bimodal ChIP-seq profile with H3K4me3 occupancy at the TSS but H3K4me1 signal both upstream and downstream of the TSS [128]. Conversely, H3K4me1 is also ubiquitous at active enhancers, but the functional relevance of this mark is not well understood. It has been proposed that MLL3/4 binds to enhancers, and recruits the coactivator p300, which acetylates H3K27 [129]. Mouse ESC (mESC) knockouts for the catalytic

domain of both MLL3 and MLL4 showed a loss of H3K4me1 and H3K27me3 at enhancers, but the overall effect on enhancer RNA (eRNA) production and gene transcription was minimal. Conversely, complete MLL3/4 knockouts have a strong reduction in enhancer RNAs (eRNAs) and diminished transcription of target gene bodies. These results suggest that the function of MLL3/4 as a long-range coactivator is unrelated to methyltransferase activity [129]. MLL3 and MLL4 are frequently mutated in a number of cancer types [130–135], and changes in enhancer function may underlie tissue specific alterations in gene expression leading to cancer pathogenesis [136,137].

Although these three groups of H3K4 methyltransferases have a well-documented role in promoter and enhancer activation, there is a substantial lack of knowledge concerning the exact loci these marks affect and whether these targets are evolutionarily conserved. Two separate planarian studies have sought to understand the effects of these enzymes in the context of neoblast differentiation by knockdown of individual H3K4 methyltransferases, and have successfully related whole-organism phenotypes with epigenetic changes at target loci using ChIP-seq (**Figure 3C and 3D**).

Smed-set1 RNAi resulted in extensive neoblast loss and worms failed to produce a significant blastema upon amputation [22,138]. ChIP-seq identified a number of stem cell genes involved in RNA/DNA binding (i.e. *piwi-1*), transcription (*soxP-1*, *soxP-2*), and chromatin modification (MLL3) that were depleted for H3K4me3 following *set1* RNAi (**Figure 3D**). Conversely, RNAi of the MLL1/2 ortholog resulted in a loss of epidermal cilia [22] consistent with locomotory defects, and an earlier report also recorded a loss of ciliated protonephridia [138]. ChIP-seq and RNA-seq following MLL1/2 RNAi identified a much narrower set of genes which were depleted for H3K4me3 and transcriptionally downregulated, including a number of ciliogenesis related genes. Interestingly, these ciliogenesis related genes were shown to have a high H3K4me3 mark in NBs, comparable with that of NB-related genes, despite having low transcript expression in this compartment. Consequently, one conclusion of this study is that MLL1/2 marks genes – specifically those involved in ciliogenesis - for later activation during differentiation keeping them in a transcriptionally poised state in stem cells (**Figure 3D**). In this way, cilia related genes may in fact be bivalent – being marked with both H3K4me3 and H3K27me3, analogous to the role of MLL2 in establishing H3K4me3 at bivalent genes in ESCs.

Three orthologs of mammalian MLL3 and MLL4 genes have been identified in the *S. mediterranea* genome. Two of these planarian orthologs are related to Drosophila Trx and the C-terminus of mammalian MLL3/4 – *Smed-Trx-1* and *Smed-Trx-2*. RNAi knockdown of *Smed-Trx-1* lead to a regenerative delay, with worms able to form a small blastema consistent with a reduction but not complete loss of mitotic neoblast activity. Conversely, *Smed-Trx-2* did not show any defects compared to wild-type worms. When double RNAi was performed with *Smed-Trx-1* and *Smed-Trx-2*, the phenotype was enhanced with worms unable to form a regenerative blastema consistent with stem cell loss [138]. Moreover, animals began to form tissue outgrowths indicative of NB over-proliferation [24]. This strengthening of phenotype with double Trx knockdown is indicative of a degree of functional redundancy between the *S. mediterranea* Trx homologs.

The singular planarian homolog of Drosophila LPT and the N-terminus of MLL3/4 contains two PHD fingers and a singular PHD-like zinc finger binding proteins indicative of chromatin binding. RNAi of *Smed-LPT* resulted in failure of stem cell differentiation of some lineages including neuronal, epidermal, and pharynx regions [24]. Moreover, *Smed-LPT* RNAi worms showed an increase in neoblast mitotic activity before the formation of epidermal outgrowths that are populated with NBs, lineage-defined NBs and epidermal progeny. Consequently, *Smed-LPT* knockdown results in proliferation and differentiation defects that have cancer-like features, which is significant as both MLL3 and 4 are tumour suppressors in mammals [130–135] . RNA-seq revealed a number of genes involved in cell proliferation and differentiation that were significantly upregulated including the serine-threonine kinase oncogenes *pim-2* and *pim-2-like*. Moreover, *Smed-p53* was significantly downregulated in this dataset, consistent with its well documented role as a tumour suppressor. For some genes, changes in transcription following *Smed-LPT* RNAi correlated with differences in H3K4me1 at the promoter region – for example, *pim-2-like* showed a decrease in H3K4me1 at the TSS indicative of its upregulation following *Smed-LPT* RNAi. However, for many genes correlations between RNA-seq and ChIP-seq data were not apparent. One likely explanation for this is that epigenetic changes at enhancers are acting to modulate changes in transcription, but currently enhancers are uncharacterised in planarians (**Figure 3D**). Future studies using ATAC-seq paired with available H3K4me1 data, would serve well to identify enhancers genome-wide, and changes to these regulatory elements following *Smed-LPT* RNAi can then be effectively assessed. Such a study may reveal novel enhancer targets that are conserved with mammalian MLL3

and MLL4, and more generally, would increase our knowledge of the effect enhancers have on transcriptional regulation during stem cell differentiation.

4.4 Polycomb repressive complex (PRC) and its role in maintaining bivalency and regulating stem cell differentiation

The term Polycomb originally referred to a mutant of *Drosophila* that displayed improper body segmentation, owing to a mis-regulation of homeotic genes [139,140]. The Polycomb Group of genes has since been used to define proteins that act as negative regulators of key developmental genes, and whose mutations result in phenotypes comparable to that of the Polycomb. Polycomb mediated gene silencing works through the post-transcriptional modification of histones at two marks: H3K27me3 and H2AK119ub [141].

The PRC2 complex is responsible for the di- and tri-methylation of lysine 27 of histone 3 (H3K27me3) via its enzymatic subunit Ezh [142–144]. This complex is broadly conserved in eukaryotes, but has been lost in some yeast species such as *S. pombe* and *S. cerevisiae* consistent with a genome-wide absence of H3K27me3 [145]. Conversely, the PRC1 complex has been traditionally thought of as being downstream of the PRC2 complex, with both complexes synergizing to recruit each other's modifying enzymes. The H3K27me3 mark, laid down by PRC2, is recognized and bound to by the chromodomain of the CBX protein component of the PRC1 complex. The E3-ligase protein (RING) of PRC1 ubiquitinates H2AK119, and this suppresses gene transcription by inhibiting RNA Pol II transcriptional elongation [146,147]. However, studies from vertebrates indicate that PRC1 also has a number of non-canonical roles independent of gene silencing owing to an existence of multiple paralogs of PRC1 components ((reviewed in [148,149])).

Although high resolution microscopy experiments have suggested that both PRC1 and PRC2 mediated histone marks are important for chromatin compaction [150,151], sequential ChIP experiments have revealed the co-occurrence of PRC components (RING1B and Ezh2) with RNA Pol II at loci in ESCs, which is mirrored by the presence of both H3K4me3 and H3K27me3 [152,153]. These bivalent loci are defined by the presence of 'poised' RNA Pol II, which has a phosphorylated Serine 5 (S5p) at the heptapeptide repeat, YSTPSP, of the C-terminal domain (CTD) of the largest subunit RPB1.

Conversely, bivalent genes lack the elongating form of RNA Pol II, characterized by the presence of phosphorylated Serine 2 (S2p) at the heptapeptide [154]. The correlation between Pol II S5p and PRC repression means that bivalent genes are prepared for transcriptional activation upon differentiation. Indeed, knockouts of PRC1 and PRC2 components lead to the aberrant expression of differentiation related genes, many of which are bivalent [155–157].

The role of PRC genes in regulating planarian stem cell differentiation remains relatively unexplored, but the scope for investigation is considerable. Three planarian genes encoding homologs of PRC2 subunits, *Smed-ezh*, *Smed-suz12-1*, *Smed-eed-1*, were shown to be expressed in planarian stem cells by *in situ* hybridisation. RNAi knockdown of these genes in wild-type worms followed by amputation did not produce an observable phenotype, with no stem cell loss. In order to sensitize animals to any subtle RNAi defects, a dose of sublethal irradiation (12.5Gy) was used to reduce the stem cell number, such that surviving stem cells would have to proliferate in order to repopulate the stem cell compartment – a recovery process that takes 14 days [158]. Utilising this assay, the PRC2 genes were shown to be necessary for stem cell clonal expansion, with worms displaying epidermal lesions and eventual lysis as a consequence of stem cell loss. Although this phenotype is reflective of a general role of PRC2 in NB biology, a genomics and transcriptomic based approaches would help in elucidating misrelated genes following RNAi.

It is clear that further analysis is needed to understand the effects of the canonical PRC complexes on the regulation of neoblast gene expression programs both during maintenance and differentiation. For example, bulk sequencing of the X1 NB compartment revealed genes involved in the differentiation process that were marked with both H3K4me3 and H3K27me3 at their promoters, indicative of bivalency [23]. In ESCs, CpG islands play an important role as PRC recruitment elements and are important in the assembly of bivalent chromatin at key developmental genes and restriction of elongating Pol II. [159,160]. Given that planarians do not have CpG islands understanding how PRC complexes localise to bivalent genes independently of this genomic reference could be relevant to mammalian biology.

4.5 PIWI, epigenetic silencing of transposable elements, and probable role in pluripotency gene regulation in planarian NBs.

A major selective force during the evolution of an organism's genome is the maintenance of genomic integrity over generations [161,162]. TEs are highly mutagenic, as they can insert into protein-coding genes, and contain repetitive sequences that can initiate ectopic recombination [163]. Given that TEs constitute a large proportion of eukaryotic genome, their repression is necessary for the maintenance of gene function and genomic stability. This is particularly true for pluri- and multi-potent stem cells that must repress TE activity in order to maintain long-term proliferation over successive generations [164,165]. In order to combat the invasion of TEs, metazoans have evolved a novel RNA class called PIWI-interacting RNAs (piRNAs). These small RNA molecules are 24-31nt long, and are transcribed from TE derived piRNA clusters in the genome. piRNAs bind to members of the PIWI (P-element induced wimpy testes) subclass of Argonaute superfamily of proteins [166–169]. PIW-piRNA effector complexes can silence TEs either by epigenetic modifications at their genomic sites (transcriptional silencing TGS) or by cleaving TE transcripts directly (post-transcriptional silencing - PTGS) (**Figure 4a**). Most animals typically have at least one nuclear expressed PIWI protein and one or two cytoplasmically expressed PIWI proteins that employ these distinct silencing modes [170]. For example, *Drosophila* germ cells express two cytoplasmic PIWI proteins, Argonaute 3 (Ago3) and Aubergine (Aub), and one nuclear called Piwi. The cytoplasmic *Drosophila* Ago3 and Aub bind to complementary TEs following transcription and directly cleave them using slicer activity. They also generate additional template piRNAs (secondary piRNAs) from the transposon debris thereby generating a piRNA self-amplification loop termed the 'ping-pong' cycle [171–173]. The nuclear *Drosophila* Piwi functions to silence TEs epigenetically by recruitment of the DNA methylation and/or histone modifying complexes that lay down the H3K9me3 mark concomitant with the formation of heterochromatin [174–176].

The mechanism by which the PIWI-piRNA effectors guide the chromatin modifying machinery to the TE locus is beginning to be elucidated in animals. For example, independent RNAi screens in *Drosophila* identified the ovary specific nuclear protein CG9754/Panoramix that when eliminated leads to TE transcriptional increases similar to Piwi knockdown [177,178]. It is likely that CG9754/Panoramix acts as a protein scaffold between the nuclear PIWI-piRNA complex and the TGS machinery that includes the H3K9 methyltransferase SETDB1 and the heterochromatin protein HP1. No significant homology for the *Drosophila* CG9754/Panoramix has been identified in mammals or other invertebrates. An

evolutionary arms race between host organism and TE parasite means that proteins involved in the PIWI-piRNA pathway have diverged significantly [179].

