DEVELOPMENTAL BIOLOGY

Variant Polycomb complexes in *Drosophila* consistent with ancient functional diversity

Hyuckjoon Kang[1,2]†, Janel R. Cabrera[1,2,3]†, Barry M. Zee[1,2], Heather A. Kang[1,2], Jenny Marie Jobe[3], Maeve B. Hegarty[3], Aurelie E. Barry[3], Alexander Glotov[4], Yuri B. Schwartz[4], Mitzi I. Kuroda[1,2]*

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Polycomb group (PcG) mutants were first identified in *Drosophila* on the basis of their failure to maintain proper *Hox* gene repression during development. The proteins encoded by the corresponding fly genes mainly assemble into one of two discrete Polycomb repressive complexes: PRC1 or PRC2. However, biochemical analyses in mammals have revealed alternative forms of PRC2 and multiple distinct types of noncanonical or variant PRC1. Through a series of proteomic analyses, we identify analogous PRC2 and variant PRC1 complexes in *Drosophila*, as well as a broader repertoire of interactions implicated in early development. Our data provide strong support for the ancient diversity of PcG complexes and a framework for future analysis in a longstanding and versatile genetic system.

INTRODUCTION

The Polycomb group (PcG) complexes, PRC1 and PRC2, each encompass numerous alternative subunits and configurations in mammalian cells (*1*–*6*). This has not been explored to the same extent in *Drosophila*, where Polycomb complexes comprise a reduced number of paralogous and accessory subunits (*7*). In the absence of extensive analyses, it has been assumed that PcG complexes have greatly diversified in mammals. However, numerous subunits of alternative PcG complexes are highly conserved, including the RING1 and YY1 binding protein (RYBP) and PcG RING finger (PCGF) proteins, which have ancient origins (*8*).

The previously described *Drosophila* dRAF and PhoRC complexes have multiple subunits in common with mammalian variant PRC1 complexes vPRC1.1 and vPRC1.6, respectively (*9*, *10*). However, several subunits thought to play defining roles in mammals were not detected in the fly complexes, including RYBP. Furthermore, orthologous PRC1.3/5 complexes have not been reported. These observations support the need for additional analyses in *Drosophila*.

Here, we use cross-linking, tandem affinity purification, and mass spectrometry (BioTAP-XL) to find that fly embryos use RYBP (CG12190) and three PCGF subunits—Psc (CG3886), Su(z)2 (CG3905), and L(3)73Ah (CG4195)—to assemble complexes related to all previously described vPRC1 subtypes. CG14073 (BCOR) is a signature subunit of PRC1.1, and CG8677 (RSF1) is a newly identified interactor. Fly PRC1.3/5 may have a conserved role in the nervous system based on Tay (CG9056), its defining subunit. Sfmbt (CG16975) interactions, which encompass the previously described PhoRC-L, suggest unexpectedly broad modularity of a potential fly vPRC1.6.

We also confirm the modularity of PRC2 in *Drosophila*, with Pcl (CG5109) and Scm (CG9495) restricted to PRC2.1 and Jarid2 (CG3654) and Jing (CG9397) restricted to PRC2.2. The conservation appears to extend to the association of PRC2.1 with stable repression and PRC2.2 with a heterogenous or transitional role. Phenotypes from overexpression of Jing or compensatory knockdown of Jarid2 provide further evidence for the importance of a proper balance between PRC2.1 and PRC2.2 during development.

