Developmental Biology 495 (2023) 63–75

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

A genome-wide computational approach to define microRNA-Polycomb/trithorax gene regulatory circuits in *Drosophila*

Jacobo Solorzano [a,c], Enrique Carrillo-de Santa Pau [b], Teresa Laguna [b,**], Ana Busturia [a,*]

[a] Centro de Biología Molecular Severo Ochoa, CSIC-UAM, Nicolas Cabrera 1, 28049, Madrid, Spain
[b] Computational Biology Group, Precision Nutrition and Cancer Research Program, IMDEA Food Institute, CEI UAM+CSIC, 28049, Madrid, Spain
[c] Centre de Recherches en Cancerologie de Toulouse, 2 Av. Hubert Curien, 31100, Toulouse, France

ARTICLE INFO

Keywords:
Polycomb
Trithorax
microRNAs
Drosophila
Computational resource
Framework
Epigenetic
GRN (gene regulatory network)
Circuits

ABSTRACT

Characterization of gene regulatory networks is fundamental to understanding homeostatic development. This process can be simplified by analyzing relatively simple genomes such as the genome of *Drosophila melanogaster*. In this work we have developed a computational framework in *Drosophila* to explore for the presence of gene regulatory circuits between two large groups of transcriptional regulators: the epigenetic group of the Polycomb/trithorax (PcG/trxG) proteins and the microRNAs (miRNAs). We have searched genome-wide for miRNA targets in *PcG/trxG* transcripts as well as for Polycomb Response Elements (PREs) in miRNA genes. Our results show that 10% of the analyzed miRNAs could be controlling *PcG/trxG* gene expression, while 40% of those miRNAs are putatively controlled by the selected set of PcG/trxG proteins. The integration of these analyses has resulted in the predicted existence of 3 classes of miRNA-PcG/trxG crosstalk interactions that define potential regulatory circuits. In the first class, *miRNA-PcG circuits* are defined by miRNAs that reciprocally crosstalk with PcG. In the second, *miRNA-trxG circuits* are defined by miRNAs that reciprocally crosstalk with trxG. In the third class, *miRNA-PcG/trxG shared circuits* are defined by miRNAs that crosstalk with both PcG and trxG regulators. These putative regulatory circuits may uncover a novel mechanism in *Drosophila* for the control of *PcG/trxG* and miRNAs levels of expression. The computational framework developed here for *Drosophila melanogaster* can serve as a model case for similar analyses in other species. Moreover, our work provides, for the first time, a new and useful resource for the *Drosophila* community to consult prior to experimental studies investigating the epigenetic regulatory networks of miRNA-PcG/trxG mediated gene expression.

1. Introduction

Regulation of gene expression is essential for the normal development and healthy life of organisms. Transcription is controlled at multiple levels and involves numerous factors and transcription cofactors that work together to provide a robust and adequate level of temporal and spatial gene expression. microRNAs (miRNAs) provide post-transcriptional control while the Polycomb Group (PcG) and the trithorax Group (trxG) of proteins function in the epigenetic regulation of transcription. miRNAs and PcG/trxG proteins are highly conserved in most eukaryotes including *Drosophila* and mammals, where they control gene expression of many genes participating in a range of biological processes (Schuettengruber et al., 2007; Shabalina and Koonin, 2008).

miRNAs are a class of small (19–24 nucleotides in length) RNAs that largely control transcription by binding to complementary sequences, known as MicroRNA Response Elements (MREs), located mainly within the 3′untranslated region (3′-UTR) of target mRNAs. Target recognition is primarily initiated through base pairing at the miRNA seed region (nucleotides 2–8 (Lewis et al., 2005), the minimal element required to engage the target mRNA. This interaction produces either mRNA translational repression or mRNA degradation, ultimately leading to a decrease in the amount of mRNA and a fine-scale tuning effect on protein levels (Baek et al., 2008). The magnitude of repression depends particularly on the complementarity between a miRNA and its target MRE.

PcG and trxG proteins are chromatin modifiers that maintain transcriptional states by controlling chromatin compaction and, thereby, accessibility to DNA by transcription factors and cofactors. They have counteracting activities: PcG represses gene expression and trxG activates

* Corresponding author.
** Corresponding author.
E-mail addresses: teresa.laguna@imdea.org (T. Laguna), abusturia@cbm.csic.es (A. Busturia).

https://doi.org/10.1016/j.ydbio.2022.12.008
Received 13 September 2022; Received in revised form 7 December 2022; Accepted 26 December 2022
Available online 31 December 2022

gene expression by either promoting a "closed" or an "open" chromatin configuration, respectively (reviewed in (Kassis et al., 2017)). PcG/trxG function as multimeric protein complexes, such as PRC1 (Polycomb Repressive Complex 1), PRC2 (Polycomb Repressive Complex 2) or the COMPASS-like complexes that bind to chromatin and modify its structure by either the covalent modification of nucleosome histones, such as methylation or ubiquitylation or by ATP-dependent chromatin remodeling (Kassis et al., 2017; Piunti and Shilatifard, 2021). Recruitment of PcG/trxG complexes to chromatin is essential for their function (Busturia et al., 1997, 2001). In vertebrates experimental evidence, although still controversial, is accumulating supporting CpG islands as targets of PcG/trxG complexes (Bauer et al., 2016; Li et al., 2017; Du et al., 2018; Owen and Davidovich, 2022). In *Drosophila*, it is well established that recruitment of PcG/trxG complexes to chromatin is mediated by binding to PREs (Polycomb Response Elements) and TREs (Trithorax Response Elements). PREs and TREs are DNA *cis*-regulatory elements that contain similar sequences, suggesting that PcG/trxG share genomic binding sites at target genes (Kassis et al., 2017; Schuettengruber et al., 2009; Schwartz et al., 2010; Kahn et al., 2014; Bredesen and Rehmsmeier, 2019). PREs contain clusters of transcription factors binding motifs, variable in number and position within the PRE (Kassis and Brown, 2013). Genome searches for PREs using a number of different PRE prediction tools (Bredesen and Rehmsmeier, 2019; Fiedler and Rehmsmeier, 2006; Khabiri and Freddolino, 2019; Ringrose et al., 2003; Zeng et al., 2012) indicate that PcG/trxG proteins may regulate the expression of roughly 30% of the *Drosophila* genes (Bredesen and Rehmsmeier, 2019).

Changes in miRNA and PcG/trxG expression levels contribute to various physiological disorders and are signatures of several types of pathologies, including cancer (Farazi et al., 2013; Chan and Morey, 2019; Fathi et al., 2021). Therefore, understanding the mechanisms controlling miRNAs and PcG/trxG levels is important for the definition of transcriptional networks controlling normal and pathological development.

A complex vertebrate regulatory network between the PcG and miRNAs gene expression has been described (Asangani et al., 2012; Bueno et al., 2008; de Nigris, 2016; Ning et al., 2015; Sato et al., 2011; Varambally et al., 2008; Wang et al., 2015; Xu et al., 2020) defining regulatory circuits that maintain each other's levels of expression. Of these circuits, the positive and negative feedback loops involving the *EZH2* (Enhancer of Zeste 2) protein of the PRC2 complex and several miRNAs are the most studied, and these have been shown to be involved in the development of a variety of pathologies (You et al., 2018; Liu et al., 2021; Jia et al., 2021; Cai et al., 2021). In *Drosophila*, less is known about possible interactions between PcG/trxG and miRNAs. It has been shown that *dme-miR-34* regulates the expression of PRC2 members, *Pcl* and *Su(z) 12*, to promote healthy brain development (Kennerdell et al., 2018). Moreover, a genome search for PREs in *Drosophila* S2 cells identified several miRNAs as PcG/trxG targets and the expression of these miRNAs is altered in *PcG* mutant backgrounds (Enderle et al., 2011).

In this study, we have investigated the regulatory interactions between miRNAs and PcG/trxG using an *in silico* approach to uncover the regulatory circuits in *Drosophila*. To this end, we have developed a computational framework based on refined prediction data generated using a combination of different predictors (Agarwal et al., 2018; Bredesen and Rehmsmeier, 2019; Khabiri and Freddolino, 2019) and experimental datasets (Wessels et al., 2019) (Fig. 1). First, we have analyzed whether each one of the miRBase v22 mature miRNAs described (Methods) showed *PcG* and *trxG* binding enrichment. Second, we have investigated the potential of each miRNA gene to be regulated by PcG/trxG through the presence of PREs in its promoter region. The results of these analyses have identified a novel and comprehensive set of miRNAs that may act as regulators of *PcG/trxG* gene expression as well as a set of miRNA genes potentially regulated by the PcG/trxG proteins. Interestingly, the combination of the results has uncovered cross-regulatory interactions between miRNAs and PcG/trxG that may define direct regulatory circuits. Our results provide a computational resource ready to be applied to the discovery of the molecular mechanisms involved in the biological processes mediated by the miRNAs and PcG/trxG networks in *Drosophila*. Moreover, the results presented here can be used as a foundation for future work to deepen our understanding on the intricate regulatory networks formed between these two large groups of transcriptional regulators.

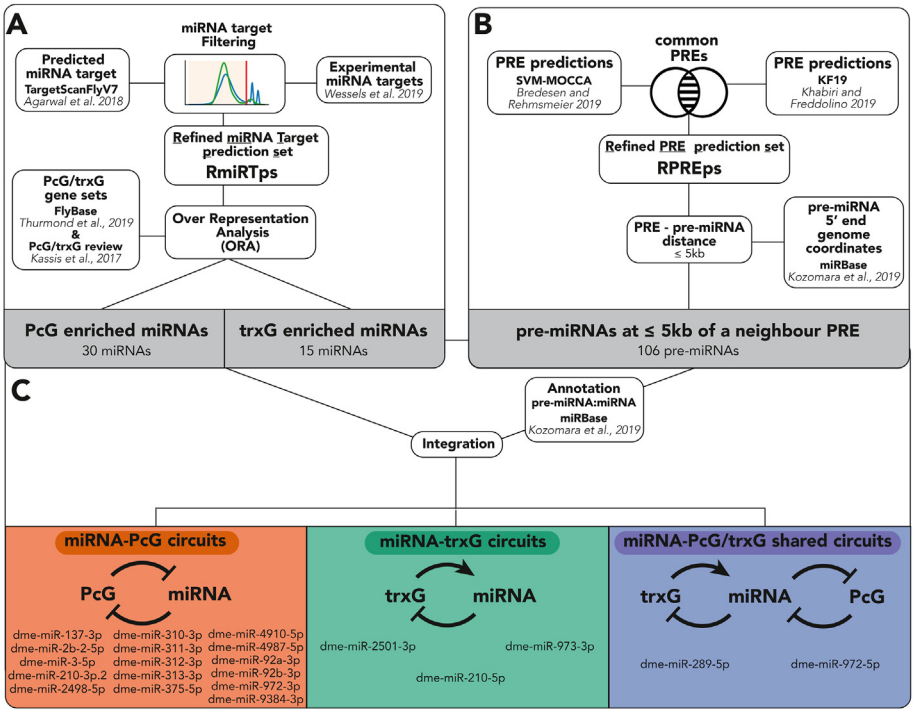

Fig. 1. Conceptual map of the computational framework A) Search for miRNA binding of PcG and trxG. TargetScanFly miRNA target predictions (Agarwal et al., 2018) and experimentally defined mature miRNA (Wessels et al., 2019) were paired and filtered to generate the Refined miRNA Target predictions set (RmiRTps). Next, an overrepresentation analysis (ORA) was applied to RmiRTps using the selected *PcG/trxG* gene sets, resulting in a list of miRNAs enriched in *PcG* or *trxG* transcripts predictive binding. **B) Search for pre-miRNAs with neighboring PREs.** Intersection of SVM-MOCCA (Bredesen and Rehmsmeier, 2019) and KF19 (Khabiri and Freddolino, 2019) predictions generated the Refined PRE prediction set (RPREps) to which the pre-miRNAs genomic coordinates were compared and filtered according to the selection criteria of ≤ 5 kb. This resulted in a list of pre-miRNAs having a PRE located at 5 kb or less from the pre-miRNAs 5′end. **C) Integration of A and B.** Following the annotation of the pre-miRNAs with the corresponding miRNAs, the combination of results from **A** and **B** generated a list of miRNAs forming circuits with PcG (16), with trxG (3) and with both PcG and trxG (2). See Methods and Fig. S1 for details.