The planarian genome has a repetitive content of 61.7%, far exceeding the 46% repeat content of the human genome, with many substantially large LTR members more than 30kb in length [180]. Previous studies have identified three major planarian PIWI proteins: SMEDWI-1, SMEDWI-2, and SMEDWI-3 in *Schmidtea mediterranea* and their respective orthologs DjPIWIA, DjPIWIB, and DjPIWIC in a sister species *Dugesia japonica*. All three of these genes are highly expressed in NBs and knockdown of *smedwi-2* and *smedwi-3* (but not *smedwi-1*) causes reductions in organismal piRNA levels, resulting in regenerative defects and lethality. One study in *D. japonica* revealed that following depletion of DjPIWIB, TEs continue to be silenced in NBs 7 days post RNAi, and NBs still retain the capacity for proliferation [181]. However, TEs were de-repressed at the onset of differentiation, and *in situ* hybridisation detected the up-regulation of a gypsy element in differentiated cells. Moreover, antibody staining revealed that DjPIWIA and DjPIWIC have cytoplasmic expression patterns restricted to the NB compartment, whereas DjPIWIB is expressed at the protein level in the nuclei of NBs and continues to be expressed in post-mitotic progeny and differentiated cells [182]. We can hypothesise on the basis of protein expression and phenotype that DjPIWIB, like *Drosophila* Piwi, may function in the epigenetic silencing of TE loci, whereas DjPIWIA and DjPIWIB function in a ping-pong cycle, cleaving cytoplasmic TE mRNAs and generating piRNAs [162,182,183] (**Figure 4A**). Thus, when DjPIWIB is lost in the NBs, TEs continue to be silenced by the cytoplasmic PIWI proteins, and it is only at the onset of differentiation, when DjPIWIA and DJPIWIC expression is lost, that TE deleterious activity increases. The question as to whether DjPIWIB/SMEDWI-2 is an epigenetic TE silencer remains outstanding, and proving so may potentially help in the understanding of NB maintenance and differentiation [180]. Since the planarian genome is replete with TEs, it is likely that the epigenetic silencing of these parasitic elements has shaped both the genome architecture and gene regulatory networks. For example, the deposition of the heterochromatic H3K9me3 mark at TEs in *Drosophila* germline stem cells has been shown to bleed to nearby gene promoters, causing their repression, or at least dampening of expression (**Figure 4B**) [184]. For planarian NBs, this would lead to a trade-off whereby TEs neighbouring highly expressed NB genes escape epigenetic silencing so that the NB gene expression program is not compromised. Thus, these TEs would be transcribed, but may be cleaved by the cytoplasmic PIWI

proteins thereby preventing genomic stress. Conversely, it is possible that genes with high expression in the differentiated compartment (Xins) are able to establish heterochromatic marks at neighbouring TEs as they have no transcriptional activity in NBs. If this hypothesis is true, genes with high Xins expression will be aberrantly expressed in NBs following DjPIWIB/SMEDWI-2 knockdown, concomitant with a loss of the H3K9me3 mark at promoters. In order for Xins genes to be expressed correctly in the differentiated compartment, the effect of DjPIWIB/SMEDWI-2 on H3K9me3 must be counteracted, or at least dampened if DjPIWIB persists in all differentiated cells [182]. Alternatively, purifying selection may remove deleterious TEs that have suppressive effects on neighbouring NB genes at least in the asexual biotype, owing to lower heterozygosity that unmasks deleterious recessive alleles [185,186]. These evolutionary scenarios can be explicitly tested in the planarian NB system and can reveal how the genome architecture of animals can be shaped by the co-evolution with parasitic TEs.

Planarian NB piRNAs themselves may mediate small RNA transgenerational inheritance between successive NB divisions. As the asexual species must persist as an adult population the somatic NBs must be collectively immortal and underpin the homeostatic turnover of adult tissue. The maintenance of genomic integrity is therefore vital and cannot be compromised. In *Drosophila*, maternal deposition of cytoplasmic piRNAs to the developing egg prior to zygotic transcription is important in kick-starting the piRNA generation system of the embryo in two distinct ways. In the nucleus, inherited piRNAs add the H3K9me3 to activate piRNA clusters in the embryo. In the cytoplasm, inherited piRNAs initiate the ping-pong cycle by providing an initial substrate for the cytoplasmic PIWI proteins [174]. A similar system may exist in planarian NBs to ensure genomic integrity between successive NB divisions. Moreover, if planarian piRNAs exist that are complementary to coding elements and are necessary for the suppression of differentiation genes, this raises an exciting possibility that NB identity can be preserved by piRNA mediated transgenerational inheritance.

5. Summary and prospects:

In this review, we have documented studies revealing the involvement of the NuRD, PRC complexes, SET1/MLL family of proteins, and PIWI in the epigenetic regulation of the planarian stem cell pluripotency and differentiation. In all four examples, elucidations of the function of these complexes

have unravelled interesting biology that is relevant to mammalian stem cells and disease. Utilisation of epigenomic techniques such as ChIP-seq and ATAC-seq will potentially reveal gene targets of these chromatin complexes that may have been previously overlooked in traditional stem cell culture based systems. Moreover, comparative analyses between the epigenomes of planarian NBs and mammalian stem cells may also uncover conserved and divergent areas of the epigenome that potentially relate to the biology, life-history, and environment of the animal in question.

Thus far, the community has only considered the roles of canonical epigenetic complexes by RNAi, but it would be worth taking a top-down approach. For example, during planarian de-growth and growth, the animal remodels itself proportionally in response to nutritional status [49,50]. How NBs are able to co-ordinate this feat at a molecular level remains an outstanding question, but the underlying mechanism is likely to involve feedback signalling from the differentiated cells mediating epigenetic changes to the entire stem cell compartment. Additionally, how stem cells read their position in the animal to activate region-specific to differentiation programmes, most likely depends on the translation of signal gradients and positional control genes by epigenetic complexes to initiate both general and cell-type specific differentiation programs [187]. The recent advent of single-cell ChIP-seq and ATAC-seq provides a promising avenue into these planarian investigations, and will help to characterize regulatory differences between individual NBs [188–191] .

Acknowledgements:

We thank Sounak Sahu and Prasad Abnave for useful discussions. Research in the authors' is funded by the MRC (MR/M000133/1). AD is funded by a BBSRC studentship (BB/J014427/1). DS is supported by a Natural Motion Scholarship.

Figure/Table Legends:

*Figure 1: Presence/absence of DNA methylation in flatworms and other species. The methylation status of the metazoan ancestor is unknown, but DNA methylation is not present in many invertebrates. The presence of DNMT and MBD genes in metazoan species were confirmed by tBLASTn in NCBI against species genomes/transcriptomes as well as cross-referencing with the findings of [63]. Although all three DNMT orthologs have been identified in the *Macrostomum lignano* genome [79], we cannot confirm the presence of MBD4/MeCP2. The presence of an MBD1/2/3 ortholog is based on a tBLASTn search using <http://www.macgenome.org/>, but we also find 2 additional MBD genes, which may be a *Macrostomum*-specific MBD1/2/3 paralogs.*

Figure 2: (A) Hypothesis as to the function of MBDs in vertebrates adapted from [100]. MBD2 binds to 5mC at promoters and recruits the NuRD complex to reinforce silencing or dampening of gene transcriptional activity (dashed arrow) of target promoters (e.g. pluripotency and differentiation-related genes). TET1 oxidises the 5methyl group of 5mC to 5hmC. MBD3 binds to 5hmC with greater affinity than 5mC. DNMT1 localization depends on the presence of 5hmC, and this leads to the restoration of 5mC at this site. Alternatively, following additional TET1-TDG-BER activity, CpG sites are fully de-methylated. This can lead to activator binding and transcriptional activity of the gene. The presence of MBD3/NuRD at these active de-methylated gene promoters is contentious, but is represented on the diagram. (B) In methylation-free invertebrates, MBD3/NuRD can either bind gene promoters and directly or indirectly via another DNA binding protein resulting in gene silencing or reduction in gene transcription.

Table 1: Planarian orthologs of NuRD subunits. tBLASTn searches of human NuRD components were made against the dd_smed_v6 transcriptome and IDs are tabulated. FACS RNA-seq proportions (i.e. X1/X2/Xins) were obtained using the dataset from [23]. X1 refers to NBs, X2 to post-mitotic cells, and Xins differentiated cells. RNAi phenotypes, post-amputation (pa) or post-RNAi in homeostatic conditions, are documented for each gene where already available

Fig. 3: (A) Domain architecture and duplication of SET1/MLL proteins in Yeast, Drosophila, Schmidtea mediterranea and humans. (B) Phylogenetic relationships SET1/MLL proteins in these four organisms. Proteins were aligned in MAFFT and RaxML was used for constructing a maximum likelihood phylogeny using the PROT GAMMA model and support for branches was estimated with 100 pseudoreplicates. (c) Cartoon showing planarian phenotypes after gene knockdown by RNAi. (D) Diagram showing hypothetical mechanistic function relating to phenotypes and neoblast ChIP-seq analyses of Set1, MLL1/2 and LPT/Trr-1/Trr-2 genes. Smed-SET1 activates stem cell genes genome-wide by the addition of H3K4me3 mark at promoters. Smed-MLL1/2 regulates the transfer of H3K4me3 at differentiation-related gene promoters (i.e. cillliogenesis-related genes) that may be bivalent and marked by H3K27me3. MLL3/4 transfer H3K4me1 to active enhancers and inactive promoters may regulate genes involved in cell proliferation and differentiation.

Figure 4 (A) Hypothesis as to the function of the three planarian genes. PIWIB (SMEDWI-2 in Schmidtea mediterranea) functions in epigenetic silencing (transcriptional gene silencing TGS) of transposons in the nucleus by recruiting a H3K9me3 methyltransferases. TEs that escape epigenetic silencing are cleaved in the cytoplasm and are post-transcriptionally silenced (PTGS) by binding of PIWIA (SMEDWI-1) and PIWIC (SMEDWI-3). These two cytoplasmic PIWI proteins participate in a ping-pong pathway to cleave TEs and produce piRNAs. (B) TEs in wild-type (WT) planarian NBs are transcriptionally silent owing to H3K9me3 deposition preventing the recruitment and/or elongation of RNA Pol II. Following knockdown of DjPiwib in Dugesia japonica or smedwi-2 in Schmidtea mediterranea, TE activity increases. In the case of a TE within an intron of a protein-coding gene, TE transcription and gene transcription both increase following knockdown.