RESULTS AND DISCUSSION

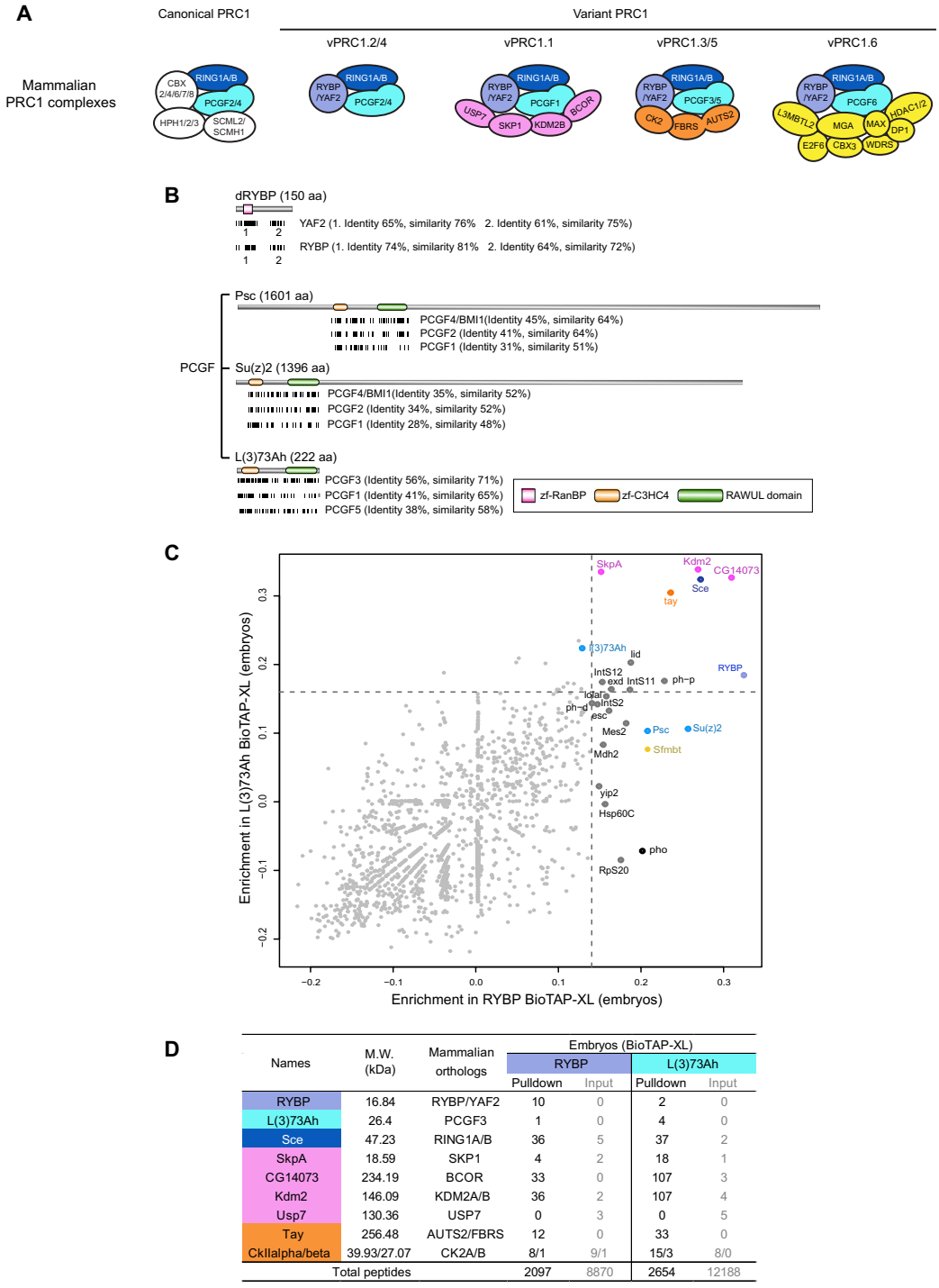
RYBP and L(3)73Ah are core subunits of *Drosophila* vPRC1 complexes

Biochemical analyses have defined two major classes of PRC1 complexes in mammals. Canonical PRC1 complexes (cPRC1.2 and cPRC1.4) function in the maintenance of stable repression that is critical for the classical role of *Hox* gene repression. Noncanonical PRC1 complexes or vPRC1 are responsible for most ubiquitination of H2A on lysine 119 (H2AK119ub) and appear to have diverse functions, including both activation and repression in the case of PRC1.3/5 (*11*–*16*). In mammals, either RING1A or RING1B is present in all cPRC1 and vPRC1 complexes. In contrast, RYBP or YY1 associated factor 2 (YAF2) is limited to vPRC1 complexes (*17*), and six PCGF proteins define the two cPRC1 and six vPRC1 complexes (Fig. 1A) (*14*, *18*).

As in mammals, Sce (CG5595, also known as dRING) is a conserved core subunit of *Drosophila* PRC1. In addition, *Drosophila* RYBP (CG12190) shows strong similarity to the mammalian paralogs, RYBP and YAF2 (Fig. 1B) and is known to interact with Sce (*19*). The orthologous relationships of PCGF proteins in flies appear somewhat less straightforward. The two well-studied orthologs, Psc and Su(z)2, play central but partially redundant roles in cPRC1 (*20*–*23*) and are most similar to mammalian PCGF4, PCGF2, and PCGF1 (Fig. 1B and fig. S1). The most evolutionarily conserved fly PCGF protein, L(3)73Ah (*8*, *24*) has not been studied with regard to its ability to participate in PcG complexes but is most related to PCGF3, the most ancient of the PCGF family (*8*). All fly PCGF orthologs are relatively distant from mammalian PCGF6 (fig. S1).