2. Methods

2.1. Data sources

miRBase v22 (Kozomara et al., 2019) was accessed to obtain pre-miRNA genomic coordinates ("*dme_miRBase.gff3*") and to annotate the correspondence between pre-miRNAs and mature miRNAs ("*miRNA.dat*"). FlyBase (V2.0 (Thurmond et al., 2019), was consulted for the generation of the *PcG/trxG* gene sets. TargetscanFlyV7 resource database was used for miRNA target predictions (www.targetscan.org/fly_72/, (Agarwal et al., 2018). Three precomputed prediction files were downloaded: 1) the *Conserved site context scores prediction* file, 2) The *Non-conserved site context scores* prediction file, and 3) the *Summary Counts, all predictions* file. In Files 1 and 2 every site with homology to a mature miRNA seed region is classified depending on the MicroRNA Response Element (MRE) site's conservation and scored by the context++ score (a measure defined by 14 features described in (Agarwal et al., 2015)), representing the predicted MRE repression. In File 3, all the MRE predictions of a given miRNA in a given transcript are integrated using the most abundant mRNA isoform, thereby providing one prediction *per* gene. These predictions are scored by the Cumulative Weighted Context Score (CWCS), representing the predicted repression of a miRNA over a gene. Additionally, in File 3 the predictions of the miRNAs of the same *seed* are grouped in one miRNA family, selecting the highest CWCS of all the individual mature miRNAs as the family CWCS. Both CWCS and context++ score are represented with negative values: the lower the value, the stronger predicted repression. For more information see TargetScanFlyV7 website, (www.targetscan.org/fly_72/, (Agarwal et al., 2018).

For the miRNA target experimental data, the *High confidence miRNA binding site set* file from (Wessels et al., 2019) was used following the recommended cut-offs, microMUMMIE_var = 0.5, see Wessels et al., for details (for details see Supplementary Data 10 in (Wessels et al., 2019). This file contains the common MREs of two separate replicate samples of an *AGO1 PAR-CLIP* experiment and includes the 3′-UTR location of the MREs associated with the top 30 expressed miRNAs in the *Drosophila* S2 cell line.

SVM-MOCCA (Bredesen and Rehmsmeier, 2019) and the predictor developed by Khabiri and Freddolino (referred to here as KF19) (Khabiri and Freddolino, 2019) were used for Polycomb Response Elements (PREs) predictions. For SVM-MOCCA predictions, we selected the "Supplementary File 3" (Bredesen and Rehmsmeier, 2019), which contains genome-wide PRE predictions with a cut-off for an expected precision of 80% in *dm6* assembly. For KF19, the 0.8 cut-off PRE predictions with a precision of 98% and a recall of 0.26, in *dm5.9* assembly, were kindly provided by the authors. The KF19 predictions were later converted to *dm6* assembly with FlyBase Coordinate converter tool (dos Santos et al., 2015).

2.2. Generation of the Refined miRNA Target predictions set (RmiRTps)

A Refined microRNA Target predictions set (RmiRTps) was generated based on experimental data to apply a sensible threshold to the predictions (Fig. 1A and Fig. S1). An intersection strategy and several filtering steps were done from TargetScanFlyV7 predictions (www.targetscan.org/fly_72/ (Agarwal et al., 2018), and *Wessels* et al. experimentally defined MREs (Wessels et al., 2019) as follows. First, the TargetScanFlyV7 *Conserved* and *Nonconserved site context scores* files (Files 1 and 2, see above and Fig. S1) were merged and the *seed* genomic coordinates were calculated from the provided position within the 3′-UTR and the 3′-UTR genomic coordinates (the *UTR genome coordinate*, TargetScanFlyV7 website). Second, the overlap between these MRE predictions and the experimentally defined MREs was generated to explore to what extent the MRE predictions were experimentally supported. Third, the resulting context++ score distribution was analyzed and compared to the score distribution of the MRE predictions. Finally, a cut-off was empirically established in a context++ score of −0.04 to filter the TargetScanFlyV7 MRE predictions and a list of the "*miRNA:transcript pairs*' below the −0.04 context++ score was extracted.

Next, the condensed miRNA families from the *Summary Counts, all predictions* file (file 3 from (Agarwal et al., 2018), see above and Fig. S1) were expanded to cover all individual miRNAs and the predictions were filtered with the list of *miRNA:transcript pairs* below the established cut-off. Only the summarized predictions included in the *miRNA:transcript pairs* list were selected, and thus excluded all the summarized predictions that were exclusively formed by MRE predictions not fulfilling the −0.04 context++ score cut-off. TargetScanFlyV7 *Summary Counts, all predictions* considers only the main transcript isoform of each gene, thus the last filtering step resulted in the RmiRTPs with unique "*miRNA:gene pairs*".

2.3. PcG/trxG gene set definition

For the definition of the *PcG* gene set, the *PcG* genes considered in Kassis et al. review (Kassis et al., 2017) with a proven *in vivo* functionality were used. For the definition of the trxG gene set the "common" genes defined as "trxG genes" in FlyBase (https://flybase.org/reports/FBgg0000303.html (Thurmond et al., 2019),) in Kassis et al. review (Kassis et al., 2017) were selected. Additionally, among the "non common genes" only those with a proven *in vivo* functionality described in Kassis et al. review and Ray et al., 2016, Brown and Kassis, 2010 were included. This resulted in the inclusion of 22 *PcG* genes for the *PcG* gene set and 26 *trxG* genes for the *trxG* gene set (Table S1).

2.4. PcG/trxG over representation analysis

The application of an over-representation analysis (ORA) allows to test whether binding of a given miRNA to a given set of transcripts is overrepresented, *i.e.* enriched in the context of all the miRNA predicted binding sites. One-sided Fisher exact test was used for testing the miRNAs predicted to target at least one *PcG* or one *trxG* transcript respectively, to obtain those miRNAs enriched in *PcG* or *trxG*. The Odds Ratio (OR) represents the statistical probability of a given miRNA binding to a *PcG* or a *trxG* transcript within all the given miRNA targets. To evaluate the false discovery rate of the ORA results, the robust *False Discovery rate* method (*rFDR*) was used with one sided continuous *p*-values parameters (Pounds and Cheng, 2006) using the "robust.fdr" function nested inside "*prot2D*" Bioconductor package (Artigaud et al., 2013). Only *rFDRs* of 22% for trxG, and 10% for PcG (corresponding with p-values ≤0.01) were considered, as the main goal of this study is to provide a *bonafide* list of candidate miRNAs for their *in vivo* experimental validation.

2.5. Generation of the refined PRE predictions set (RPREps)

The Refined PRE predictions set (RPREps) was produced by intersecting SVM-MOCCA (Bredesen and Rehmsmeier, 2019) and KF19 (Khabiri and Freddolino, 2019) PRE predictions (Fig. 1B, Fig. S2). The goal of the intersection was to improve the confidence of PRE predictions, given the technical and biological differences between KF19 and SVM-MOCCA, such as the biological sample (isolated mesodermal embryonic cells *versus SG4* embryonic cells), the technical experimental procedure of the training set (ChIP-seq *versus* ChIP-chip), the model features and the model construction. KF19 predictions were converted to the *dm6* assembly using *FlyBase Coordinate converter tool* (dos Santos et al., 2015) and then used *BEDTools* (Quinlan and Hall, 2010) to obtain the overlap regions of the intersecting KF19 and SVM-MOCCA PRE predictions, resulting in the RPREps (Fig. S2).

2.6. Association of each miRNA with the closest PRE

Following KF19 criteria (Khabiri and Freddolino, 2019), a distance of ≤ 5 kb between the PREs of the RPREps and the pre-miRNA *5'* end from

miRBase v22 (Kozomara et al., 2019) was used to select the pre-miRNAs potentially regulated by PcG/trxG proteins. Of those, we found that some pre-miRNA contained a PRE within the pre-miRNA transcription unit and were classified as overlap (Fig. 4).

2.7. Integration of the miRNAs enriched in PcG/trxG transcripts and the pre-miRNAs potentially regulated by PcG/trxG proteins

pre-miRNAs at a distance of ≤ 5 kb to a PRE were annotated with their corresponding mature miRNAs using the miRBase v22 "*miRNA.dat*" file (Kozomara et al., 2019) and then integrated with the miRNA-PcG/trxG ORA results. This resulted in a list of mature miRNAs forming the regulatory circuits with PcG or trxG proteins (Fig. 1C). *dme-miR-10404* and *dme-miR-11182* were excluded due to differences between the miRBase version utilized in this analysis and TargetScan-FlyV7 predictions (Agarwal et al., 2018).

A network was produced to show the relation between miRNAs, PcG and trxG by exploring the individual PcG/trxG targets of each miRNA using Cytoscape (Shannon et al., 2003) (Fig. 5). For this, miRNAs were considered as source nodes and *PcG/trxG* transcripts as target nodes. Edges between source and target nodes are drawn based on RmiRTPs, considering the expected repression (*i.e.*, CWCS) of the predictions as the edge weight (the thicker the edge the stronger the expected repression). For representation of Fig. 5C, source nodes were filtered to keep only miRNAs involved in regulatory circuits (See Data availability).

2.8. Data availability

The original files, as well as the RmiRTps and RPREps and all the code used in this study are available at: https://github.com/j-solor/Drosophila-miRNA-PcG-circuits/

Network available for analysis at NDEx:

https://www.ndexbio.org/#/networkset/998959df-1a4e-11ed-ac45-0ac135e8bacf?accesskey=6e232187b6745a196eca1d0321b2f8befd6dbd737e83a4407e6c9ce29e9842a0.