Figure 1

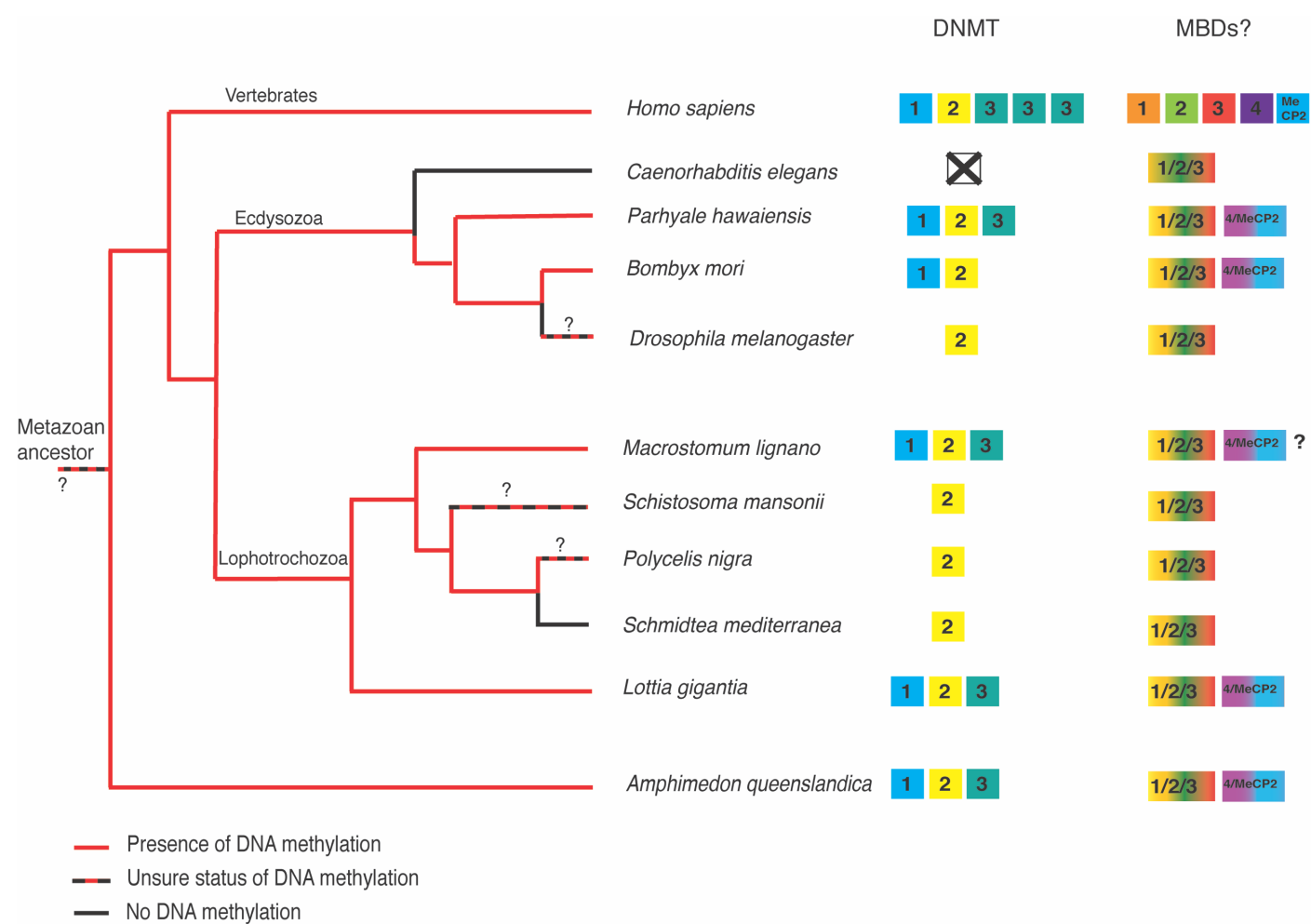
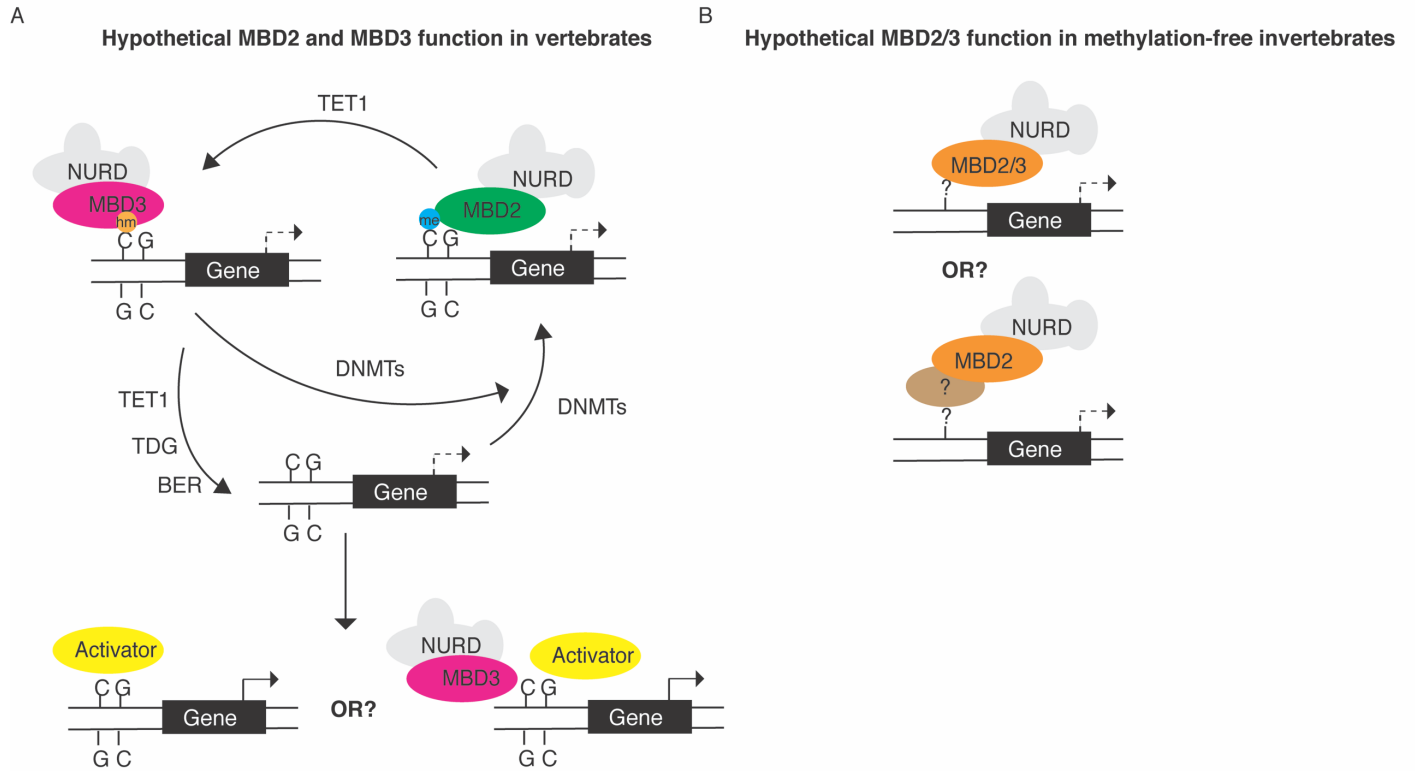


Figure 2



Gene	Function	Dresden Transcriptome ID	X1 Prop	X2 Prop	Xins Prop	NBs	Mitotic Activity	Early Epidermal Progeny	Late Epidermal Progeny	Other Phenotypes	Ref
MBD2/3	NuRD	dd_smed_v6_3054_0_2	43	52	5	No NB loss D7pa or 3 weeks post-RNAi homeostasis	No reduction D7pa or 3 weeks post-RNAi homeostasis	Increase D14pa No overall change, but anterior accumulation 3 weeks post-RNAi homeostasis	Reduction D7pa and 3 weeks post-RNAi homeostasis	Loss of gut branches, eyes, pharynx, brain neurons D7pa. Protonephridia and VNCs still form D7pa.	[81]
HDAC-1	NuRD, CoREST/REST, NCoR/SMRT, Sin3, SHIP1	dd_smed_v6_695_0_1	46	49	5	Reduction D2pa or Day 10 post-RNAi homeostasis	Complete loss D1pa or reduction D4 post-RNAi homeostasis	Reduction D2pa or D10 post-RNAi homeostasis	Reduction D3pa or D12 post-RNAi homeostasis	No blastema formation following amputation	[111]
MTA1-like-1	NuRD and other complexes involved in DNA damage, Inflammation, EMT transition	dd_smed_v6_5860_0_2	44	49	6	-	-	-	-	-	-
MTA1-like-2		dd_smed_v6_3995_0_1	36	53	10	-	-	-	-	-	-
LSD-1	NuRD, CoREST/CtBP, SIRT1	dd_smed_v6_7431_0_1	47	44	9	-	-	-	-	-	-
Rbap46/48-1	NuRD, Sin3	dd_smed_v6_5055_0_1	61	34	5	-	-	-	-	-	-
Rbap46/48-2		dd_smed_v6_2065_0_1	45	50	5	-	Reduction D14pa	-	-	-	[113]
Rbap46/48-3		dd_smed_v6_5609_0_1	53	44	4	-	-	-	-	-	-
CHD4	NuRD, Sin3	dd_smed_v6_2331_0_1	33	56	10	Reduction D16-18 post-RNAi homeostasis	Reduction D16-18 post-RNAi homeostasis	Reduction D16-18 post-RNAi homeostasis	Reduction D6 post-RNAi homeostasis	No blastema formation following amputation	[108]
CHD3/CHD5		dd_smed_v6_9090_0_1	53	38	10	-	-	-	-	-	-
p66a	NuRD	dd_smed_v6_3115_0_1	28	62	11	Increase D14 post-RNAi homeostasis	Slight reduction D8pa	No change	Reduction D14 post-RNAi homeostasis	Loss of eye PRNs, brain neurons D7pa. Protonephridia and VNCs still form D8pa.	[114]

Table 1

Figure 3

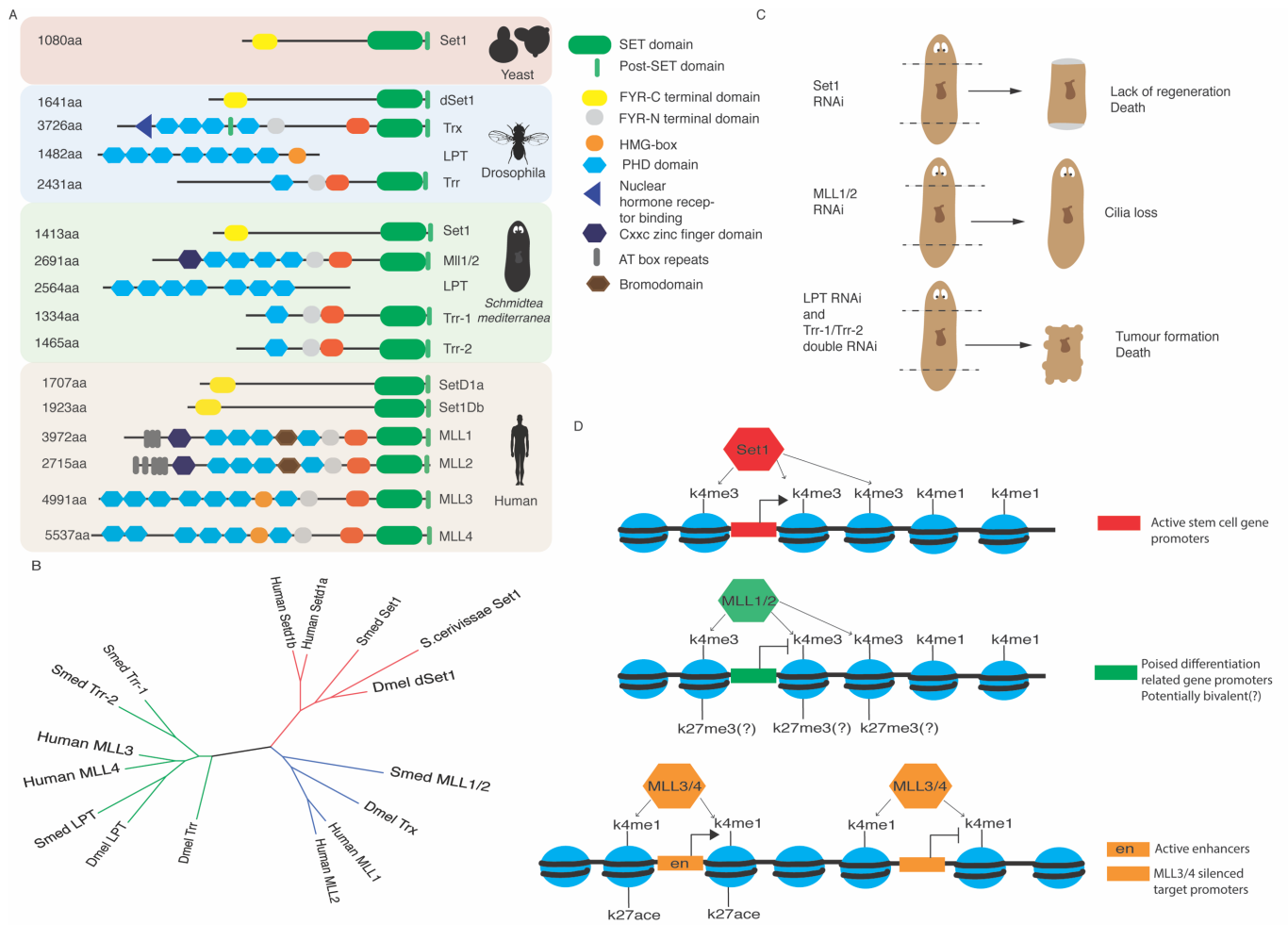
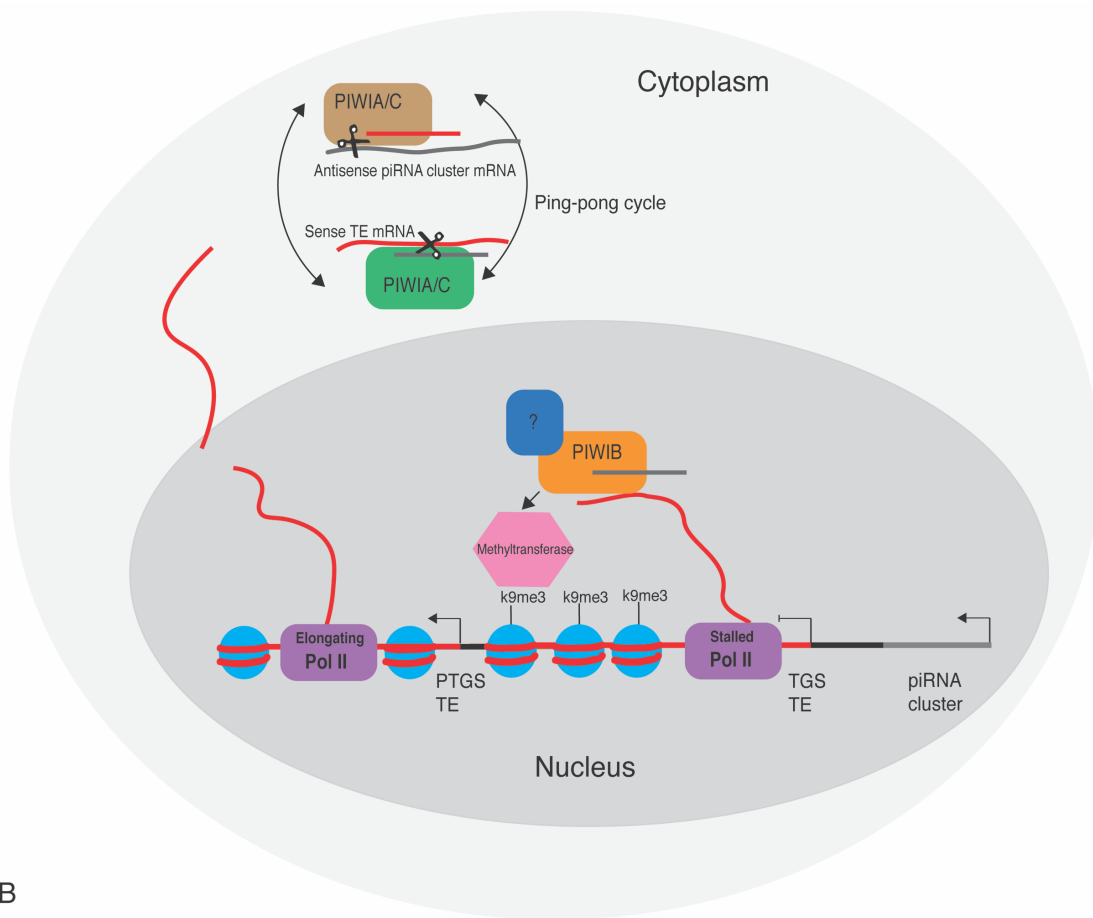
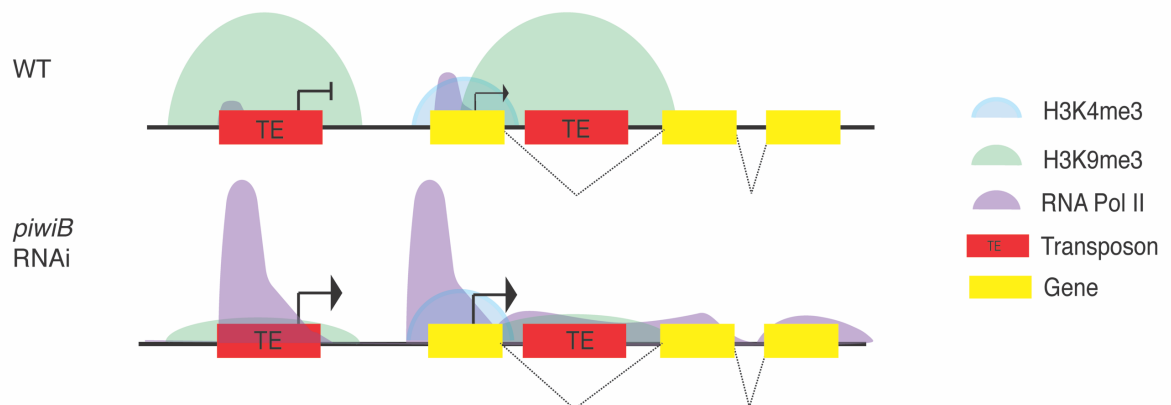