Ideally, we would identify all candidate PRC1 subunits in flies by affinity-purifying Sce-associated complexes. However, we were unable to recover functional Sce protein after tagging either the N or C terminus with the BioTAP epitope (fig. S2A). Therefore, we proceeded to tag RYBP and L(3)73Ah, established that they were functional as fusion proteins by viability rescue tests (fig. S2A), and performed BioTAP-XL affinity purification and mass spectrometry in the respective transgenic 12- to 24-hour embryos (Fig. 1C). Perhaps because of the small sizes of the two bait proteins, only 10 peptides

[1]Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA. [2]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. [3]Biology Department, Emmanuel College, Boston, MA 02115, USA. [4]Department of Molecular Biology, Umeå University, SE-90187 Umeå, Sweden.
*Corresponding author. Email: mkuroda@genetics.med.harvard.edu
†These authors contributed equally to this work.



Names	M.W. (kDa)	Mammalian orthologs	Embryos (BioTAP-XL)			
			RYBP		L(3)73Ah	
			Pulldown	Input	Pulldown	Input
RYBP	16.84	RYBP/YAF2	10	0	2	0
L(3)73Ah	26.4	PCGF3	1	0	4	0
Sce	47.23	RING1A/B	36	5	37	2
SkpA	18.59	SKP1	4	2	18	1
CG14073	234.19	BCOR	33	0	107	3
Kdm2	146.09	KDM2A/B	36	2	107	4
Usp7	130.36	USP7	0	3	0	5
Tay	256.48	AUTS2/FBRS	12	0	33	0
CkIIalpha/beta	39.93/27.07	CK2A/B	8/1	9/1	15/3	8/0
Total peptides			2097	8870	2654	12188

Fig. 1. RYBP and L(3)73Ah copurify orthologs of mammalian vPRC1 subunits. (**A**) Mammalian PRC1 complex components. PRC1 complexes contain core subunits RING1A/B and PCGF, and variant PRC1 (vPRC1) complexes commonly contain RYBP or YAF2. Each PRC1 subtype is defined by distinct PCGF proteins and additional accessory proteins, as indicated. (**B**) Top: Sequence alignment of *Drosophila* RYBP (NP_001286742.1) with mammalian YAF2 (XP_011536030.1) and RYBP (NP_036366.3). Bottom: Alignment of Psc (NP_001286368.1), Su(z)2 (NP_001260933.1), and L(3)73Ah (NP_001246797.1) with mammalian PCGF orthologs PCGF1 (NP_116062.2), PCGF2 (NP_001356543.1), PCGF3 (NP_001304765.1), PCGF4/BMI1 (NP_005171.4), and PCGF5 (NP_001243478.1). Identical sequences of conserved regions are depicted as black lines, and percent sequence identity and percent similarity (identity + conservative substitutions) between two protein sequences are described in parentheses. Conserved domains are also shown. (**C**) Enrichment plots ($\log_{10}$ fold enrichment of normalized spectral abundance factors (NSAFs) in pulldown with respect to input) for proteins (individual dots) identified in RYBP and L(3)73Ah BioTAP-XL experiments from *Drosophila* embryos. Dashed lines denote the 99th percentile threshold of enriched proteins in RYBP pulldown (*x* axis) and in L(3)73Ah pulldown (*y* axis). (**D**) Peptide counts of *Drosophila* PRC1.1 and PRC1.3 subunit orthologs copurified by RYBP and L(3)73Ah embryonic BioTAP-XL compared to input. Proteins are color-coded according to their mammalian vPRC1 subunit orthologs in (A). Counts of all peptides detected in BioTAP-XL pulldowns and inputs are indicated as total peptides at the bottom of the table. See data file S1 for the full set of results.

of RYBP and 4 peptides of L(3)73Ah were recovered in their corresponding affinity purifications. Nevertheless, the pulldowns enriched for many candidate vPRC1 subunits at substantive levels (Fig. 1D). For example, several orthologs of mammalian PRC1.1 and PRC1.3/5 were copurified in both embryonic RYBP and L(3)73Ah BioTAP-XL pulldowns, strongly suggesting that these variant complexes are conserved in *Drosophila*. The RYBP pulldown additionally copurified Sfmbt, Pho (Pleiohomeotic, CG17743), and Psc/Su(z)2, as well as PRC1.1. and PRC1.3/5 subunit orthologs, suggesting that RYBP could be a core subunit of all vPRC1 complexes, as it is in mammals.