3. Results

3.1. Search for mature microRNAs potentially regulating PcG/trxG gene expression

To search for miRNAs potentially regulating *PcG* and *trxG* gene expression, we studied all of the 469 miRNAs so far identified in the *Drosophila* genome. We generated a Refined miRNA Target predictions set (RmiRTps) with 410,855 miRNA*:Gene* interactions for a total of 13,212 genes and an average number of ~31 mature miRNAs targeting each gene (Fig. 1A). This curated miRNA Target prediction set was produced by matching TargetScanFlyV7 predictions (Agarwal et al., 2018) with the transcriptome-wide map of miRNAs target sites developed by *CLIP*-seq (crosslinking and immunoprecipitation followed by sequencing) experiments in a *Drosophila* S2 cell line (Wessels et al., 2019). This strategy allowed a reduction in the number of spurious and residual miRNA predictions (Fig. S1, Methods), resulting in the selection of only those miRNAs with at least one *PcG* or one *trxG* transcript target. Therefore, 343 miRNAs were tested for binding enrichment in PcG targets and 335 miRNAs for binding enrichment in trxG targets.

Next, we sought to identify the miRNAs enriched in the putative regulation of *PcG/trxG* genes by applying an overrepresentation analysis (ORA, Methods) to test whether a miRNA could interact more specifically with either *PcG* or *trxG* than with any other protein-coding gene sequences (Fig. 1A). To do this, we first defined the *PcG/trxG* gene sets (Methods) that included 22 *PcG* genes and 26 *trxG* genes respectively (Table S1). The application of ORA to search for the miRNAs enriched in *PcG* regulation resulted in 35 out of 343 mature miRNAs predicted to significantly target *PcG* genes (*p*-value <0.01). Additionally, the application of ORA to find the miRNAs enriched in the putative regulation of *trxG*, resulted in 15 out of the 335 mature miRNAs predicted to significantly target *trxG* genes (*p*-value <0.01). Out of these 50 (35 + 15) miRNAs, 6 target both *PcG* and *trxG* and thus, resulting in a total of 39 enriched miRNAs (Fig. 1A, Table S2). As shown in Fig. 2, there are miRNAs showing high Odds Ratio (OR) values for *PcG* enrichment such as *dme-mir-8-5p* (OR = 6.02, p-value = 6.6×10^{-4}), miRNAs showing high OR for enrichment in both *PcG* and *trxG* such as *dme-mir-2283-5p* (OR = 7.59, p-value = 1.0×10^{-5} for PcG; OR = 6.09, p-value = 1.2×10^{-5} for trxG) and miRNAs showing high OR for *trxG* enrichments such as *dme-mir-973-3p* (OR = 5.89, p-value = 3.3×10^{-3}) or *dme-mir-963-5p* (OR = 7.40, p-value = 1.3×10^{-3}). Notably, our results predict that *dme-mir-34-5p* targets *Pcl*, *Su(z)12, ph-d* and *Jing* transcripts (Fig. 3) which is in accordance to the results indicating that *dme-mir-34-5p* targets, *in vivo*, the *Pcl* and *Su(z)12* transcripts (Kennerdell et al., 2018), thus validating our methods.

In summary, we identified miRNAs enriched in *PcG* binding and 15 miRNAs enriched in *trxG* binding, thus suggesting potentially regulatory interactions. Moreover, taking into account that miRNA binding enrichment analysis was performed with 343 miRNAs for *PcG* and 335 microRNAs for *trxG*, our results suggest that roughly 10% and 3% of the mature miRNAs putatively regulate *PcG* and *trxG* genes, respectively.

3.2. Search for pre-miRNAs potentially regulated by PcG/trxG proteins

To identify pre-miRNAs likely to be regulated by PcG/trxG proteins we search for the presence of PREs in the vicinity of pre-miRNA genes. As PREs serve as binding sites for both PcG and trxG proteins to regulate chromatin structure (Kassis et al., 2017), the outcome of this search provides a list of pre-miRNAs predicted to be regulated by both the PcG and trxG proteins. We generated a Refined PRE predictions set (RPREps) to obtain higher confidence in the predicted PREs. For that, we selected the common PRE predicted sites by SVM-MOCCA (Bredesen and Rehmsmeier, 2019) and by the predictor developed by Khabiri and Freddolino (2019), referred in the text as KF19, for higher confidence in the predicted PREs. This search identified 4258 PREs of 759bp average length in the RPREps (Fig. S2). Next, we calculated the distance between each pre-miRNA and its closest PRE using the coordinate information of the 260 pre-miRNAs of the miRBase database (Kozomara et al., 2019) and our RPREps. We considered a miRNA to be potentially regulated by PcG/trxG if the 5'end of the pre-miRNA was located 5 kb or less from the closest PRE, a similar strategy to that was used in Khabiri and Freddolino (2019). We found that 40% (106 out of 260) of the pre-miRNAs are potentially regulated by PREs (Fig. 4, Table S2). Of note, 16 of the 106 pre-miRNAs overlap with a PRE, *i.e.*, contain a PRE within the pre-miRNA transcription unit (Fig. 4). This is the case for *dme-miR*-313 that together with *dme-miRNA-310*, *dme-miRNA-311* and *dme-miRNA-312* form the *dme-miR-310s* cluster (Ryazansky et al., 2011), suggesting a coordinated control of all these miRNAs by PcG/trxG proteins. These findings suggest an epigenetic mechanism of expression control by PcG/trxG proteins for nearly half of the pre-miRNAs.

3.3. Analysis of miRNA-PcG/trxG crosstalk interactions: Definition of regulatory circuits

Our results indicated that PcG/trxG proteins could interact with 40% of pre-miRNAs (Fig. 4), and that 50 mature miRNAs are enriched in *PcG/trxG* transcript binding (Fig. 2). This led us to search for the presence of crosstalk interactions that could define regulatory circuits between miRNAs and PcG/trxG.

First, we selected those miRNAs that contained a PRE in their vicinity (≤5 kb) (Fig. 4) and that additionally showed a significant enrichment of its target sequence in *PcG* transcripts (Fig. 2). The results showed 16 crosstalk interactions defining potential *miRNA-PcG regulatory circuits* (Fig. 5A, Table 1). For example, the regulatory circuit formed by the *dme-miRNA-313-3p* (and the entire *miR-310s* cluster) with the *Pc*, *Su(z)2* or *pho* transcripts. *dme-miRNA-313* "overlaps" (Fig. 5AB) with the

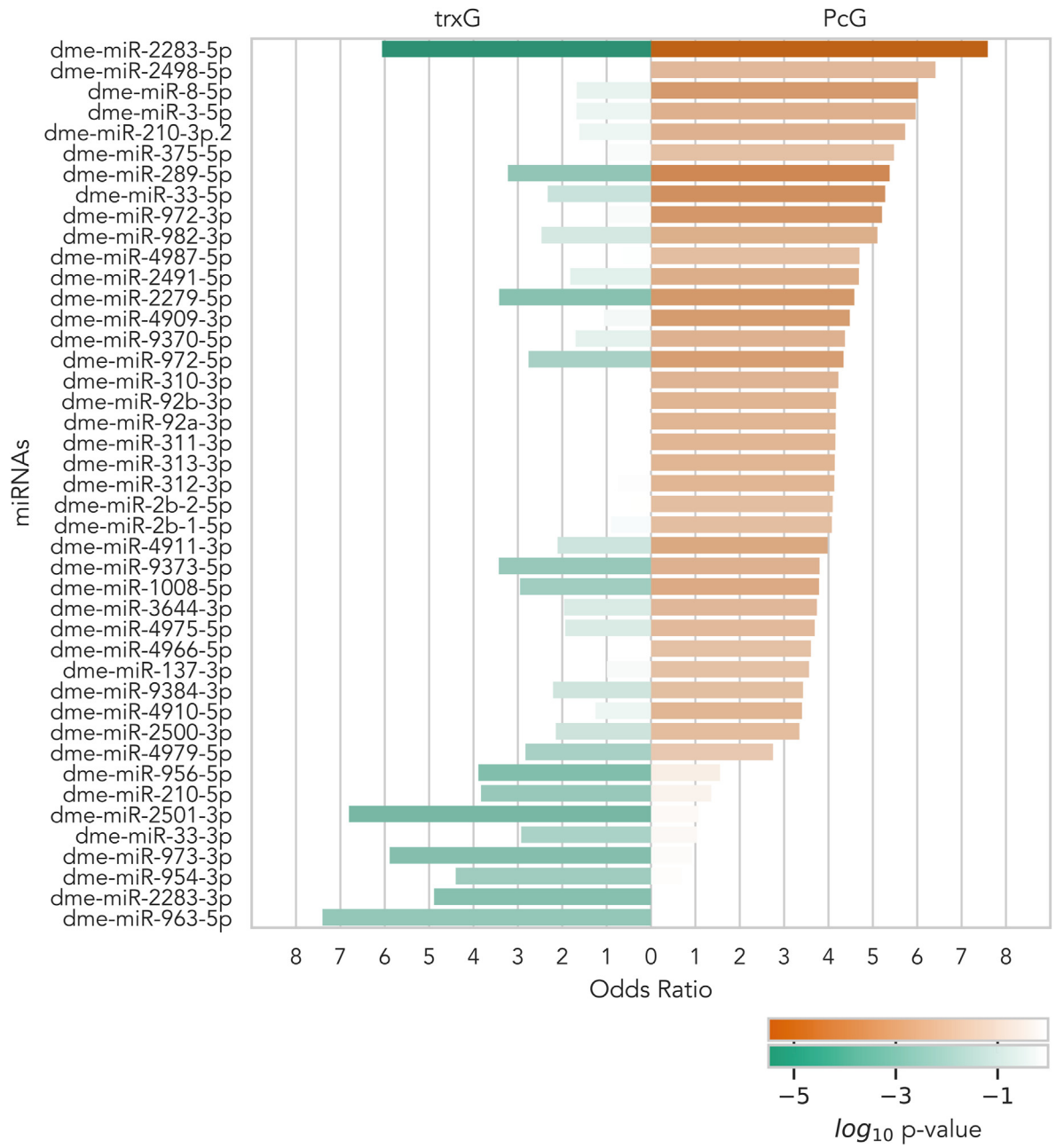

Fig. 2. Over-representation analysis of miRNAs binding PcG/trxG Barplot showing the Odds Ratio (OR) of the miRNAs significantly enriched in PcG and trxG as the x-axis. OR represents the probability of a given miRNA binding to a *PcG* or a trxG transcript within all the given miRNA targets. *p*-values are represented as the color gradient in a logarithmic scale, $\log_{10}$ p-values ≤ -2 are significant. Brown color gradient for PcG enrichment and green color gradient for trxG enrichment.

regulatory predicted PRE and shows significant enrichment for *PcG* transcript binding (OR = 4.14, *p*-value = 4.7 x 10^{-3}), with predicted targets in *Pc* (Cumulative Weighted Context Score (CWCS) = −0.606), *Su(z)*$_2$ (CWCS = −0.342) and *pho* (CWCS = −0.481) (Figs. 3 and 5B), which all code for PcG proteins members of the PRC1, PRC2, PhoRC, and dRAF complexes involved the establishment of repressive transcriptional states (Kassis et al., 2017).