Figure 4

A



B



Bibliography:

- [1] O. Rechavi, I. Lev, Principles of Transgenerational Small RNA Inheritance in *Caenorhabditis elegans*, *Curr. Biol.* 27 (2017) R720–R730. doi:10.1016/j.cub.2017.05.043.
- [2] S. Anava, R. Posner, O. Rechavi, The soft genome, *Worm.* 3 (2014) e989798. doi:10.4161/21624054.2014.989798.
- [3] A. Bird, DNA methylation patterns and epigenetic memory, *Genes Dev.* 16 (2002) 6–21. doi:10.1101/gad.947102.
- [4] S. Tweedie, J. Charlton, V. Clark, A. Bird, Methylation of genomes and genes at the invertebrate-vertebrate boundary., *Mol. Cell. Biol.* 17 (1997) 1469–1475. doi:10.1128/MCB.17.3.1469.
- [5] G.P. Dunwell, Thomas L ; Pfeifer, *Drosophila* genomic methylation: new evidence and new questions, *Epigenomics.* 6 (2015) 459–461. doi:10.2217/epi.14.46.*Drosophila*.
- [6] S. Tweedie, H.H. Ng, A.L. Barlow, B.M. Turner, B. Hendrich, A. Bird, Vestiges of a DNA methylation system in *Drosophila melanogaster*?, *Nat. Genet.* 23 (1999) 389–390. doi:10.1038/70490.
- [7] F. Lyko, B.H. Ramsahoye, R. Jaenisch, DNA methylation in *Drosophila melanogaster*., *Nature.* 408 (2000) 538–540. doi:10.1038/35046205.
- [8] K. Luger, A.W. Mäder, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature.* 389 (1997) 251–260. doi:10.1038/38444.
- [9] B.D. Strahl, C.D. Allis, The language of covalent histone modifications, *Nature.* 403 (2000) 41–45. doi:10.1038/47412.
- [10] T. Kouzarides, Chromatin Modifications and Their Function, *Cell.* 128 (2007) 693–705. doi:10.1016/j.cell.2007.02.005.
- [11] S.R. Bhaumik, E. Smith, A. Shilatifard, Covalent modifications of histones during development and disease pathogenesis, *Nat. Struct. Mol. Biol.* 14 (2007) 1008–1016. doi:10.1038/nsmb1337.
- [12] A.G. Smith, J.K. Heath, D.D. Donaldson, G.G. Wong, J. Moreau, M. Stahl, D. Rogers, Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides, *Nature.* 336 (1988) 688–690. doi:10.1038/336688a0.
- [13] R.L. Williams, D.J. Hilton, S. Pease, T. a Willson, C.L. Stewart, D.P. Gearing, E.F. Wagner, D. Metcalf, N. a Nicola, N.M. Gough, Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells., *Nature.* 336 (1988) 684–687. doi:10.1038/336684a0.
- [14] Q.-L. Ying, J. Wray, J. Nichols, L. Batlle-Morera, B. Doble, J. Woodgett, P. Cohen, A. Smith, The ground state of embryonic stem cell self-renewal, *Nature.* 453 (2008) 519–523. doi:10.1038/nature06968.
- [15] J. Martin Gonzalez, S.M. Morgani, R.A. Bone, K. Bonderup, S. Abelchian, C. Brakebusch, J.M. Brickman, Embryonic Stem Cell Culture Conditions Support Distinct States Associated with Different Developmental Stages and Potency, *Stem Cell Reports.* 7 (2016) 177–191. doi:10.1016/j.stemcr.2016.07.009.
- [16] H. Marks, T. Kalkan, R. Menafrá, S. Denisov, K. Jones, H. Hofemeister, J. Nichols, A. Kranz, A. Francis Stewart, A. Smith, H.G. Stunnenberg, The transcriptional and epigenomic foundations of ground state pluripotency, *Cell.* 149 (2012) 590–604. doi:10.1016/j.cell.2012.03.026.

- [17] M. Yagi, S. Kishigami, A. Tanaka, K. Semi, E. Mizutani, S. Wakayama, T. Wakayama, T. Yamamoto, Y. Yamada, Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation, *Nature*. 548 (2017) 224–227. doi:10.1038/nature23286.
- [18] M. Yagi, S. Yamanaka, Y. Yamada, Epigenetic foundations of pluripotent stem cells that recapitulate in vivo pluripotency, *Lab. Investig.* 97 (2017) 1133–1141. doi:10.1038/labinvest.2017.87.
- [19] K. Takahashi, S. Yamanaka, Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors, *Cell*. 126 (2006) 663–676. doi:10.1016/j.cell.2006.07.024.
- [20] R.D. Hawkins, G.C. Hon, L.K. Lee, Q. Ngo, R. Lister, M. Pelizzola, L.E. Edsall, S. Kuan, Y. Luu, S. Klugman, J. Antosiewicz-Bourget, Z. Ye, C. Espinoza, S. Agarwahl, L. Shen, V. Ruotti, W. Wang, R. Stewart, J.A. Thomson, J.R. Ecker, B. Ren, Distinct epigenomic landscapes of pluripotent and lineage-committed human cells, *Cell Stem Cell*. 6 (2010) 479–491. doi:10.1016/j.stem.2010.03.018.
- [21] O. Bar-Nur, H.A. Russ, S. Efrat, N. Benvenisty, Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells, *Cell Stem Cell*. 9 (2011) 17–23. doi:10.1016/j.stem.2011.06.007.
- [22] E.M. Duncan, A.D. Chitsazan, C.W. Seidel, A.S. Alvarado, Erratum: Set1 and MLL1/2 Target Distinct Sets of Functionally Different Genomic Loci In Vivo (*Cell Reports* (2015) 13 (12) (p2741–2755)(S2211124715014084)(10.1016/j.celrep.2015.11.059)), *Cell Rep.* 17 (2016) 930. doi:10.1016/j.celrep.2016.09.071.
- [23] D. Kao, Y. Mihaylova, S. Hughes, A. Lai, A. Aboobaker, Epigenetic analyses of the planarian genome reveals conservation of bivalent promoters in animal stem cells., *Bioarxiv*. (2017). doi:https://doi.org/10.1101/122135.
- [24] Y. Mihaylova, D. Kao, S. Hughes, A. Lai, F. Jaber-Hijazi, N. Kosaka, P. Abnave, A. Aboobaker, MLL3/4 Prevents Stem Cell Hyperplasia And Controls Differentiation Programs In A Planarian Cancer Stem Cell Model, *Bioarxiv*. (2017). doi:https://doi.org/10.1101/126540.
- [25] J.C. Rink, Stem cell systems and regeneration in planaria, *Dev. Genes Evol.* 223 (2013) 67–84. doi:10.1007/s00427-012-0426-4.
- [26] A.A. Aboobaker, Planarian stem cells: A simple paradigm for regeneration, *Trends Cell Biol.* 21 (2011) 304–311. doi:10.1016/j.tcb.2011.01.005.
- [27] C.E. Adler, A. Sánchez Alvarado, Types or States? Cellular Dynamics and Regenerative Potential, *Trends Cell Biol.* 25 (2015) 687–696. doi:10.1016/j.tcb.2015.07.008.
- [28] P. Önal, D. Grün, C. Adamidi, A. Rybak, J. Solana, G. Mastrobuoni, Y. Wang, H.P. Rahn, W. Chen, S. Kempa, U. Ziebold, N. Rajewsky, Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells, *EMBO J.* 31 (2012) 2755–2769. doi:10.1038/emboj.2012.110.
- [29] J. Solana, D. Kao, Y. Mihaylova, F. Jaber-Hijazi, S. Malla, R. Wilson, A. Aboobaker, Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNAseq, RNA interference and irradiation approach, *Genome Biol.* 13 (2012). doi:10.1186/gb-2012-13-3-r19.
- [30] R.M. Labbé, M. Irimia, K.W. Currie, A. Lin, S.J. Zhu, D.D.R. Brown, E.J. Ross, V. Voisin, G.D. Bader, B.J. Blencowe, B.J. Pearson, A Comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals, *Stem Cells*.

- 30 (2012) 1734–1745. doi:10.1002/stem.1144.
- [31] A. Alié, T. Hayashi, I. Sugimura, M. Manuel, W. Sugano, A. Mano, N. Satoh, K. Agata, N. Funayama, The ancestral gene repertoire of animal stem cells, *Proc. Natl. Acad. Sci.* (2015) 201514789. doi:10.1073/pnas.1514789112.
 - [32] J. Solana, Closing the circle of germline and stem cells: the Primordial Stem Cell hypothesis, *Evodevo.* 4 (2013) 2. doi:10.1186/2041-9139-4-2.
 - [33] R.A. Blassberg, D.A. Felix, B. Tejada-Romero, A.A. Aboobaker, PBX/extradenticle is required to re-establish axial structures and polarity during planarian regeneration, *Development.* 140 (2013) 730–739. doi:10.1242/dev.082982.
 - [34] M. Iglesias, J.L. Gomez-Skarmeta, E. Salo, T. Adell, Silencing of Smed- catenin1 generates radial-like hypercephalized planarians, *Development.* 135 (2008) 1215–1221. doi:10.1242/dev.020289.
 - [35] C.P. Petersen, P.W. Reddien, Smed-??catenin-1 is required for anteroposterior blastema polarity in planarian regeneration, *Science* (80-.). 319 (2008) 327–330. doi:10.1126/science.1149943.
 - [36] C.P. Petersen, P.W. Reddien, A wound-induced Wnt expression program controls planarian regeneration polarity, *Proc. Natl. Acad. Sci.* 106 (2009) 17061–17066. doi:10.1073/pnas.0906823106.
 - [37] S. Yazawa, Y. Umesono, T. Hayashi, H. Tarui, K. Agata, Planarian Hedgehog/Patched establishes anterior-posterior polarity by regulating Wnt signaling, *Proc. Natl. Acad. Sci.* 106 (2009) 22329–22334. doi:10.1073/pnas.0907464106.
 - [38] J.C. Rink, K.A. Gurley, S.A. Elliott, A.S. Alvarado, Planarian Hh signaling regulates regeneration polarity and links Hh pathway evolution to cilia, *Science* (80-.). 326 (2009) 1406–1410. doi:10.1126/science.1178712.
 - [39] M.A. Gaviño, P.W. Reddien, A Bmp/Admp regulatory circuit controls maintenance and regeneration of dorsal-ventral polarity in planarians, *Curr. Biol.* 21 (2011) 294–299. doi:10.1016/j.cub.2011.01.017.
 - [40] M.L. Scimone, L.E. Cote, T. Rogers, P.W. Reddien, Two FGFR1-Wnt circuits organize the planarian anteroposterior axis, *Elife.* 5 (2016). doi:10.7554/eLife.12845.
 - [41] R. Lander, C.P. Petersen, Wnt, Ptk7, and FGFR1 expression gradients control trunk positional identity in planarian regeneration, *Elife.* 5 (2016). doi:10.7554/eLife.12850.
 - [42] O. Wurtzel, L.E. Cote, A. Poirier, R. Satija, A. Regev, P.W. Reddien, A Generic and Cell-Type-Specific Wound Response Precedes Regeneration in Planarians, *Dev. Cell.* 35 (2015) 632–645. doi:10.1016/j.devcel.2015.11.004.
 - [43] M. Lucila Scimone, L.E. Cote, P.W. Reddien, Orthogonal muscle fibres have different instructive roles in planarian regeneration, *Nature.* 551 (2017) 623–628. doi:10.1038/nature24660.
 - [44] J.N. Witchley, M. Mayer, D.E. Wagner, J.H. Owen, P.W. Reddien, Muscle cells provide instructions for planarian regeneration, *Cell Rep.* 4 (2013) 633–641. doi:10.1016/j.celrep.2013.07.022.
 - [45] D. Kao, D. Felix, A. Aboobaker, The planarian regeneration transcriptome reveals a shared but temporally shifted regulatory program between opposing head and tail scenarios, *BMC Genomics.* 14 (2013). doi:10.1186/1471-2164-14-797.
 - [46] M.L. Scimone, K.M. Kravarik, S.W. Lapan, P.W. Reddien, Neoblast specialization in regeneration of the planarian *schmidtea mediterranea*, *Stem Cell Reports.* 3 (2014) 339–352. doi:10.1016/j.stemcr.2014.06.001.
 - [47] H. Reuter, M. März, M.C. Vogg, D. Eccles, L. Grífol-Boldú, D. Wehner, S. Owlarn, T.

- Adell, G. Weidinger, K. Bartscherer, β -Catenin-Dependent Control Of Positional Information Along The AP body axis in planarians involves a teashirt family member, *Cell Rep.* 10 (2015) 253–265. doi:10.1016/j.celrep.2014.12.018.
- [48] S. Owlarn, K. Bartscherer, Go ahead, grow a head! A planarian's guide to anterior regeneration, *Regeneration.* 3 (2016) 139–155. doi:10.1002/reg2.56.
- [49] C. González-Estévez, D.A. Felix, G. Rodríguez-Esteban, A. Aziz Aboobaker, Decreased neoblast progeny and increased cell death during starvation-induced planarian degrowth, *Int. J. Dev. Biol.* 56 (2012) 83–91. doi:10.1387/ijdb.113452cg.
- [50] M. Mangel, M.B. Bonsall, A. Aboobaker, Feedback control in planarian stem cell systems, *BMC Syst. Biol.* 10 (2016). doi:10.1186/S12918-016-0261-8.
- [51] E.M. De Robertis, Wnt signaling in axial patterning and regeneration: Lessons from planaria, *Sci. Signal.* 3 (2010). doi:10.1126/scisignal.3127pe21.
- [52] K.C. Tu, B.J. Pearson, A. Sánchez Alvarado, TORC1 is required to balance cell proliferation and cell death in planarians, *Dev. Biol.* 365 (2012) 458–469. doi:10.1016/j.ydbio.2012.03.010.
- [53] M. Almuedo-Castillo, X. Crespo, F. Seebeck, K. Bartscherer, E. Salò, T. Adell, JNK Controls the Onset of Mitosis in Planarian Stem Cells and Triggers Apoptotic Cell Death Required for Regeneration and Remodeling, *PLoS Genet.* 10 (2014). doi:10.1371/journal.pgen.1004400.
- [54] B. Tejada-Romero, J.-M. Carter, Y. Mihaylova, B. Neumann, A.A. Aboobaker, JNK signalling is necessary for a Wnt- and stem cell-dependent regeneration programme., *Development.* 142 (2015) 2413–2424. doi:10.1242/dev.115139.
- [55] J.A. Law, S.E. Jacobsen, Establishing, maintaining and modifying DNA methylation patterns in plants and animals, *Nat. Rev. Genet.* 11 (2010) 204–220. doi:10.1038/nrg2719.
- [56] T.P. Jurkowski, A. Jeltsch, On the evolutionary origin of eukaryotic DNA methyltransferases and Dnmt2, *PLoS One.* 6 (2011). doi:10.1371/journal.pone.0028104.
- [57] A. Zemach, D. Zilberman, Evolution of eukaryotic DNA methylation and the pursuit of safer sex, *Curr. Biol.* 20 (2010). doi:10.1016/j.cub.2010.07.007.
- [58] F. Tuorto, F. Herbst, N. Alerasool, S. Bender, O. Popp, G. Federico, S. Reitter, R. Liebers, G. Stoecklin, H.-J. Gröne, G. Dittmar, H. Glimm, F. Lyko, The tRNA methyltransferase Dnmt2 is required for accurate polypeptide synthesis during haematopoiesis., *Embo J.* 34 (2015) 2350–62. doi:10.15252/emboj.201591382.
- [59] M.G. Goll, F. Kirpekar, K.A. Maggert, J.A. Yoder, C.L. Hsieh, X. Zhang, K.G. Golic, S.E. Jacobsen, T.H. Bestor, Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2, *Science* (80-.). 311 (2006) 395–398. doi:10.1126/science.1120976.
- [60] S. Phalke, O. Nickel, D. Walluscheck, F. Hortig, M.C. Onorati, G. Reuter, Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2, *Nat. Genet.* 41 (2009) 696–702. doi:10.1038/ng.360.
- [61] M. Schaefer, F. Lyko, Lack of evidence for DNA methylation of *Invader4* retroelements in *Drosophila* and implications for Dnmt2-mediated epigenetic regulation, *Nat. Genet.* 42 (2010) 920–921. doi:10.1038/ng1110-920.
- [62] R. Albalat, Evolution of DNA-methylation machinery: DNA methyltransferases and methyl-DNA binding proteins in the amphioxus *Branchiostoma floridae*, *Dev. Genes Evol.* 218 (2008) 691–701. doi:10.1007/s00427-008-0247-7.

- [63] R. Albalat, J. Martí-Solans, C. Cañestro, Dna methylation in amphioxus: From ancestral functions to new roles in vertebrates, *Brief. Funct. Genomics*. 11 (2012) 142–155. doi:10.1093/bfpg/els009.
- [64] A.M. Deaton, A. Bird, CpG islands and the regulation of transcription, *Genes Dev*. 25 (2011) 1010–1022. doi:10.1101/gad.2037511.
- [65] A. Bird, M. Taggart, M. Frommer, O.J. Miller, D. Macleod, A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA, *Cell*. 40 (1985) 91–99. doi:10.1016/0092-8674(85)90312-5.
- [66] A. Zemach, I.E. McDaniel, P. Silva, D. Zilberman, Genome-wide evolutionary analysis of eukaryotic DNA methylation., *Science*. 328 (2010) 916–9. doi:10.1126/science.1186366.
- [67] B.K. Duncan, J.H. Miller, Mutagenic deamination of cytosine residues in DNA, *Nature*. 287 (1980) 560–561. doi:10.1038/287560a0.
- [68] R.J. Britten, W.F. Baron, D.B. Stout, E.H. Davidson, Sources and evolution of human Alu repeated sequences, *Evolution (N. Y.)*. 85 (1988) 4770–4774. doi:10.1073/pnas.85.13.4770.
- [69] J. Sved, A. Bird, The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model., *Proc. Natl. Acad. Sci.* 87 (1990) 4692–4696. doi:10.1073/pnas.87.12.4692.
- [70] M. Bulmer, Neighboring base effects on substitution rates in pseudogenes., *Mol. Biol. Evol.* 3 (1986) 322–329. doi:10.1093/oxfordjournals.molbev.a040401.
- [71] S. Rošić, R. Amouroux, C.E. Requena, A. Gomes, M. Emperle, T. Beltran, J.K. Rane, S. Linnett, M.E. Selkirk, P.H. Schiffer, A.J. Bancroft, R.K. Grencis, A. Jeltsch, P. Hajkova, P. Sarkies, Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity, *Nat. Genet.* (2018). doi:10.1038/s41588-018-0061-8.
- [72] M. Rosado Fantappiè, E. Rodrigues Pereira Gimba, F.D. Rumjanek, Lack of DNA methylation in *Schistosoma mansoni*, *Exp. Parasitol.* 98 (2001) 162–166. doi:10.1006/expr.2001.4630.
- [73] K.K. Geyer, I.W. Chalmers, N. MacKintosh, J.E. Hirst, R. Geoghegan, M. Badets, P.M. Brophy, K. Brehm, K.F. Hoffmann, Cytosine methylation is a conserved epigenetic feature found throughout the phylum Platyhelminthes, *BMC Genomics*. 14 (2013). doi:10.1186/1471-2164-14-462.
- [74] G. Raddatz, P.M. Guzzardo, N. Olova, M.R. Fantappie, M. Rampp, M. Schaefer, W. Reik, G.J. Hannon, F. Lyko, Dnmt2-dependent methylomes lack defined DNA methylation patterns, *Proc. Natl. Acad. Sci.* 110 (2013) 8627–8631. doi:10.1073/pnas.1306723110.
- [75] K.K. Geyer, C.M. Rodríguez López, I.W. Chalmers, S.E. Munshi, M. Truscott, J. Heald, M.J. Wilkinson, K.F. Hoffmann, Cytosine methylation regulates oviposition in the pathogenic blood fluke *Schistosoma mansoni*, *Nat. Commun.* 2 (2011). doi:10.1038/ncomms1433.
- [76] M. Schaefer, S. Hagemann, K. Hanna, F. Lyko, Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines, *Cancer Res.* 69 (2009) 8127–8132. doi:10.1158/0008-5472.CAN-09-0458.
- [77] C.J. Marsit, Influence of environmental exposure on human epigenetic regulation., *J. Exp. Biol.* 218 (2015) 71–79. doi:10.1242/jeb.106971.
- [78] M. Zauri, G. Berridge, M.L. Thézéas, K.M. Pugh, R. Goldin, B.M. Kessler, S.