Second, we applied the same strategy to search for miRNA-trxG crosstalk interactions. The results showed 3 potential *miRNA-trxG regulatory circuits* (Fig. 5A, Table 1), for example, the circuit formed by *dme-miR-210-5p* with the *trx*, *osa* or *Kis* transcripts. *dme-miR-210-5p* is located at 1952 bp from the predicted regulatory PRE (Fig. 5AB) and showed significant enrichment for binding to trxG transcripts (OR = 3.83, p-value = 3.9 x × 10^{-3}), with predicted targets in *trx* (CWCS = −0.921), *osa* (CWCS = −1.166) and *Kis* (CWCS = −1.474) (Figs. 3 and 5B) coding for trxG proteins members of the TAC1, COMPASS-TRX, ASH1, BAP, PBAP, KIS complexes all of which are involved in the establishment of active transcriptional states (Kassis et al., 2017).

Finally, we selected those miRNAs that contained a PRE in their vicinity (≤5 kb) (Fig. 4) and that show a significant enrichment of its targets both in *PcG* and in *trxG* transcripts (Fig. 2). Interestingly, the combined results showed crosstalk interactions between miRNAs and both PcG and trxG thus defining 2 "*shared*" regulatory circuits. We identified *dme-miR-972-5p* and *dme-miR-289-5p* as potential miRNAs involved in these circuits (Fig. 5A, Table 1). For instance, *dme-miR-972-5p* is located 635bp from a predicted regulatory PRE and shows significant enrichment for *PcG* transcripts (*OR* = 4.72, p-value = 8.1 x 10-^{4}) with predicted targets in 15 out of 22 *PcG* (Figs. 3 and 5AB). Moreover, *dme-miR-972-5p* shows significant enrichment for trxG transcripts (*OR* = 2.76, p-value = 8.5x 10^{-3}) with predicted targets in 15 out of 26 trxG

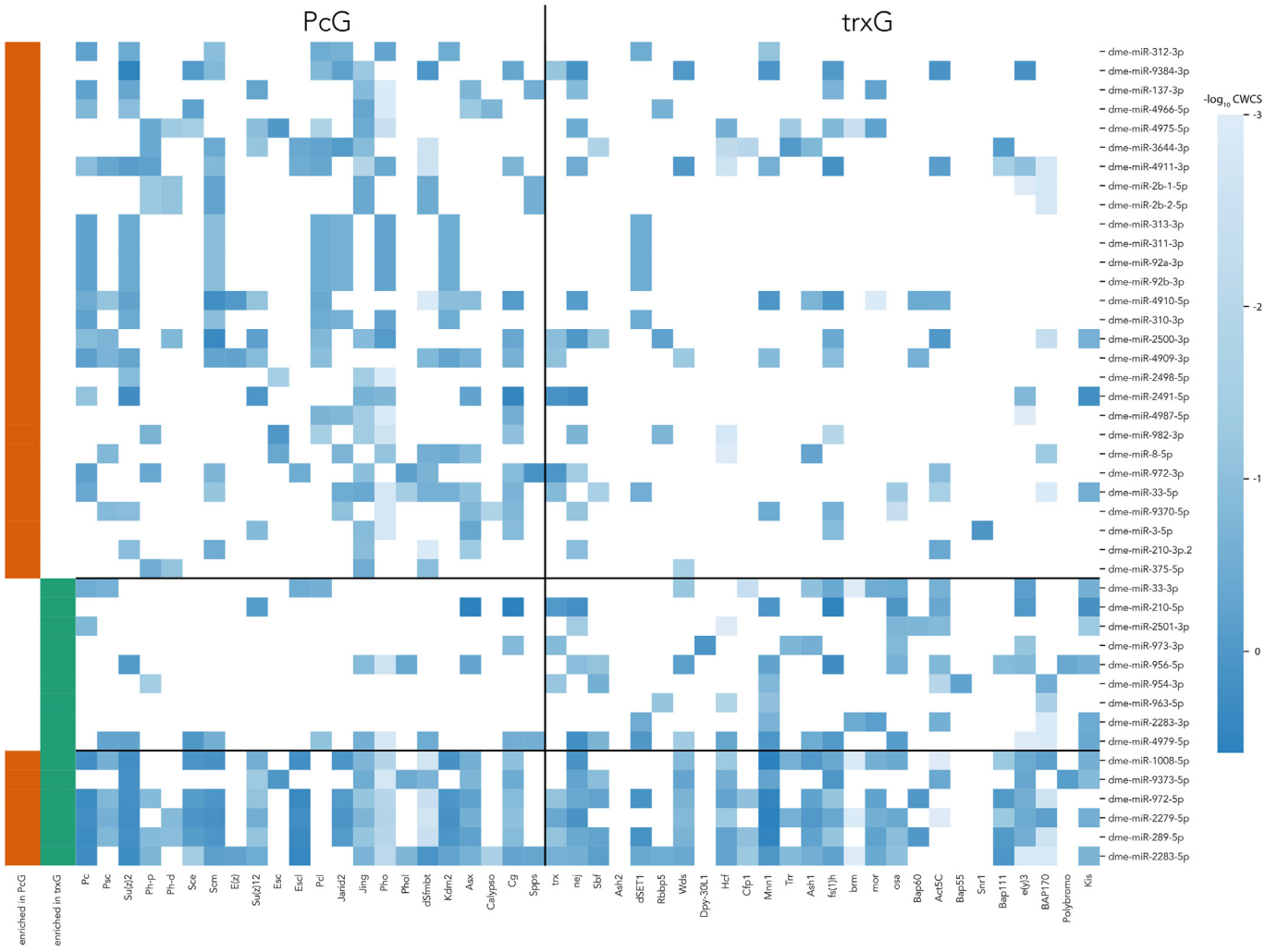

Fig. 3. Predicted PcG/trxG targets of the significantly enriched miRNAs Heatmap indicating the Cumulative weighted context score (CWCS) shown in logarithmic scale (blue color gradient). This represents the predicted strength of repression of a given miRNA (as specified) on a given *PcG* or *trxG* transcript (as specified). Orange bar (left) indicates the miRNAs significantly enriched in *PcG* binding. Green bar (left) indicates the miRNAs significantly enriched in *trxG* binding.

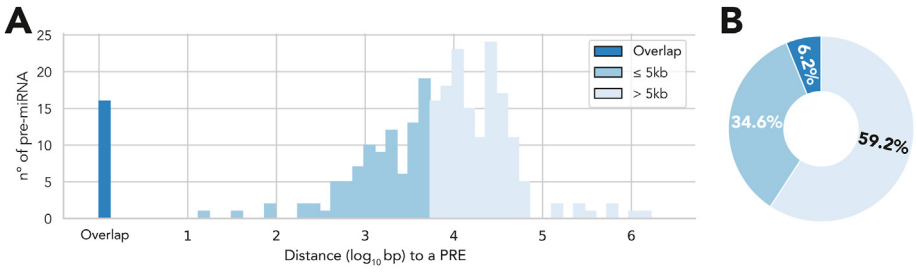

Fig. 4. Distribution of the pre-miRNAs according to their distance from the closest PRE A) Representation of the number of pre-miRNAs (260 total) and their 5'end distance (in logarithmic scale) to the closest PRE. Shown are the number of pre-miRNAs containing a PRE within the pre-miRNA transcription unit (overlap), located at ≤ 5 kb and located at >5 kb. **B)** Pie chart representing the percentage of the three indicated classes of pre-miRNAs.

transcripts (Figs. 3 and 5B).

Thus, our results identified 21 potential regulatory circuits between miRNAs and PcG/trxG in which members could function to control each other's levels of expression resulting in a complex regulatory network (Fig. 5C). Of those, 16 are formed by crosstalk interactions between miRNAs and PcG, 3 by crosstalk interactions between miRNAs and trxG and 2 between miRNAs and both PcG and trxG (shared regulatory circuits).

4. Discussion

A challenge in current biological studies is the identification of regulatory networks controlling the molecular and cellular mechanisms directing homeostatic levels of gene expression during normal development and upon stress stimuli. Regulatory network identification can be simplified by analyzing organisms with relatively simple genomes such as *Drosophila melanogaster*. In addition, genome-wide approaches allow the investigation of the functional interdependence between selected

Table 1
Description of the miRNAs involved in the regulatory circuits. In the "PcG genes" and "trxG genes" columns, the Cumulative Weighted Context Score is displayed for the predicted miRNA targets.

mature miRNA	pre-miRNA	Distance to PRE (bp)	Location closest PRE	PcG enrichment odds ratio	PcG enrichment p-valuie	trxG enrichment odds ratio	trxG enrichment p-value	PcG genes [Cummulative Weighted Context score]	trxG genes [Cummulative Weighted Context score]
dme-miR-4987-5p	dme-mir-4987	1625	2L:10842101..10842650	4.69826272477903	0.008129240404036	0.635257731958762	0.795681983020144	Pcl [-0.196], Jarid2 [-0.163], Jing [-0.037], Pho [-0.001], Cg [-0.276]	e(y)3 [-0.000]
dme-miR-2b-2-5p	dme-mir-2b-2	1187	2L:19571451..19572350	4.09633757961783	0.008086165051436	0.43463768115942	0.898603414179656	Ph-p [-0.057], Ph-d [-0.077], Scm [-0.513], Jing [-0.383], dSfmbt [-0.350], Spps [-0.262]	BAP170 [-0.002]
dme-miR-4910-5p	dme-mir-4910	870	2L:4818751..4819250	3.40322580645161	0.004662987082664	1.25114800423878	0.377504747288209	Pc [-0.299], Psc [-0.079], Su(z)2 [-0.555], Scm [-1.677], E(z) [-0.947], Su(z)12 [-0.079], Pcl [-0.394], dSfmbt [-0.001], Kdm2 [-0.105], Asx [-0.118], Cg [-0.827]	nej [-0.783], Mnn1 [-1.030], Ash1 [-0.192], fs(1)h [-1.189], mor [-0.001], Bap60 [-0.363], Act5C [-0.338]
dme-miR-375-5p	dme-mir-375	4942	2L:852051..852600	5.47640791476407	0.009760365974831	0.979455252918287	0.647137457593319	Ph-p [-0.306], Ph-d [-0.075], Jing [-0.302], dSfmbt [-0.115]	Wds [-0.026]
dme-miR-137-3p	dme-mir-137	1586	2R:16066930..16067395	3.56179183135704	0.009925710209183	0.992229903445123	0.595722445208553	Pc [-0.577], Su(z)2 [-0.435], Su(z)12 [-0.463], Jing [-0.133], Pho [-0.001], Asx [-0.180], Spps [-0.273]	nej [-0.133], fs(1)h [-0.205], mor [-0.530]
dme-miR-3-5p	dme-mir-3	2018	2R:19658896..19659645	5.96256684491978	0.003138678689138	1.68026386645126	0.350251823728148	Su(z)12 [-0.133], Jing [-0.123], Pho [-0.002], Asx [-0.441], Cg [-0.122]	fs(1)h [-0.123], Snr1 [-1.159]
dme-miR-2498-5p	dme-mir-2498	75	2R:20584196..20584745	6.41103518096684	0.005775287942213			Su(z)2 [-0.125], Esc [-0.032], Jing [-0.039], Pho [-0.001]	
dme-miR-310-3p	dme-mir-310	356	2R:20584196..20584745	4.22290388548057	0.004323871839355	0.36	0.935339409176805	Pc [-0.606], Su(z)2 [-0.342], Scm [-0.103], Pcl [-0.321], Jarid2 [-0.244], Pho [-0.481], Kdm2 [-0.329]	dSET1 [-0.313]
dme-miR-311-3p	dme-mir-311	236	2R:20584196..20584745	4.15555555555555	0.00468703052517	0.354281414597441	0.937992048476882	Pc [-0.606], Su(z)2 [-0.342], Scm [-0.103], Pcl [-0.321], Jarid2 [-0.244],	dSET1 [-0.313]