- Kriaucionis, CDA directs metabolism of epigenetic nucleosides revealing a therapeutic window in cancer, *Nature*. 524 (2015) 114–118. doi:10.1038/nature14948.
- [79] K. Wasik, J. Gurtowski, X. Zhou, O.M. Ramos, M.J. Delás, G. Battistoni, O. El Demerdash, I. Falciatori, D.B. Vizoso, A.D. Smith, P. Ladurner, L. Schärer, W.R. McCombie, G.J. Hannon, M. Schatz, Genome and transcriptome of the regeneration-competent flatworm, *Macrostomum lignano*, *Proc. Natl. Acad. Sci.* 112 (2015) 201516718. doi:10.1073/pnas.1516718112.
- [80] J. Wudarski, D. Simanov, K. Ustyantsev, K. De Mulder, M. Grelling, M. Grudniewska, F. Beltman, L. Glazenburg, T. Demircan, J. Wunderer, W. Qi, D.B. Vizoso, P.M. Weissert, D. Olivieri, S. Mouton, V. Guryev, A. Aboobaker, L. Schärer, P. Ladurner, E. Berezikov, Efficient transgenesis and annotated genome sequence of the regenerative flatworm model *Macrostomum lignano*, (n.d.). doi:10.1038/s41467-017-02214-8.
- [81] F. Jaber-Hijazi, P.J.K.P. Lo, Y. Mihaylova, J.M. Foster, J.S. Benner, B. Tejada Romero, C. Chen, S. Malla, J. Solana, A. Ruzov, A. Aziz Aboobaker, Planarian MBD2/3 is required for adult stem cell pluripotency independently of DNA methylation, *Dev. Biol.* 384 (2013) 141–153. doi:10.1016/j.ydbio.2013.09.020.
- [82] X. Zou, W. Ma, I.A. Solov'Yov, C. Chipot, K. Schulten, Recognition of methylated DNA through methyl-CpG binding domain proteins, *Nucleic Acids Res.* 40 (2012) 2747–2758. doi:10.1093/nar/gkr1057.
- [83] I. Ohki, N. Shimotake, N. Fujita, J.G. Jee, T. Ikegami, M. Nakao, M. Shirakawa, Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA, *Cell*. 105 (2001) 487–497. doi:10.1016/S0092-8674(01)00324-5.
- [84] N. Reynolds, M. Salmon-Divon, H. Dvinge, A. Hynes-Allen, G. Balasooriya, D. Leaford, A. Behrens, P. Bertone, B. Hendrich, NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression, *EMBO J.* 31 (2012) 593–605. doi:10.1038/emboj.2011.431.
- [85] N. Reynolds, P. Latos, A. Hynes-Allen, R. Loos, D. Leaford, A. O'Shaughnessy, O. Mosaku, J. Signolet, P. Brennecke, T. Kalkan, I. Costello, P. Humphreys, W. Mansfield, K. Nakagawa, J. Strouboulis, A. Behrens, P. Bertone, B. Hendrich, NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment, *Cell Stem Cell*. 10 (2012) 583–594. doi:10.1016/j.stem.2012.02.020.
- [86] J. Ogas, S. Kaufmann, J. Henderson, C. Somerville, PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in *Arabidopsis*, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 13839–13844. doi:10.1073/pnas.96.24.13839.
- [87] Y. Unhavaithaya, T.H. Shin, N. Miliaras, J. Lee, T. Oyama, C.C. Mello, MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in *C. elegans*, *Cell*. 111 (2002) 991–1002. doi:10.1016/S0092-8674(02)01202-3.
- [88] T. von Zelewsky, F. Palladino, K. Brunschwig, H. Tobler, a Hajnal, F. Müller, The *C. elegans* Mi-2 chromatin-remodelling proteins function in vulval cell fate determination., *Development*. 127 (2000) 5277–5284.
- [89] J. Nitarska, J.G. Smith, W.T. Sherlock, M.M.G. Hillege, A. Nott, W.D. Barshop, A.A. Vashisht, J.A. Wohlschlegel, R. Mitter, A. Riccio, A Functional Switch of NuRD Chromatin Remodeling Complex Subunits Regulates Mouse Cortical Development, *Cell Rep.* 17 (2016) 1683–1698. doi:10.1016/j.celrep.2016.10.022.
- [90] Y. Wang, H. Zhang, Y. Chen, Y. Sun, F. Yang, W. Yu, J. Liang, L. Sun, X. Yang, L. Shi, R. Li,

- Y. Li, Y. Zhang, Q. Li, X. Yi, Y. Shang, LSD1 Is a Subunit of the NuRD Complex and Targets the Metastasis Programs in Breast Cancer, *Cell*. 138 (2009) 660–672. doi:10.1016/j.cell.2009.05.050.
- [91] A. Adamo, B. Sesé, S. Boue, J. Castaño, I. Paramonov, M.J. Barrero, J.C.I. Belmonte, LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells, *Nat. Cell Biol.* 13 (2011) 652–661. doi:10.1038/ncb2246.
- [92] W.A. Whyte, S. Bilodeau, D.A. Orlando, H.A. Hoke, G.M. Frampton, C.T. Foster, S.M. Cowley, R.A. Young, Enhancer decommissioning by LSD1 during embryonic stem cell differentiation, *Nature*. 482 (2012) 221–225. doi:10.1038/nature10805.
- [93] G. Hu, P.A. Wade, NuRD and pluripotency: A complex balancing act, *Cell Stem Cell*. 10 (2012) 497–503. doi:10.1016/j.stem.2012.04.011.
- [94] B. Hendrich, A. Bird, Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins, *Mol. Cell. Biol.* 18 (1998) 6538–6547. doi:10.1128/MCB.18.11.6538.
- [95] Y. Zhang, H.H. Ng, H. Erdjument-Bromage, P. Tempst, A. Bird, D. Reinberg, Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation, *Genes Dev.* 13 (1999) 1924–1935. doi:10.1101/gad.13.15.1924.
- [96] M. Mellén, P. Ayata, S. Dewell, S. Kriaucionis, N. Heintz, MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system, *Cell*. 151 (2012) 1417–1430. doi:10.1016/j.cell.2012.11.022.
- [97] O. Yildirim, R. Li, J.H. Hung, P.B. Chen, X. Dong, L.S. Ee, Z. Weng, O.J. Rando, T.G. Fazzio, Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells, *Cell*. 147 (2011) 1498–1510. doi:10.1016/j.cell.2011.11.054.
- [98] X. Lu, B.S. Zhao, C. He, TET family proteins: Oxidation activity, interacting molecules, and functions in diseases, *Chem. Rev.* 115 (2015) 2225–2239. doi:10.1021/cr500470n.
- [99] M. Tahiliani, K.P. Koh, Y. Shen, W.A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L.M. Iyer, D.R. Liu, L. Aravind, A. Rao, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, *Science* (80-.). 324 (2009) 930–935. doi:10.1126/science.1170116.
- [100] S.J. Hainer, K.N. McCannell, J. Yu, L.S. Ee, L.J. Zhu, O.J. Rando, T.G. Fazzio, DNA methylation directs genomic localization of Mbd2 and Mbd3 in embryonic stem cells, *Elife*. 5 (2016). doi:10.7554/eLife.21964.
- [101] S.A. Denslow, P.A. Wade, The human Mi-2/NuRD complex and gene regulation, *Oncogene*. 26 (2007) 5433–5438. doi:10.1038/sj.onc.1210611.
- [102] T. Baubec, R. Ivánek, F. Lienert, D. Schübeler, Methylation-dependent and -independent genomic targeting principles of the mbd protein family, *Cell*. 153 (2013) 480–492. doi:10.1016/j.cell.2013.03.011.
- [103] T. Shimbo, Y. Du, S.A. Grimm, A. Dhasarathy, D. Mav, R.R. Shah, H. Shi, P.A. Wade, MBD3 Localizes at Promoters, Gene Bodies and Enhancers of Active Genes, *PLoS Genet.* 9 (2013). doi:10.1371/journal.pgen.1004028.
- [104] K. Günther, M. Rust, J. Leers, T. Boettger, M. Scharfe, M. Jarek, M. Bartkuhn, R. Renkawitz, Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences, *Nucleic Acids Res.* 41 (2013) 3010–3021. doi:10.1093/nar/gkt035.
- [105] R. Menafrá, H.G. Stunnenberg, MBD2 and MBD3: Elusive functions and mechanisms, *Front. Genet.* 5 (2014). doi:10.3389/fgene.2014.00428.

- [106] B. Hendrich, S. Tweedie, The methyl-CpG binding domain and the evolving role of DNA methylation in animals, *Trends Genet.* 19 (2003) 269–277. doi:10.1016/S0168-9525(03)00080-5.
- [107] J.M. Cramer, D. Pohlmann, F. Gomez, L. Mark, B. Kornegay, C. Hall, E. Siraliev-Perez, N.M. Walavalkar, M.J. Sperlazza, S. Bilinovich, J.W. Prokop, A.L. Hill, D.C. Williams, Methylation specific targeting of a chromatin remodeling complex from sponges to humans, *Sci. Rep.* 7 (2017). doi:10.1038/srep40674.
- [108] M.L. Scimone, J. Meisel, P.W. Reddien, The Mi-2-like Smed-CHD4 gene is required for stem cell differentiation in the planarian *Schmidtea mediterranea*, *Development.* 137 (2010) 1231–1241. doi:10.1242/dev.042051.
- [109] G.T. Eisenhoffer, H. Kang, A.S. Alvarado, Molecular Analysis of Stem Cells and Their Descendants during Cell Turnover and Regeneration in the Planarian *Schmidtea mediterranea*, *Cell Stem Cell.* 3 (2008) 327–339. doi:10.1016/j.stem.2008.07.002.
- [110] S.J. Zhu, B.J. Pearson, The Retinoblastoma pathway regulates stem cell proliferation in freshwater planarians, *Dev. Biol.* 373 (2013) 442–452. doi:10.1016/j.ydbio.2012.10.025.
- [111] S.M.C. Robb, A.S. Alvarado, Histone Modifications and Regeneration in the Planarian *Schmidtea mediterranea*, *Curr. Top. Dev. Biol.* 108 (2014) 71–93. doi:10.1016/B978-0-12-391498-9.00004-8.
- [112] L. Bonuccelli, L. Rossi, A. Lena, V. Scarcelli, G. Rainaldi, M. Evangelista, P. Iacopetti, V. Gremigni, A. Salvetti, An RbAp48-like gene regulates adult stem cells in planarians., *J. Cell Sci.* 123 (2010) 690–698. doi:10.1242/jcs.053900.
- [113] A. Hubert, J.M. Henderson, M.W. Cowles, K.G. Ross, M. Hagen, C. Anderson, C.J. Szeterlak, R.M. Zayas, A functional genomics screen identifies an Importin- α homolog as a regulator of stem cell function and tissue patterning during planarian regeneration, *BMC Genomics.* 16 (2015). doi:10.1186/s12864-015-1979-1.
- [114] C. Vásquez-Doorman, C.P. Petersen, The NuRD complex component *p66* suppresses photoreceptor neuron regeneration in planarians, *Regeneration.* 3 (2016) 168–178. doi:10.1002/reg2.58.
- [115] S. Smith, B. Stillman, Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro., *Cell.* 58 (1989) 15–25. doi:10.1016/0092-8674(89)90398-X.
- [116] K. Marheineke, T. Krude, Nucleosome assembly activity and intracellular localization of human CAF-1 changes during the cell division cycle, *J. Biol. Chem.* 273 (1998) 15279–86. doi:10.1074/jbc.273.24.15279.
- [117] A. O’Shaughnessy-Kirwan, J. Signolet, I. Costello, S. Gharbi, B. Hendrich, Constraint of gene expression by the chromatin remodelling protein CHD4 facilitates lineage specification, *Development.* 142 (2015) 2586–2597. doi:10.1242/dev.125450.
- [118] T. Miller, N.J. Krogan, J. Dover, H. Erdjument-Bromage, P. Tempst, M. Johnston, J.F. Greenblatt, A. Shilatifard, COMPASS: a complex of proteins associated with a trithorax-related SET domain protein., *Proc. Natl. Acad. Sci. U. S. A.* 98 (2001) 12902–7. doi:10.1073/pnas.231473398.
- [119] N.J. Krogan, J. Dover, S. Khorrami, J.F. Greenblatt, J. Schneider, M. Johnston, A. Shilatifard, COMPASS, a histone H3 (lysine 4) methyltransferase required for telomeric silencing of gene expression, *J. Biol. Chem.* 277 (2002) 10753–10755. doi:10.1074/jbc.C200023200.
- [120] J.C. Eissenberg, A. Shilatifard, Histone H3 lysine 4 (H3K4) methylation in development