(*continued on next page*)

Table 1 (*continued*)

mature miRNA	pre-miRNA	Distance to PRE (bp)	Location closest PRE	PcG enrichment odds ratio	PcG enrichment p-valuie	trxG enrichment odds ratio	trxG enrichment p-value	PcG genes [Cummulative Weighted Context score]	trxG genes [Cummulative Weighted Context score]
								Pho [-0.481], Kdm2 [-0.329]	
dme-miR-312-3p	dme-mir-312	75	2R:20584196..20584745	4.13122807017543	0.004826604182234	0.734394506866416	0.758521611488091	Pc [-0.606], Su(z)2 [-0.342], Scm [-0.103], Pcl [-0.321], Jarid2 [-0.244], Pho [-0.481], Kdm2 [-0.329]	dSET1 [-0.313], Mnn1 [-0.075]
dme-miR-313-3p	dme-mir-313	0	2R:20584196..20584745	4.14162270786234	0.004766399534726	0.353098274568642	0.938537008793288	Pc [-0.606], Su(z)2 [-0.342], Scm [-0.103], Pcl [-0.321], Jarid2 [-0.244], Pho [-0.481], Kdm2 [-0.329]	dSET1 [-0.313]
dme-miR-289-5p	dme-mir-289	2329	3L:13623151..13623700	5.37753222836095	0.000172688985084	3.22504604051565	0.002823276295614	Pc [-2.046], Psc [-0.160], Su(z)2 [-1.605], Ph-p [-0.134], Ph-d [-0.108], Sce [-0.785], Scm [-1.068], Su(z)12 [-0.080], Escl [-2.322], Jarid2 [-0.818], Jing [-0.068], Pho [-0.010], dSfmbt [-0.005], Kdm2 [-1.050], Asx [-0.620], Cg [-0.124]	trx [-0.075], nej [-0.403], Sbf [-0.521], dSET1 [-1.334], Wds [-0.216], Hcf [-0.531], Cfp1 [-0.092], Mnn1 [-3.006], Ash1 [-0.563], fs(1)h [-0.070], mor [-0.576], osa [-0.103], Bap60 [-0.976], Bap111 [-0.872], e(y)3 [-0.277], BAP170 [-0.004]
dme-miR-2501-3p	dme-mir-2501	1137	3L:8109551..8109970	1.07066052227342	0.616431988714722	6.80669077757685	0.000644684384588	Pc [-0.161]	nej [-0.041], Hcf [-0.000], osa [-0.184], Bap60 [-0.190], Act5C [-0.144], Kis [-0.049]
dme-miR-92a-3p	dme-mir-92a	423	3R:25646929..25647678	4.16255362099419	0.004647730883487	0.354875659382064	0.937717829011951	Pc [-0.606], Su(z)2 [-0.342], Scm [-0.103], Pcl [-0.321], Jarid2 [-0.244], Pho [-0.481], Kdm2 [-0.329]	dSET1 [-0.313]
dme-miR-92b-3p	dme-mir-92b	3070	3R:25647829..25648328	4.16957290876926	0.004608686337308	0.355471698113207	0.937442443701222	Pc [-0.606], Su(z)2 [-0.342], Scm [-0.103], Pcl [-0.321], Jarid2 [-0.244], Pho [-0.481], Kdm2 [-0.329]	dSET1 [-0.313]
dme-miR-9384-3p	dme-mir-9384	0	3R:4748529..4749078	3.42643581725268	0.008613674436318	2.20698758636184	0.068058853628999	Su(z)2 [-3.392], Sce [-0.811], Scm [-0.154], Pcl [-0.172],	trx [-0.071], nej [-1.342], Wds [-0.907], Mnn1 [-1.206], fs(1)h

(*continued on next page*)

Table 1 (*continued*)

mature miRNA	pre-miRNA	Distance to PRE (bp)	Location closest PRE	PcG enrichment odds ratio	PcG enrichment p-valuie	trxG enrichment odds ratio	trxG enrichment p-value	PcG genes [Cummulative Weighted Context score]	trxG genes [Cummulative Weighted Context score]
								Jarid2 [-0.603], Jing [-0.058], dSfmbt [-1.039], Cg [-0.448]	[-0.881], Act5C [-1.277], e(y)3 [-1.468]
dme-miR-210-3p.2	dme-mir-210	1952	X:18125718..18126267	5.72794117647058	0.00369715955157	1.6144375324002	0.367768836018748	Su(z)2 [-0.110], Scm [-0.300], Jing [-0.120], dSfmbt [-0.001], Asx [-0.049]	nej [-0.110], Act5C [-0.415]
dme-miR-210-5p	dme-mir-210	1952	X:18125718..18126267	1.35567704924344	0.407210783239273	3.83031489639419	0.003858869992258	Su(z)12 [-0.758], Asx [-3.526], Cg [-3.578]	trx [-0.921], nej [-1.609], Mnn1 [-1.025], fs(1)h [-3.397], osa [-1.166], Act5C [-0.439], e(y)3 [-0.932], Kis [-1.474]
dme-miR-972-3p	dme-mir-972	635	X:19549968..19550517	5.20626187856356	0.00050721606566	0.976233385897724	0.605956509336639	Pc [-0.653], Ph-p [-0.422], Scm [-0.180], Escl [-0.328], Jing [-0.102], Phol [-0.622], dSfmbt [-0.200], Cg [-0.127], Spps [-1.325]	trx [-1.072], nej [-0.042], Act5C [-0.108]
dme-miR-973-3p	dme-mir-973	1011	X:19549968..19550517	0.927380952380952	0.667920256535042	5.88897637795275	0.001318090013727	Cg [-0.147]	trx [-0.140], Dpy-30L1 [-1.176], Trr [-0.178], Ash1 [-0.261], osa [-0.152], e(y)3 [-0.068]
dme-miR-972-5p	dme-mir-972	635	X:19549968..19550517	4.33913731697665	0.000805666185207	2.76001007302946	0.008545496503637	Pc [-2.046], Psc [-0.160], Su(z)2 [-1.605], Ph-p [-0.134], Sce [-0.785], Scm [-1.068], Su(z)12 [-0.080], Escl [-2.322], Jarid2 [-0.818], Jing [-0.068], Pho [-0.010], dSfmbt [-0.005], Kdm2 [-1.050], Asx [-0.620], Cg [-0.124]	trx [-0.075], nej [-0.403], Sbf [-0.521], dSET1 [-1.334], Wds [-0.216], Hcf [-0.531], Cfp1 [-0.092], Mnn1 [-3.006], Ash1 [-0.563], fs(1)h [-0.070], mor [-0.576], Bap60 [-0.976], Bap111 [-0.872], e(y)3 [-0.277], BAP170 [-0.004]

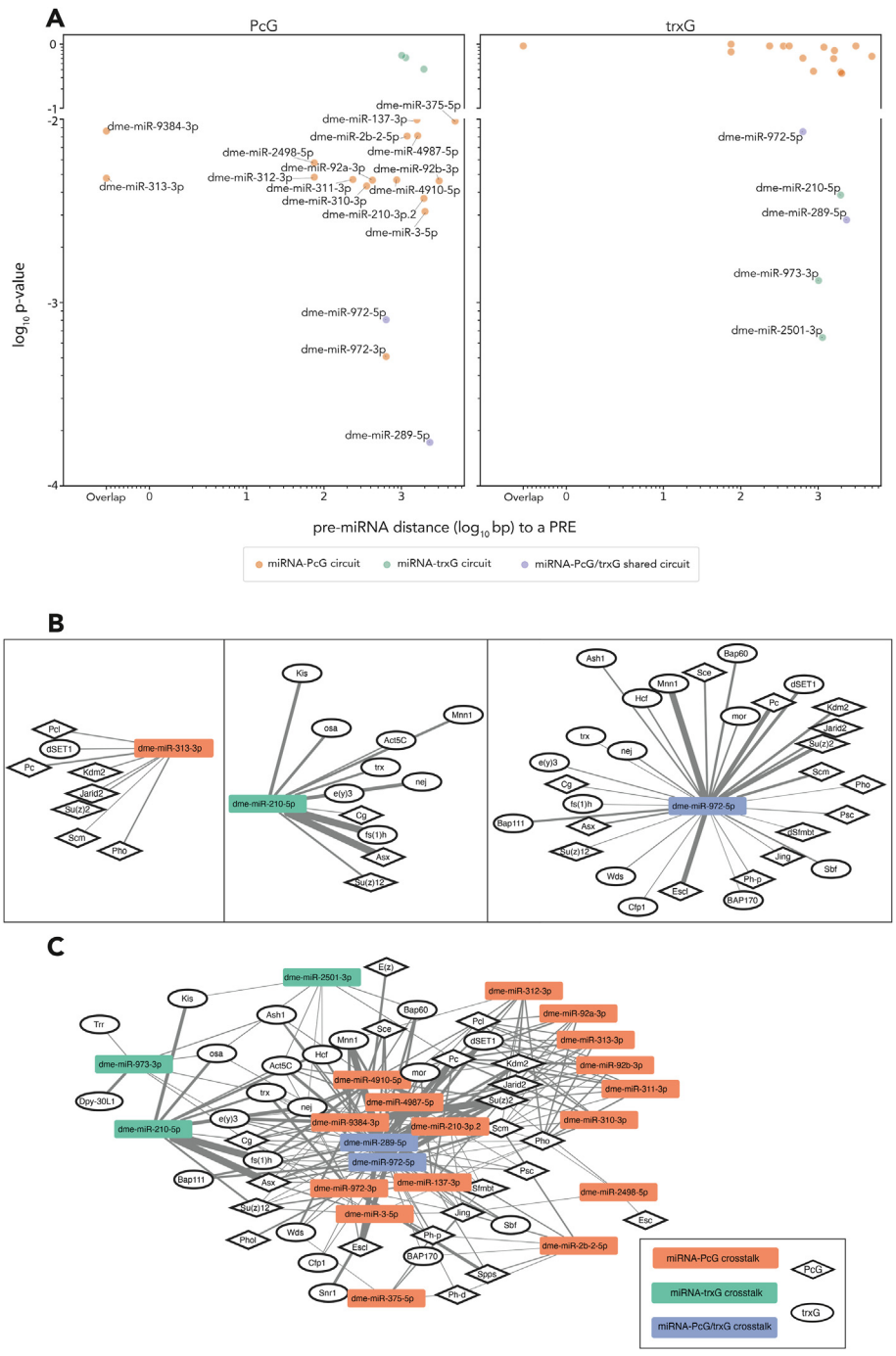

Fig. 5. miRNAs defining PcG/trxG regulatory circuits A) Representation of the miRNA enrichment p-values in relation to the distance to the closest PRE. Left panel: miRNAs (orange dots) defining *miRNA-PcG circuits* with a significant p-value. Right panel: miRNAs (green dots) forming *miRNA-trxG* circuits with a significant p-value. Also indicated are the miRNAs with a significant p-value (purple dots) forming *miRNA-PcG/trxG shared circuits*. p-values are in a logarithmic scale in y-axis, $\log_{10}$ p-values ≤ -2 are significant. In the x-axis, pre-miRNA distance to PRE is represented in the logarithmic scale. **B) Network representations of miRNA-PcG/trxG circuits**. Panels from left to right show examples of *miRNA-PcG circuit* (*dme-miR-313-3p* network, left), *miRNA-trxG circuit* (*dme-miR-210-5p* network, middle), *miRNA-PcG/trxG circuit* (*dme-miR-972-5p* network, right). **C) Network representation with the integration of the 16 miRNAs forming all three classes of regulatory circuits**. In panels B and C color rectangles nodes represent miRNAs; diamond and ellipse nodes represent target transcripts belonging to the *PcG* and *trxG* respectively. Edge weight represents the Cumulative Weighted Context Score (CWCS) of each target prediction.

components and the complex relationships between them.

We have developed an *in silico* genome-wide framework in *Drosophila* (Fig. 1) to explore the existence of gene regulatory circuits between two large groups of regulators: the *Polycomb* and *trithorax* Groups genes (*PcG/trxG*) and the microRNAs (miRNAs), relevant for organismal development and human disease pathologies. Several studies have reported interactions between these groups of regulators (Asangani et al., 2012; Bueno et al., 2008; Ning et al., 2015; Varambally et al., 2008), in some cases finding a reciprocal control through crosstalk interactions thus defining regulatory circuits between vertebrate PcG and miRNAs (de Nigris, 2016; Cao et al., 2011; Schubert et al., 2013; Jiang et al., 2019). However, the existence of such regulatory circuits in *Drosophila* has not yet been explored.

In this work, we searched for miRNA target enrichment for *PcG/trxG* transcripts as well as for Polycomb Response Elements (PREs) in the vicinity of pre-miRNAs. The intersection and combination of these analyses resulted in the predicted existence of regulatory circuits, providing a high value resource for the experimental analysis of these regulatory *Drosophila* networks.

Our results identify 35 miRNAs as potential regulators of *PcG* gene expression and 15 miRNAs as potential regulators *trxG* gene expression (Figs. 2 and 3, Table S2). Considering that a total of 343 and 335 miRNAs were included in our study to analyze *PcG* and *trxG* enrichment, respectively, our results indicate that roughly 10% and 3% of the miRNAs may regulate *PcG* and *trxG* gene expression. Moreover, our results indicate that 40% of pre-miRNAs (106 out of 260 pre-miRNAs) could be potentially regulated by PcG/trxG proteins through the binding to PREs located in their vicinity (Fig. 4, Table S2). This percentage is higher than

in a previous genome-wide analysis performed in *Drosophila* S2 cells where it was reported that 41 out of the 157 pre-miRNAs described at the time were potentially regulated by neighboring PREs (Enderle et al., 2011).

The intersection of the PcG/trxG enriched miRNAs with the miRNAs potentially regulated by PREs revealed the existence of putative miRNA-PcG/trxG regulatory circuits. Our analysis strategy revealed 3 classes of miRNA-PcG/trxG crosstalk interactions in *Drosophila* that define potential 21 regulatory circuits involving 21 miRNAs, 21 *PcG* and 21 *trxG* (Fig. 5). In the first class, for example the regulatory circuit formed by the *dme-miRNA-313-3p* and the *Pc*, *Su(z)2* or *pho* transcripts (Fig. 5B), miRNAs that reciprocally crosstalk with PcG define the *miRNA-PcG circuits*. Into this class, fall 16 of the 21 potential regulatory circuits. These are composed of two reciprocal repressive elements, the miRNA and the PcG, forming *a double negative feedback loop* (Ferro et al., 2019). Modeling studies of this type of loop comprising miRNAs and epigenetic machinery associate it with cell fate decisions (Osella et al., 2014). Examples described in humans include the hsa-miR-214/EZH2 circuit involved in skeletal muscle cell differentiation (Juan et al., 2009) and the hsa-miR-138/SIRT1 circuit involved in axon regeneration (Liu et al., 2013). In the second class, for instance, the circuit formed between *dme-miR-210-5p* and the *trx*, *osa* or *Kis* transcripts (Fig. 5B), miRNAs that reciprocally crosstalk with trxG define the *miRNA-trxG circuits* (3 out of 16). These are composed of a repressive element (the miRNA) and an activating element (the trxG) forming a *positive feedback loop* (Ferro et al., 2019). This second kind of circuits, similar to ones described between miRNAs and activating transcription factors, allow for a tight control of the activating element levels (Ferro et al., 2019). A known human example of this class is the hsa-miR-17–92/E2F1 circuit that is critical for balancing E2F1 protein levels important for the control of cell cycle progression (Li et al., 2011). The third class, consisting of miRNAs that reciprocally crosstalk with both *PcG* and *trxG*, such as the circuit formed between *dme-miR-972-5p* and 15 and 15 *PcG* and *trxG* transcripts respectively (Fig. 5B), define the *miRNAs-PcG/trxG shared circuits* (2 out of 21). In these circuits, a negative regulator (the miRNA) represses both a positive (trxG) and a negative (PcG) regulator. This class of crosstalk interactions is the most complex and so far undescribed functionally.

Of note, our results are not necessarily comprehensive. There might exist more circuits than the ones described here, both due to the intrinsic nature of miRNA transcription (Yin et al., 2015), and/or due to the technical compromise decision established in the different steps of the computational workflow. The case of *dme-mir-8* (hsa-mir-200 ortholog in humans) is illustrative (Cao et al., 2011; Enderle et al., 2011; Polytarchou et al., 2012). The closest predicted PRE is located 24 kb from *dme-miR-8* pre-miRNA, thus it has not been included in our defined miRNA-forming circuits in spite of its significant enrichment in PcG (Fig. 2).

Examination of the 21 miRNAs forming PcG/trxG regulatory circuits revealed conserved miRNA families. This is the case for the TargetScanFlyV7 *AUUGCAC* family that share the same *seed* sequence and include the *dme-miR-310s* cluster, composed of *dme-miR-310-3p*, *dme-miR-311-3p*, *dme-miR-312-3p* and *dme-miR-313-3p* (Ryazansky et al., 2011), as well as *dme-miR-92a-3p* and *dme-miR-92b-3p*. All of the TargetScanFlyV7 *AUUGCAC* family members form circuits of the *miRNA-PcG* class (Fig. 5A). The *dme-miR-310s* cluster is conserved across different *Drosophila* species (Lu et al., 2008) and has a role in the establishment of complex morphological patterns through the regulation of Hox genes (Kaschula et al., 2018). Inactivation of these miRNAs through shRNA produces either pupal lethality or homeotic-like phenotypes (Unpublished results Regojo and Busturia,), thus making the *dme-miR-310s-PcG* circuit identified in this work a promising crosstalk interaction for *in vivo* validation. Moreover, *dme-miR-92a-3p* and *dme-miR-92b-3p* are conserved in all metazoans (Marco et al., 2013). Among other functions, *dme-miR-92a* controls circadian rhythms (Chen and Rosbash, 2017) with a target gene, *l(3)73AH*, that has been predicted to be a member of the PCR1 complex (Gaudet et al., 2011; Irminger-Finger and Nöthiger, 1995). Moreover, human orthologs of *dme-miR-92a* and *dme-miR-92b* have been found to be up-regulated in many cancer types (Chen et al., 2016; Huo et al., 2016; Motoyama et al., 2009; Nass et al., 2009; Poliseno et al., 2010; Ren et al., 2016; Shrestha et al., 2014; Szurián et al., 2017; Xiang et al., 2015), making *dme-miR-92a-3*p and *dme-miR-92b-3p* excellent candidates for translational studies and *in vivo* validation.

The novel integrative computational framework developed in this work has permitted a genome-wide analysis of putative *miRNA-PcG/trxG regulatory circuits*. Notably, this is the first time such an investigation has been performed in *Drosophila* and this framework can serve as a basis for similar analyses in other species. The work identified a group of miRNAs that form regulatory circuits with PcG/trxG potentially regulating each other 's expression. Interestingly, the integration of these circuits in Fig. 5C shows a more complex and higher level of interactions than those defined in this work, to be further explored. The results provided here serve as a useful and concise resource for the *Drosophila* community to consult prior to experimental studies investigating the *Drosophila* regulatory networks of miRNA-PcG/trxG mediated gene expression.

5. Conclusions

- A computational framework has been developed to search for interactions between microRNAs (miRNAs) and Polycomb/trithorax (PcG/trxG) from a genome-wide perspective in *Drosophila melanogaster*. This framework may serve as a basis for similar analysis in another species.
- The search has predicted that 35 and 15 mature microRNAs potentially regulate the expression of PcG and trxG genes respectively. Also, the search predicted that 160 pre-miRNAs are potentially regulated by PcG/trxG proteins.
- This integrative approach resulted in the discovery of crosstalk regulatory interactions that define 16 miRNA-*PcG regulatory circuits*, 3 miRNA-*trxG circuits* and 2 miRNA-*PcG/trxG shared circuits*.
- A new and useful resource is provided to consult prior to experimental studies investigating the *Drosophila* regulatory networks of miRNA-PcG/trxG mediated gene expression.
- These regulatory circuits may uncover a putative novel mechanism in Drosophila to control PcG/trxG and miRNAs levels of expression

Data availability

I have shared the link to my data/code at the Attach File Step

Acknowledgements

We thank Dr. Peter Freddolino (University of Michigan Medical School, USA) for kindly providing us with the Polycomb Response Element genome-wide predictor (Khabiri and Freddolino, 2019) and Keith Harshman for carefully reading the manuscript. This work was supported by PID2020-114533 GB-C21 grant from Spanish Agencia Estatal de Investigación/Ministerio de Ciencia e Innovación and by institutional grants from Fundación Areces and Banco Santander.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ydbio.2022.12.008.

Table S1. PcG/trxG gene sets. Gene sets obtained from the integration of Kassis et al., (2017) (Kassis et al., 2017) review, Ray et al. (2016), Brown and Kassis (2010) and Flybase v2.0 (Thurmond et al., 2019). See Methods.

Table S2. Description of the 460 miRNAs analyzed. "Distance_PRE" is the pre-miRNA 5'end distance to a PRE. "PRE_location" is the genomic location of the pre-miRNA closest PRE. "PcG_OR" and "trxG_OR" represent the PcG and trxG enrichment odds ratio respectively. "PcG_pval" and "trxG_pval" represent PcG and trxG enrichment p-value respectively. In the "PcG_genes" and "trxG_genes" columns, the Cumulative Weighted

Context Score (CWCS) is displayed for the predicted miRNA targets.

References

Agarwal, V., Bell, G.W., Nam, J.-W., Bartel, D.P., 2015. Predicting effective microRNA target sites in mammalian mRNAs. Elife 4, e05005. https://doi.org/10.7554/eLife.05005.

Agarwal, V., Subtelny, A.O., Thiru, P., Ulitsky, I., Bartel, D.P., 2018. Predicting microRNA targeting efficacy in Drosophila. Genome Biol. 19, 152. https://doi.org/10.1186/s13059-018-1504-3.