- and differentiation, *Dev. Biol.* 339 (2010) 240–249. doi:10.1016/j.ydbio.2009.08.017.
- [121] X. Zhang, W. Novera, Y. Zhang, L.-W. Deng, MLL5 (KMT2E): structure, function, and clinical relevance, *Cell. Mol. Life Sci.* 74 (2017) 2333–2344. doi:10.1007/s00018-017-2470-8.
 - [122] B.M. Emerling, J. Bonifas, C.P. Kratz, S. Donovan, B.R. Taylor, E.D. Green, M.M. Le Beau, K.M. Shannon, MLL5, a homolog of *Drosophila trithorax* located within a segment of chromosome band 7q22 implicated in myeloid leukemia, *Oncogene*. 21 (2002) 4849–4854. doi:10.1038/sj.onc.1205615.
 - [123] M.B. Ardehali, A. Mei, K.L. Zobeck, M. Caron, J.T. Lis, T. Kusch, *Drosophila* Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription, *EMBO J.* 30 (2011) 2817–2828. doi:10.1038/emboj.2011.194.
 - [124] M. Mohan, H.-M. Herz, E.R. Smith, Y. Zhang, J. Jackson, M.P. Washburn, L. Florens, J.C. Eissenberg, A. Shilatifard, The COMPASS Family of H3K4 Methylases in *Drosophila*, *Mol. Cell. Biol.* 31 (2011) 4310–4318. doi:10.1128/MCB.06092-11.
 - [125] M. Wu, P.F. Wang, J.S. Lee, S. Martin-Brown, L. Florens, M. Washburn, A. Shilatifard, Molecular Regulation of H3K4 Trimethylation by Wdr82, a Component of Human Set1/COMPASS, *Mol. Cell. Biol.* 28 (2008) 7337–7344. doi:10.1128/MCB.00976-08.
 - [126] T.A. Milne, S.D. Briggs, H.W. Brock, M.E. Martin, D. Gibbs, C.D. Allis, J.L. Hess, MLL targets SET domain methyltransferase activity to Hox gene promoters, *Mol. Cell.* 10 (2002) 1107–1117. doi:10.1016/S1097-2765(02)00741-4.
 - [127] S. Denissov, H. Hofemeister, H. Marks, A. Kranz, G. Ciotta, S. Singh, K. Anastassiadis, H.G. Stunnenberg, A.F. Stewart, Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant, *Development*. 141 (2014) 526–537. doi:10.1242/dev.102681.
 - [128] J. Cheng, R. Blum, C. Bowman, D. Hu, A. Shilatifard, S. Shen, B.D. Dynlacht, A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers, *Mol. Cell.* 53 (2014) 979–992. doi:10.1016/j.molcel.2014.02.032.
 - [129] K.M. Dorigi, T. Swigut, T. Henriques, N. V. Bhanu, B.S. Scruggs, N. Nady, C.D. Still, B.A. Garcia, K. Adelman, J. Wysocka, Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation, *Mol. Cell.* 66 (2017) 568–576.e4. doi:10.1016/j.molcel.2017.04.018.
 - [130] J. Lee, D.-H. Kim, S. Lee, Q.-H. Yang, D.K. Lee, S.-K. Lee, R.G. Roeder, J.W. Lee, A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4., *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 8513–8. doi:10.1073/pnas.0902873106.
 - [131] R.D. Morin, M. Mendez-Lago, A.J. Mungall, R. Goya, K.L. Mungall, R.D. Corbett, N.A. Johnson, T.M. Severson, R. Chiu, M. Field, S. Jackman, M. Krzywinski, D.W. Scott, D.L. Trinh, J. Tamura-Wells, S. Li, M.R. Firme, S. Rogic, M. Griffith, S. Chan, O. Yakovenko, I.M. Meyer, E.Y. Zhao, D. Smailus, M. Moksa, S. Chittaranjan, L. Rimsza, A. Brooks-Wilson, J.J. Spinelli, S. Ben-Neriah, B. Meissner, B. Woolcock, M. Boyle, H. McDonald, A. Tam, Y. Zhao, A. Delaney, T. Zeng, K. Tse, Y. Butterfield, I. Birol, R. Holt, J. Schein, D.E. Horsman, R. Moore, S.J.M. Jones, J.M. Connors, M. Hirst, R.D. Gascoyne, M.A. Marra, Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma, *Nature*. 476 (2011) 298–303. doi:10.1038/nature10351.
 - [132] D.W. Parsons, M. Li, X. Zhang, S. Jones, R.J. Leary, J.C.H. Lin, S.M. Boca, H. Carter, J. Samayoa, C. Bettgowda, G.L. Gallia, G.I. Jallo, Z.A. Binder, Y. Nikolsky, J. Hartigan, D.R. Smith, D.S. Gerhard, D.W. Fults, S. VandenBerg, M.S. Berger, S.K.N. Marie, S.M.O.

- Shinjo, C. Clara, P.C. Phillips, J.E. Minturn, J.A. Biegel, A.R. Judkins, A.C. Resnick, P.B. Storm, T. Curran, Y. He, B.A. Rasheed, H.S. Friedman, S.T. Keir, R. McLendon, P.A. Northcott, M.D. Taylor, P.C. Burger, G.J. Riggins, R. Karchin, G. Parmigiani, D.D. Bigner, H. Yan, N. Papadopoulos, B. Vogelstein, K.W. Kinzler, V.E. Velculescu, The genetic landscape of the childhood cancer medulloblastoma, *Science* (80-.). 331 (2011) 435–439. doi:10.1126/science.1198056.
- [133] D.T.W. Jones, N. Jäger, M. Kool, T. Zichner, B. Hutter, M. Sultan, Y.J. Cho, T.J. Pugh, V. Hovestadt, A.M. Stütz, T. Rausch, H.J. Warnatz, M. Ryzhova, S. Bender, D. Sturm, S. Pleier, H. Cin, E. Pfaff, L. Sieber, A. Wittmann, M. Remke, H. Witt, S. Hutter, T. Tzaridis, J. Weischenfeldt, B. Raeder, M. Avci, V. Amstislavskiy, M. Zapatka, U.D. Weber, Q. Wang, B. Lasitschka, C.C. Bartholomae, M. Schmidt, C. von Kalle, V. Ast, C. Lawerenz, J. Eils, R. Kabbe, V. Benes, P. van Sluis, J. Koster, R. Volckmann, D. Shih, M.J. Betts, R.B. Russell, S. Coco, G.P. Tonini, U. Schüller, V. Hans, N. Graf, Y.J. Kim, C. Monoranu, W. Roggendorf, A. Unterberg, C. Herold-Mende, T. Milde, A.E. Kulozik, A. von Deimling, O. Witt, E. Maass, J. Rössler, M. Ebinger, M.U. Schuhmann, M.C. Frühwald, M. Hasselblatt, N. Jabado, S. Rutkowski, A.O. von Bueren, D. Williamson, S.C. Clifford, M.G. McCabe, V.P. Collins, S. Wolf, S. Wiemann, H. Lehrach, B. Brors, W. Scheurlen, J. Felsberg, G. Reifenberger, P.A. Northcott, M.D. Taylor, M. Meyerson, S.L. Pomeroy, M.L. Yaspo, J.O. Korbel, A. Korshunov, R. Eils, S.M. Pfister, P. Lichter, Dissecting the genomic complexity underlying medulloblastoma, *Nature*. 488 (2012) 100–105. doi:10.1038/nature11284.
- [134] T.J. Pugh, S.D. Weeraratne, T.C. Archer, D.A. Pomeranz Krummel, D. Auclair, J. Bochicchio, M.O. Carneiro, S.L. Carter, K. Cibulskis, R.L. Erlich, H. Greulich, H. Greulich, N.J. Lennon, A. McKenna, J. Meldrum, A.H. Ramos, M.G. Ross, C. Russ, E. Shefler, A. Sivachenko, B. Sogoloff, P. Stojanov, P. Tamayo, J.P. Mesirov, V. Amani, N. Teider, S. Sengupta, J.P. Francois, P.A. Northcott, M.D. Taylor, F. Yu, G.R. Crabtree, A.G. Kautzman, S.B. Gabriel, G. Getz, N. Jäger, D.T.W. Jones, P. Lichter, S.M. Pfister, T.M. Roberts, J.L. Dargatzis, S.L. Pomeroy, Y.J. Cho, Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations, *Nature*. 488 (2012) 106–110. doi:10.1038/nature11329.
- [135] S.P. Cleary, W.R. Jeck, X. Zhao, K. Chen, S.R. Selitsky, G.L. Savich, T.X. Tan, M.C. Wu, G. Getz, M.S. Lawrence, J.S. Parker, J. Li, S. Powers, H. Kim, S. Fischer, M. Guindi, A. Ghanekar, D.Y. Chiang, Identification of driver genes in hepatocellular carcinoma by exome sequencing, *Hepatology*. 58 (2013) 1693–1702. doi:10.1002/hep.26540.
- [136] J.E. Lee, C. Wang, S. Xu, Y.W. Cho, L. Wang, X. Feng, A. Baldrige, V. Sartorelli, L. Zhuang, W. Peng, K. Ge, H3K4 mono- And di-methyltransferase MLL4 is required for enhancer activation during cell differentiation, *Elife*. 2013 (2013). doi:10.7554/eLife.01503.
- [137] D. Hu, X. Gao, M.A. Morgan, H.-M. Herz, E.R. Smith, A. Shilatifard, The MLL3/MLL4 Branches of the COMPASS Family Function as Major Histone H3K4 Monomethylases at Enhancers, *Mol. Cell. Biol.* 33 (2013) 4745–4754. doi:10.1128/MCB.01181-13.
- [138] A. Hubert, J.M. Henderson, K.G. Ross, M.W. Cowles, J. Torres, R.M. Zayas, Epigenetic regulation of planarian stem cells by the SET1/MLL family of histone methyltransferases, *Epigenetics*. 8 (2013) 79–91. doi:10.4161/epi.23211.
- [139] G. Jürgens, A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*, *Nature*. 316 (1985) 153–155. doi:10.1038/316153a0.
- [140] E.B. Lewis, A gene complex controlling segmentation in *Drosophila*, *Nat.* 276565-

570. 276 (1978) 565–570.
<http://www.nature.com/nature/journal/v276/n5688/abs/276565a0.html>.
- [141] J.A. Simon, R.E. Kingston, Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put, *Mol. Cell.* 49 (2013) 808–824. doi:10.1016/j.molcel.2013.02.013.
 - [142] J. Müller, C.M. Hart, N.J. Francis, M.L. Vargas, A. Sengupta, B. Wild, E.L. Miller, M.B. O'Connor, R.E. Kingston, J.A. Simon, Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex, *Cell.* 111 (2002) 197–208. doi:10.1016/S0092-8674(02)00976-5.
 - [143] R. Cao, L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R.S. Jones, Y. Zhang, Role of histone H3 lysine 27 methylation in polycomb-group silencing, *Science* (80-.). 298 (2002) 1039–1043. doi:10.1126/science.1076997.
 - [144] A. Kuzmichev, K. Nishioka, H. Erdjument-Bromage, P. Tempst, D. Reinberg, Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein., *Genes Dev.* 16 (2002) 2893–905. doi:10.1101/gad.1035902.
 - [145] S. Shaver, J.A. Casas-Mollano, R.L. Cerny, H. Cerutti, Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga *Chlamydomonas*, *Epigenetics.* 5 (2010) 301–312. doi:10.4161/epi.5.4.11608.
 - [146] R. Cao, Y. Tsukada, Y. Zhang, Role of Bmi-1 and Ring1A in H2A Ubiquitylation and Hox Gene Silencing Table S1 mHox, *Mol. Cell.* 20 (2005) 1–4. doi:10.1016/j.molcel.2005.12.002.
 - [147] W. Zhou, P. Zhu, J. Wang, G. Pascual, K.A. Ohgi, J. Lozach, C.K. Glass, M.G. Rosenfeld, Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation, *Mol. Cell.* 29 (2008) 69–80. doi:10.1016/j.molcel.2007.11.002.
 - [148] J. Gil, A. O'Loghlen, PRC1 complex diversity: Where is it taking us?, *Trends Cell Biol.* 24 (2014) 632–641. doi:10.1016/j.tcb.2014.06.005.
 - [149] S. Aranda, G. Mas, L. Di Croce, Regulation of gene transcription by Polycomb proteins, *Sci. Adv.* 1 (2015) e1500737–e1500737. doi:10.1126/sciadv.1500737.
 - [150] N.J. Francis, R.E. Kingston, C.L. Woodcock, Chromatin compaction by a polycomb group protein complex., *Science.* 306 (2004) 1574–7. doi:10.1126/science.1100576.
 - [151] A.N. Boettiger, B. Bintu, J.R. Moffitt, S. Wang, B.J. Beliveau, G. Fudenberg, M. Imaev, L.A. Mirny, C.T. Wu, X. Zhuang, Super-resolution imaging reveals distinct chromatin folding for different epigenetic states, *Nature.* 529 (2016) 418–422. doi:10.1038/nature16496.
 - [152] B.J. Lesch, D.C. Page, Poised chromatin in the mammalian germ line, *Development.* 141 (2014) 3619–3626. doi:10.1242/dev.113027.
 - [153] B.E. Bernstein, T.S. Mikkelsen, X. Xie, M. Kamal, D.J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S.L. Schreiber, E.S. Lander, A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells, *Cell.* 125 (2006) 315–326. doi:10.1016/j.cell.2006.02.041.
 - [154] E. Brookes, A. Pombo, Modifications of RNA polymerase II are pivotal in regulating gene expression states, *EMBO Rep.* 10 (2009) 1213–1219. doi:10.1038/embor.2009.221.
 - [155] V. Azuara, P. Perry, S. Sauer, M. Spivakov, H.F. Jørgensen, R.M. John, M. Gouti, M. Casanova, G. Warnes, M. Merckenschlager, A.G. Fisher, Chromatin signatures of