Artigaud, S., Gauthier, O., Pichereau, V., 2013. Identifying differentially expressed proteins in two-dimensional electrophoresis experiments: inputs from transcriptomics statistical tools. Bioinformatics 29, 2729–2734. https://doi.org/10.1093/bioinformatics/btt464.

Asangani, I.A., Harms, P.W., Dodson, L., Pandhi, M., Kunju, L.P., Maher, C.A., Fullen, D.R., Johnson, T.M., Giordano, T.J., Palanisamy, N., Chinnaiyan, A.M., 2012. Genetic and epigenetic loss of microRNA-31 leads to feed-forward expression of EZH2 in melanoma. Oncotarget 3, 1011–1025. https://doi.org/10.18632/oncotarget.622.

Baek, D., Villén, J., Shin, C., Camargo, F.D., Gygi, S.P., Bartel, D.P., 2008. The impact of microRNAs on protein output. Nature 455, 64–71. https://doi.org/10.1038/nature07242.

Bauer, M., Trupke, J., Ringrose, L., 2016. The quest for mammalian Polycomb response elements: are we there yet? Chromosoma 125, 471–496. https://doi.org/10.1007/s00412-015-0539-4.

Bredesen, B.A., Rehmsmeier, M., 2019. DNA sequence models of genome-wide Drosophila melanogaster Polycomb binding sites improve generalization to independent Polycomb Response Elements. Nucleic Acids Res. 47, 7781–7797. https://doi.org/10.1093/nar/gkz617.

Brown, J.L., Kassis, J.A., 2010. Spps, a Drosophila Sp1/KLF family member, binds to PREs and is required for PRE activity late in development. Development 137, 2597–2602. https://doi.org/10.1242/dev.047761.

Bueno, M.J., Pérez de Castro, I., Gómez de Cedrón, M., Santos, J., Calin, G.A., Cigudosa, J.C., Croce, C.M., Fernández-Piqueras, J., Malumbres, M., 2008. Genetic and epigenetic silencing of MicroRNA-203 enhances ABL1 and BCR-ABL1 oncogene expression. Cancer Cell 13, 496–506. https://doi.org/10.1016/j.ccr.2008.04.018.

Busturia, A., Lloyd, A., Bejarano, F., Zavortink, M., Xin, H., Sakonju, S., 2001. The MCP silencer of the Drosophila Abd-B gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression. Development 128, 2163–2173.

Busturia, A., Wightman, C.D., Sakonju, S., 1997. A silencer is required for maintenance of transcriptional repression throughout Drosophila development. Development 124, 4343–4350.

Cai, S., Weng, Y., Miao, F., 2021. MicroRNA-194 inhibits PRC1 activation of the Wnt/β-catenin signaling pathway to prevent tumorigenesis by elevating self-renewal of non-side population cells and side population cells in esophageal cancer stem cells. Cell Tissue Res. 384, 353–366. https://doi.org/10.1007/s00441-021-03412-z.

Cao, Q., Mani, R.-S., Ateeq, B., Dhanasekaran, S.M., Asangani, I.A., Prensner, J.R., Kim, J.H., Brenner, J.C., Jing, X., Cao, X., Wang, R., Li, Y., Dahiya, A., Wang, L., Pandhi, M., Lonigro, R.J., Wu, Y.-M., Tomlins, S.A., Palanisamy, N., Qin, Z., Yu, J., Maher, C.A., Varambally, S., Chinnaiyan, A.M., 2011. Coordinated regulation of polycomb group complexes through microRNAs in cancer. Cancer Cell 20, 187–199. https://doi.org/10.1016/j.ccr.2011.06.016.

Chan, H.L., Morey, L., 2019. Emerging roles for polycomb-group proteins in stem cells and cancer. Trends Biochem. Sci. 44, 688–700. https://doi.org/10.1016/j.tibs.2019.04.005.

Chen, X., Rosbash, M., 2017. MicroRNA-92a is a circadian modulator of neuronal excitability in Drosophila. Nat. Commun. 8, 14707. https://doi.org/10.1038/ncomms14707.

Chen, Z., Wu, Y., Meng, Q., Xia, Z., 2016. Elevated microRNA-25 inhibits cell apoptosis in lung cancer by targeting RGS3. Vitro Cell. Dev. Biol. - Anim. 52, 62–67. https://doi.org/10.1007/s11626-015-9947-2.

de Nigris, F., 2016. Epigenetic regulators: polycomb-miRNA circuits in cancer. Biochim. Biophys. Acta BBA - Gene Regul. Mech. 1859, 697–704. https://doi.org/10.1016/j.bbagrm.2016.03.005.

dos Santos, G., Schroeder, A.J., Goodman, J.L., Strelets, V.B., Crosby, M.A., Thurmond, J., Emmert, D.B., Gelbart, W.M., 2015. the FlyBase Consortium. Nucleic Acids Res. 43, D690–D697. https://doi.org/10.1093/nar/gku1099. FlyBase: introduction of the Drosophila melanogaster Release 6 reference genome assembly and large-scale migration of genome annotations.

Du, J., Kirk, B., Zeng, J., Ma, J., Wang, Q., 2018. Three classes of response elements for human PRC2 and MLL1/2–Trithorax complexes. Nucleic Acids Res. 46, 8848–8864. https://doi.org/10.1093/nar/gky595.

Enderle, D., Beisel, C., Stadler, M.B., Gerstung, M., Athri, P., Paro, R., 2011. Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. Genome Res. 21, 216–226. https://doi.org/10.1101/gr.114348.110.

Farazi, T.A., Hoell, J.I., Morozov, P., Tuschl, T., 2013. MicroRNAs in human cancer. In: Schmitz, U., Wolkenhauer, O., Vera, J. (Eds.), MicroRNA Cancer Regulation. Springer Netherlands, Dordrecht, pp. 1–20. https://doi.org/10.1007/978-94-007-5590-1_1.

Fathi, M., Ghafouri-Fard, S., Abak, A., Taheri, M., 2021. Emerging roles of miRNAs in the development of pancreatic cancer. Biomed. Pharmacother. 141, 111914. https://doi.org/10.1016/j.biopha.2021.111914.

Ferro, E., Enrico Bena, C., Grigolon, S., Bosia, C., 2019. From endogenous to synthetic microRNA-mediated regulatory circuits: an overview. Cells 8, 1540. https://doi.org/10.3390/cells8121540.

Fiedler, T., Rehmsmeier, M., 2006. jPREdictor: a versatile tool for the prediction of cis-regulatory elements. Nucleic Acids Res. 34, W546–W550. https://doi.org/10.1093/nar/gkl250.

Gaudet, P., Livstone, M.S., Lewis, S.E., Thomas, P.D., 2011. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Briefings Bioinf. 12, 449–462. https://doi.org/10.1093/bib/bbr042.

Huo, J., Zhang, Y., Li, R., Wang, Y., Wu, J., Zhang, D., 2016. Upregulated MicroRNA-25 mediates the migration of melanoma cells by targeting DKK3 through the WNT/β-Catenin pathway. Int. J. Mol. Sci. 17, 1124. https://doi.org/10.3390/ijms17111124.

Irminger-Finger, I., Nöthiger, R., 1995. The Drosophila melanogaster gene lethal(3) 73Ah encodes a ring finger protein homologous to the oncoproteins MEL-18 and BMI-1. Gene 163, 203–208. https://doi.org/10.1016/0378-1119(95)00326-2.

Jia, Y., Ding, X., Zhou, L., Zhang, L., Yang, X., 2021. Mesenchymal stem cells-derived exosomal microRNA-139-5p restrains tumorigenesis in bladder cancer by targeting PRC1. Oncogene 40, 246–261. https://doi.org/10.1038/s41388-020-01486-7.

Jiang, M., Xu, B., Li, X., Shang, Y., Chu, Y., Wang, W., Chen, D., Wu, N., Hu, S., Zhang, S., Li, M., Wu, K., Yang, X., Liang, J., Nie, Y., Fan, D., 2019. Correction: O-GlcNAcylation promotes colorectal cancer metastasis via the miR-101-O-GlcNAc/EZH2 regulatory feedback circuit. Oncogene 38, 5744–5745. https://doi.org/10.1038/s41388-019-0834-2.

Juan, A.H., Kumar, R.M., Marx, J.G., Young, R.A., Sartorelli, V., 2009. Mir-214-Dependent regulation of the polycomb protein Ezh2 in skeletal muscle and embryonic stem cells. Mol. Cell 36, 61–74. https://doi.org/10.1016/j.molcel.2009.08.008.

Kahn, T.G., Stenberg, P., Pirrotta, V., Schwartz, Y.B., 2014. Combinatorial interactions are required for the efficient recruitment of pho repressive complex (PhoRC) to polycomb response elements. PLoS Genet. 10, e1004495. https://doi.org/10.1371/journal.pgen.1004495.

Kaschula, R., Pinho, S., Alonso, C.R., 2018. microRNA-dependent regulation of Hox gene expression sculpts fine-grain morphological patterns in a Drosophila appendage. Development dev 161133. https://doi.org/10.1242/dev.161133.

Kassis, J.A., Brown, J.L., 2013. Polycomb group response elements in Drosophila and vertebrates. In: Advances in Genetics. Elsevier, pp. 83–118. https://doi.org/10.1016/B978-0-12-407677-8.00003-8.

Kassis, J.A., Kennison, J.A., Tamkun, J.W., 2017. Polycomb and trithorax group genes in Drosophila. Genetics 206, 1699–1725. https://doi.org/10.1534/genetics.115.185116.

Kennerdell, J.R., Liu, N., Bonini, N.M., 2018. MiR-34 inhibits polycomb repressive complex 2 to modulate chaperone expression and promote healthy brain aging. Nat. Commun. 9, 4188. https://doi.org/10.1038/s41467-018-06592-5.

Khabiri, M., Freddolino, P.L., 2019. Genome-wide prediction of potential polycomb response elements and their functions (preprint). Syst. Biol. https://doi.org/10.1101/516500.

Kozomara, A., Birgaoanu, M., Griffiths-Jones, S., 2019. miRBase: from microRNA sequences to function. Nucleic Acids Res. 47, D155–D162. https://doi.org/10.1093/nar/gky1141.

Lewis, B.P., Burge, C.B., Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are MicroRNA targets. Cell 120, 15–20. https://doi.org/10.1016/j.cell.2004.12.035.

Li, H., Liefke, R., Jiang, J., Kurland, J.V., Tian, W., Deng, P., Zhang, W., He, Q., Patel, D.J., Bulyk, M.L., Shi, Y., Wang, Z., 2017. Polycomb-like proteins link the PRC2 complex to CpG islands. Nature 549, 287–291. https://doi.org/10.1038/nature23881.

Li, Yichen, Li, Yumin, Zhang, H., Chen, Y., 2011. MicroRNA-mediated positive feedback loop and optimized bistable switch in a cancer network Involving miR-17-92. PLoS One 6, e26302. https://doi.org/10.1371/journal.pone.0026302.

Liu, C.-M., Wang, R.-Y., Saijilafu, Jiao, Z.-X., Zhang, B.-Y., Zhou, F.-Q., 2013. MicroRNA-138 and SIRT1 form a mutual negative feedback loop to regulate mammalian axon regeneration. Genes Dev. 27, 1473–1483. https://doi.org/10.1101/gad.209619.112.

Liu, T., Cai, Jian, Cai, Jing, Wang, Z., Cai, L., 2021. EZH2-miRNA positive feedback promotes tumor growth in ovarian cancer. Front. Oncol. 10, 608393. https://doi.org/10.3389/fonc.2020.608393.

Lu, J., Fu, Y., Kumar, S., Shen, Y., Zeng, K., Xu, A., Carthew, R., Wu, C.-I., 2008. Adaptive evolution of newly emerged micro-RNA genes in Drosophila. Mol. Biol. Evol. 25, 929–938. https://doi.org/10.1093/molbev/msn040.

Marco, A., Ninova, M., Ronshaugen, M., Griffiths-Jones, S., 2013. Clusters of microRNAs emerge by new hairpins in existing transcripts. Nucleic Acids Res. 41, 7745–7752. https://doi.org/10.1093/nar/gkt534.

Motoyama, k, Inoue, H., Takatsuno, Y., Tanaka, F., Mimori, K., Uetake, H., Kenichi Sugihara, K., Mori, M., 2009. Over- and under-expressed microRNAs in human colorectal cancer. Int. J. Oncol. 34. https://doi.org/10.3892/ijo_00000233.

Nass, D., Rosenwald, S., Meiri, E., Gilad, S., Tabibian-Keissar, H., Schlosberg, A., Kuker, H., Sion-Vardy, N., Tobar, A., Kharenko, O., Sitbon, E., Lithwick Yanai, G., Elyakim, E., Cholakh, H., Gibori, H., Spector, Y., Bentwich, Z., Barshack, I., Rosenfeld, N., 2009. MiR-92b and miR-9/9* are specifically expressed in brain primary tumors and can Be used to differentiate primary from metastatic brain tumors. Brain Pathol. 19, 375–383. https://doi.org/10.1111/j.1750-3639.2008.00184.x.

Ning, X., Shi, Z., Liu, X., Zhang, A., Han, L., Jiang, K., Kang, C., Zhang, Q., 2015. DNMT1 and EZH2 mediated methylation silences the microRNA-200b/a/429 gene and promotes tumor progression. Cancer Lett. 359, 198–205. https://doi.org/10.1016/j.canlet.2015.01.005.

Osella, M., Riba, A., Testori, A., CorÃ, D., Caselle, M., 2014. Interplay of microRNA and epigenetic regulation in the human regulatory network. Front. Genet. 5. https://doi.org/10.3389/fgene.2014.00345.

Owen, B.M., Davidovich, C., 2022. DNA binding by polycomb-group proteins: searching for the link to CpG islands. Nucleic Acids Res. 50, 4813–4839. https://doi.org/10.1093/nar/gkac290.

Piunti, A., Shilatifard, A., 2021. The roles of Polycomb repressive complexes in mammalian development and cancer. Nat. Rev. Mol. Cell Biol. 22, 326–345. https://doi.org/10.1038/s41580-021-00341-1.

Poliseno, L., Salmena, L., Riccardi, L., Fornari, A., Song, M.S., Hobbs, R.M., Sportoletti, P., Varmeh, S., Egia, A., Fedele, G., Rameh, L., Loda, M., Pandolfi, P.P., 2010. Identification of the miR-106b ~ 25 MicroRNA cluster as a proto-oncogenic PTEN -targeting intron that cooperates with its host gene MCM7 in transformation. Sci. Signal. 3. https://doi.org/10.1126/scisignal.2000594.

Polytarchou, C., Iliopoulos, D., Struhl, K., 2012. An integrated transcriptional regulatory circuit that reinforces the breast cancer stem cell state. Proc. Natl. Acad. Sci. USA 109, 14470–14475. https://doi.org/10.1073/pnas.1212811109.

Pounds, S., Cheng, C., 2006. Robust estimation of the false discovery rate. Bioinformatics 22, 1979–1987. https://doi.org/10.1093/bioinformatics/btl328.

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

Ray, P., De, S., Mitra, A., Bezstarosti, K., Demmers, J.A.A., Pfeifer, K., Kassis, J.A., 2016. Combgap contributes to recruitment of Polycomb group proteins in Drosophila. Proc. Natl. Acad. Sci. U.S.A. 113, 3826–3831. https://doi.org/10.1073/pnas.1520926113.

Regojo, A., Busturia, A., Unpublished results. Análisis del circuito regulador Polycomb-microRNAs en Drosophila melanogaster.

Ren, C., Wang, W., Han, C., Chen, H., Fu, D., Luo, Y., Yao, H., Wang, D., Ma, L., Zhou, L., Han, D., Shen, M., 2016. Expression and prognostic value of miR-92a in patients with gastric cancer. Tumor Biol. 37, 9483–9491. https://doi.org/10.1007/s13277-016-4865-9.

Ringrose, L., Rehmsmeier, M., Dura, J.-M., Paro, R., 2003. Genome-wide prediction of polycomb/trithorax response elements in Drosophila melanogaster. Dev. Cell 5, 759–771. https://doi.org/10.1016/S1534-5807(03)00337-X.

Ryazansky, S.S., Gvozdev, V.A., Berezikov, E., 2011. Evidence for post-transcriptional regulation of clustered microRNAs in Drosophila. BMC Genom. 12, 371. https://doi.org/10.1186/1471-2164-12-371.

Sato, F., Tsuchiya, S., Meltzer, S.J., Shimizu, K., 2011. MicroRNAs and epigenetics: MicroRNAs and epigenetics. FEBS J. 278, 1598–1609. https://doi.org/10.1111/j.1742-4658.2011.08089.x.

Schubert, M., Spahn, M., Kneitz, S., Scholz, C.J., Joniau, S., Stroebel, P., Riedmiller, H., Kneitz, B., 2013. Distinct microRNA expression profile in prostate cancer patients with early clinical failure and the impact of let-7 as prognostic marker in high-risk prostate cancer. PLoS One 8, e65064. https://doi.org/10.1371/journal.pone.0065064.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., Cavalli, G., 2007. Genome regulation by polycomb and trithorax proteins. Cell 128, 735–745. https://doi.org/10.1016/j.cell.2007.02.009.

Schuettengruber, B., Ganapathi, M., Leblanc, B., Portoso, M., Jaschek, R., Tolhuis, B., van Lohuizen, M., Tanay, A., Cavalli, G., 2009. Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. PLoS Biol. 7, e1000013. https://doi.org/10.1371/journal.pbio.1000013.

Schwartz, Y.B., Kahn, T.G., Stenberg, P., Ohno, K., Bourgon, R., Pirrotta, V., 2010. Alternative epigenetic chromatin states of polycomb target genes. PLoS Genet. 6, e1000805. https://doi.org/10.1371/journal.pgen.1000805.

Shabalina, S., Koonin, E., 2008. Origins and evolution of eukaryotic RNA interference. Trends Ecol. Evol. 23, 578–587. https://doi.org/10.1016/j.tree.2008.06.005.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504. https://doi.org/10.1101/gr.1239303.

Shrestha, S., Hsu, S., Huang, W., Huang, Hsi-Yuan, Chen, W., Weng, S., Huang, Hsien-Da, 2014. A systematic review of microRNA expression profiling studies in human gastric cancer. Cancer Med. 3, 878–888. https://doi.org/10.1002/cam4.246.

Szurián, K., Csala, I., Piurkó, V., Deák, L., Matolcsy, A., Reiniger, L., 2017. Quantitative miR analysis in chronic lymphocytic leukaemia/small lymphocytic lymphoma – proliferation centres are characterized by high miR-92a and miR-155 and low miR-150 expression. Leuk. Res. 58, 39–42. https://doi.org/10.1016/j.leukres.2017.04.002.

Thurmond, J., Goodman, J.L., Strelets, V.B., Attrill, H., Gramates, L.S., Marygold, S.J., Matthews, B.B., Millburn, G., Antonazzo, G., Trovisco, V., Kaufman, T.C., Calvi, B.R., , the FlyBase Consortium, Perrimon, N., Gelbart, S.R., Agapite, J., Broll, K., Crosby, L., Santos, G., dos Emmert, D., Gramates, L.S., Falls, K., Jenkins, V., Matthews, B., Sutherland, C., Tabone, C., Zhou, P., Zytkovicz, M., Brown, N., Antonazzo, G., Attrill, H., Garapati, P., Holmes, A., Larkin, A., Marygold, S., Millburn, G., Pilgrim, C., Trovisco, V., Urbano, P., Kaufman, T., Calvi, B., Czoch, B., Goodman, J., Strelets, V., Thurmond, J., Cripps, R., Baker, P., 2019. FlyBase 2.0: the next generation. Nucleic Acids Res. 47, D759–D765. https://doi.org/10.1093/nar/gky1003.

Varambally, S., Cao, Q., Mani, R.-S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K., Brenner, J.C., Yu, J., Kim, J.H., Han, B., Tan, P., Kumar-Sinha, C., Lonigro, R.J., Palanisamy, N., Maher, C.A., Chinnaiyan, A.M., 2008. Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science 322, 1695–1699. https://doi.org/10.1126/science.1165395.

Wang, G.G., Konze, K.D., Tao, J., 2015. Polycomb genes, miRNA, and their deregulation in B-cell malignancies. Blood 125, 1217–1225. https://doi.org/10.1182/blood-2014-10-606822.

Wessels, H.-H., Lebedeva, S., Hirsekorn, A., Wurmus, R., Akalin, A., Mukherjee, N., Ohler, U., 2019. Global identification of functional microRNA-mRNA interactions in Drosophila. Nat. Commun. 10, 1626. https://doi.org/10.1038/s41467-019-09586-z.

Xiang, J., Hang, J.-B., Che, J.-M., Li, H.-C., 2015. miR-25 is up-regulated in non-small cell lung cancer and promotes cell proliferation and motility by targeting FBXW7. Int. J. Clin. Exp. Pathol. 8, 9147–9153.

Xu, Y.-J., Liu, P.-P., Ng, S.-C., Teng, Z.-Q., Liu, C.-M., 2020. Regulatory networks between Polycomb complexes and non-coding RNAs in the central nervous system. J. Mol. Cell Biol. 12, 327–336. https://doi.org/10.1093/jmcb/mjz058.

Yin, S., Yu, Y., Reed, R., 2015. Primary microRNA processing is functionally coupled to RNAP II transcription in vitro. Sci. Rep. 5, 11992. https://doi.org/10.1038/srep11992.

You, Y., Que, K., Zhou, Y., Zhang, Z., Zhao, X., Gong, J., Liu, Z., 2018. MicroRNA-766-3p inhibits tumour progression by targeting Wnt3a in hepatocellular carcinoma. Mol. Cell 41, 830–841. https://doi.org/10.14348/molcells.2018.0181.

Zeng, J., Kirk, B.D., Gou, Y., Wang, Q., Ma, J., 2012. Genome-wide polycomb target gene prediction in Drosophila melanogaster. Nucleic Acids Res. 40, 5848–5863. https://doi.org/10.1093/nar/gks209.