- pluripotent cell lines, *Nat. Cell Biol.* 8 (2006) 532–538. doi:10.1038/ncb1403.
- [156] L.A. Boyer, K. Plath, J. Zeitlinger, T. Brambrink, L.A. Medeiros, T.I. Lee, S.S. Levine, M. Wernig, A. Tajonar, M.K. Ray, G.W. Bell, A.P. Otte, M. Vidal, D.K. Gifford, R.A. Young, R. Jaenisch, Polycomb complexes repress developmental regulators in murine embryonic stem cells, *Nature*. 441 (2006) 349–353. doi:10.1038/nature04733.
- [157] D. Pasini, A.P. Bracken, J.B. Hansen, M. Capillo, K. Helin, The polycomb group protein Suz12 is required for embryonic stem cell differentiation., *Mol. Cell. Biol.* 27 (2007) 3769–79. doi:10.1128/MCB.01432-06.
- [158] D.E. Wagner, J.J. Ho, P.W. Reddien, Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis, *Cell Stem Cell*. 10 (2012) 299–311. doi:10.1016/j.stem.2012.01.016.
- [159] A. Tanay, A.H. O'Donnell, M. Damelin, T.H. Bestor, Hyperconserved CpG domains underlie Polycomb-binding sites., *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 5521–5526. doi:10.1073/pnas.0609746104.
- [160] S.-M. Lee, J. Lee, K.-M. Noh, W.-Y. Choi, S. Jeon, G.T. Oh, J. Kim-Ha, Y. Jin, S.-W. Cho, Y.-J. Kim, Intragenic CpG islands play important roles in bivalent chromatin assembly of developmental genes, *Proc. Natl. Acad. Sci.* 114 (2017) E1885–E1894. doi:10.1073/pnas.1613300114.
- [161] C. Ernst, D.T. Odom, C. Kutter, The emergence of piRNAs against transposon invasion to preserve mammalian genome integrity, *Nat. Commun.* 8 (2017). doi:10.1038/s41467-017-01049-7.
- [162] S. Sahu, A. Dattani, A.A. Aboobaker, Secrets from immortal worms: What can we learn about biological ageing from the planarian model system?, *Semin. Cell Dev. Biol.* 70 (2017) 108–121. doi:10.1016/j.semcdb.2017.08.028.
- [163] D.J. Hedges, P.L. Deininger, Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity, *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* 616 (2007) 46–59. doi:10.1016/j.mrfmmm.2006.11.021.
- [164] C. Juliano, J. Wang, H. Lin, Uniting Germline and Stem Cells: The Function of Piwi Proteins and the piRNA Pathway in Diverse Organisms, *Annu. Rev. Genet.* 45 (2011) 447–469. doi:10.1146/annurev-genet-110410-132541.
- [165] J.C. van Wolfswinkel, Piwi and potency: PIWI proteins in animal stem cells and regeneration, *Integr. Comp. Biol.* 54 (2014) 700–713. doi:10.1093/icb/icu084.
- [166] A. Aravin, D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, N. Iovino, P. Morris, M.J. Brownstein, S. Kuramochi-Miyagawa, T. Nakano, M. Chien, J.J. Russo, J. Ju, R. Sheridan, C. Sander, M. Zavolan, T. Tuschl, A novel class of small RNAs bind to MILI protein in mouse testes, *Nature*. 442 (2006) 203–207. doi:10.1038/nature04916.
- [167] A. Girard, R. Sachidanandam, G.J. Hannon, M.A. Carmell, A germline-specific class of small RNAs binds mammalian Piwi proteins, *Nature*. 442 (2006) 199–202. doi:10.1038/nature04917.
- [168] S.T. Grivna, E. Beyret, Z. Wang, H. Lin, A novel class of small RNAs in mouse spermatogenic cells, *Genes Dev.* 20 (2006) 1709–1714. doi:10.1101/gad.1434406.
- [169] T. Watanabe, A. Takeda, T. Tsukiyama, K. Mise, T. Okuno, H. Sasaki, N. Minami, H. Imai, Identification and characterization of two novel classes of small RNAs in the mouse germline: Retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes, *Genes Dev.* 20 (2006) 1732–1743. doi:10.1101/gad.1425706.
- [170] E.-M. Weick, E.A. Miska, piRNAs: from biogenesis to function, *Development*. 141 (2014) 3458–3471. doi:10.1242/dev.094037.

- [171] L.S. Gunawardane, K. Saito, K.M. Nishida, K. Miyoshi, Y. Kawamura, T. Nagami, H. Siomi, M.C. Siomi, A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*, *Science* (80-.). 315 (2007) 1587–1590. doi:10.1126/science.1140494.
- [172] J. Brennecke, A.A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, G.J. Hannon, Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in *Drosophila*, *Cell*. 128 (2007) 1089–1103. doi:10.1016/j.cell.2007.01.043.
- [173] H. Huang, Y. Li, K.E. Szulwach, G. Zhang, P. Jin, D. Chen, AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification, *J. Cell Biol.* 206 (2014) 217–230. doi:10.1083/jcb.201401002.
- [174] A. Le Thomas, E. Stuwe, S. Li, J. Du, G. Marinov, N. Rozhkov, Y.C.A. Chen, Y. Luo, R. Sachidanandam, K.F. Toth, D. Patel, A.A. Aravin, Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing, *Genes Dev.* 28 (2014) 1667–1680. doi:10.1101/gad.245514.114.
- [175] S. Shpiz, I. Olovnikov, A. Sergeeva, S. Lavrov, Y. Abramov, M. Savitsky, A. Kalmykova, Mechanism of the piRNA-mediated silencing of *Drosophila* telomeric retrotransposons, *Nucleic Acids Res.* 39 (2011) 8703–8711. doi:10.1093/nar/gkr552.
- [176] N. V. Rozhkov, M. Hammell, G.J. Hannon, Multiple roles for Piwi in silencing *Drosophila* transposons, *Genes Dev.* 27 (2013) 400–412. doi:10.1101/gad.209767.112.
- [177] Y. Yu, J. Gu, Y. Jin, Y. Luo, J.B. Preall, J. Ma, B. Czech, G.J. Hannon, Panoramix enforces piRNA-dependent cotranscriptional silencing, *Science* (80-.). 350 (2015) 339–342. doi:10.1126/science.aab0700.
- [178] S. Bagheri-Fam, A. Argentaro, T. Svingen, A.N. Combes, A.H. Sinclair, P. Koopman, V.R. Harley, D. Basquin, A. Spierer, F. Begeot, D.E. Koryakov, A.L. Todeschini, S. Ronssey, C. Vieira, P. Spierer, M. Delattre, D. Clynes, C. Jelinska, B. Xella, H. Ayyub, C. Scott, M. Mitson, S. Taylor, D.R. Higgs, R.J. Gibbons, K. Huyhn, M.B. Renfree, J. a Graves, A.J. Pask, and L.A. Simon J Elsässer, Kyung-Min Noh, Nichole Diaz, C David Allis, Banaszynski, P. Tang, D.J. Park, J. a Marshall Graves, V.R. Harley, G. Sienski, J. Batki, K. Senti, D. Dönertas, L. Tirian, K. Meixner, J. Brennecke, C.C. Giauque, S.E. Bickel, A.K. Singh, S.C. Lakhota, Silencio / CG9754 connects the Piwi – piRNA complex to the cellular heterochromatin machinery, *Chromosoma*. 11 (2015) 173–189. doi:10.1007/s00412-015-0540-y.
- [179] S.H. Lewis, H. Salmela, D.J. Obbard, Duplication and diversification of dipteran argonaute genes, and the evolutionary divergence of Piwi and Aubergine, *Genome Biol. Evol.* 8 (2016) 507–518. doi:10.1093/gbe/evw018.
- [180] M.A. Grohme, S. Schloissnig, A. Rozanski, M. Pippel, G.R. Young, S. Winkler, H. Brandl, I. Henry, A. Dahl, S. Powell, M. Hiller, E. Myers, J.C. Rink, The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms, *Nat.* 2018. 554 (2018) 56–61. doi:10.1038/nature25473.
- [181] N. Shibata, L. Rouhana, K. Agata, Cellular and molecular dissection of pluripotent adult somatic stem cells in planarians, *Dev. Growth Differ.* 52 (2010) 27–41. doi:10.1111/j.1440-169X.2009.01155.x.
- [182] N. Shibata, M. Kashima, T. Ishiko, O. Nishimura, L. Rouhana, K. Misaki, S. Yonemura, K. Saito, H. Siomi, M.C. Siomi, K. Agata, Inheritance of a Nuclear PIWI from Pluripotent Stem Cells by Somatic Descendants Ensures Differentiation by Silencing

- Transposons in Planarian, *Dev. Cell.* 37 (2016) 226–237. doi:10.1016/j.devcel.2016.04.009.
- [183] M.E. Tharp, A. Bortvin, DjPiwiB: A Rich Nuclear Inheritance for Descendants of Planarian Stem Cells, *Dev. Cell.* 37 (2016) 204–206. doi:10.1016/j.devcel.2016.04.022.
 - [184] G. Sienski, D. Dönertas, J. Brennecke, Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression, *Cell.* 151 (2012) 964–980. doi:10.1016/j.cell.2012.10.040.
 - [185] J.D. Hollister, B.S. Gaut, Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression, *Genome Res.* 19 (2009) 1419–1428. doi:10.1101/gr.091678.109.
 - [186] S.C.H. Barrett, D. Charlesworth, Effects of a change in the level of inbreeding on the genetic load., *Nature.* 352 (1991) 522–524. doi:10.1038/352522a0.
 - [187] O. Wurtzel, I.M. Oderberg, P.W. Reddien, Planarian Epidermal Stem Cells Respond to Positional Cues to Promote Cell-Type Diversity, *Dev. Cell.* 40 (2017) 491–504.e5. doi:10.1016/j.devcel.2017.02.008.
 - [188] A. Rotem, O. Ram, N. Shores, R.A. Sperling, A. Goren, D.A. Weitz, B.E. Bernstein, Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state, *Nat. Biotechnol.* 33 (2015) 1165–1172. doi:10.1038/nbt.3383.
 - [189] S.J. Clark, H.J. Lee, S.A. Smallwood, G. Kelsey, W. Reik, Single-cell epigenomics: Powerful new methods for understanding gene regulation and cell identity, *Genome Biol.* 17 (2016). doi:10.1186/s13059-016-0944-x.
 - [190] J.D. Buenrostro, B. Wu, U.M. Litzenburger, D. Ruff, M.L. Gonzales, M.P. Snyder, H.Y. Chang, W.J. Greenleaf, Single-cell chromatin accessibility reveals principles of regulatory variation, *Nature.* 523 (2015) 486–490. doi:10.1038/nature14590.
 - [191] J. Cao, J.S. Packer, V. Ramani, D.A. Cusanovich, C. Huynh, R. Daza, X. Qiu, C. Lee, S.N. Furlan, F.J. Steemers, A. Adey, R.H. Waterston, C. Trapnell, J. Shendure, Comprehensive single-cell transcriptional profiling of a multicellular organism, *Science* (80-.). 357 (2017) 661–667. doi:10.1126/science.aam8940.