Kdm2 and CG14073 (BCOR) are unique to vPRC1.1, while Tay defines vPRC1.3/5

From our previous results, we could not exclude the possibility that orthologs of mammalian PRC1.1 and PRC1.3/5 subunits enriched in both RYBP and L(3)73Ah pulldowns might be subunits of one composite complex in *Drosophila*. Therefore, we performed Kdm2 (CG11033) BioTAP-XL in embryos to determine whether its presence could discriminate between fly vPRC1.1 and PRC1.3/5. We were unable to test BioTAP-tagged Kdm2 for functionality, as Kdm2 mutants are viable (*25*); however, we successfully copurified Kdm2-interacting orthologs of mammalian PRC1.1. Whether interactions are expressed as enrichment plots (Fig. 2, A and B) or in a heatmap (Fig. 2C), it is apparent that Tay, the fly ortholog of mammalian signature PRC1.3/5 subunits AUTS2/FBRS/FBRSL1, was not copurified in BioTAP-XL pulldown of Kdm2. From this, we conclude that PRC1.1 and PRC1.3/5 exist as separate vPRC1 complexes in *Drosophila*, as in mammals (Fig. 2D). Fly *tay* was found in a behavioral screen of locomotor mutants (*26*), while AUTS2 is named for its association with autism in humans (*13*, *27*), suggesting a conserved link to the nervous system in the two evolutionarily distant species.

Although the previously described dRAF complex was identified after Kdm2 copurification with Psc (*10*), our Kdm2 and L(3)73Ah BioTAP-XL pulldowns did not enrich for Psc or Su(z)2, while orthologs of PRC1.1 subunits, e.g., CG14073 (BCOR) and SkpA (CG16983), strongly interacted with Kdm2. We cannot exclude the existence of the dRAF complex in flies, but Kdm2 seems to mainly form a complex with CG14073, SkpA, Sce, and L(3)73Ah in *Drosophila* embryos.

Unexpectedly, the Kdm2 BioTAP-XL pulldown also strongly copurified CG8677, the *Drosophila* ortholog of RSF1 (remodeling and spacing factor 1). In mammals, RSF1 specifically recognizes H2AK119ub, the chromatin mark catalyzed principally by vPRC1 (*28*). As L(3)73Ah also copurified CG8677, this result supports the discovery of RSF1 as a PcG-related H2Aub reader (*28*). Usp7 (CG1490), a ubiquitin-specific protease, was absent from our purifications, raising the possibility that the recovery of CG8677 (RSF1) could be related to stabilization or accessibility of H2Aub.

Potential relationship of PhoRC to PRC1.6, with an expanded repertoire of interactions

On the basis of the prior discovery of PhoRC and PhoRC-L as stable, soluble Sfmbt complexes (Fig. 2D) (*9*, *29*) and the recovery of Pho, Sfmbt, Psc, and Su(z)2 in our RYBP pulldown (Fig. 2, B and C), we created BioTAP-Sfmbt as a candidate bait to further explore the possible relationship between PhoRC-L and vPRC1.6 in flies. We found that BioTAP-Sfmbt associated with larval polytene chromosomes as expected for a PcG protein (fig. S2B) but failed to rescue mutants to viability (fig. S2A). Previously, transgenic Sfmbt, tandem-tagged with protein A and calmodulin-binding peptide, also failed to rescue mutant animals. However, it restored *Hox* gene repression in mutant clones and was used successfully to affinity-purify PhoRC-L (*9*). Therefore, we proceeded with proteomic analysis of BioTAP-tagged Sfmbt in both embryos and a stable S2 cell line.

Sfmbt is the *Drosophila* ortholog of L3MBTL2, a subunit of mammalian PRC1.6, but it is also equidistant in identity and similarity to mammalian MBTD1 (Fig. 3A). In addition, fly Sfmbt has a sterile alpha motif (SAM) domain, while its mammalian counterparts do not. Perhaps for these reasons, the Sfmbt pulldowns generated a much longer list of interactors than typical for the relatively discrete PcG complexes that we have analyzed in the past. Subunits of fly PhoRC-L and orthologs of mammalian PRC1.6 subunits were enriched but broadly distributed in the enrichment rankings (Fig. 3, B and C). To confirm this, we compared the enrichments between the embryo and S2 cell experiments (Fig. 3D). Dashed lines denote the top 1% enrichment in S2 cells (*x* axis) and in 12- to 24-hours embryos (*y* axis) showing general agreement between the two distinct biological samples. The total peptide counts of selected proteins enriched in Sfmbt BioTAP-XL pulldowns are shown in Fig. 3E compared to their cognate inputs. Two notable orthologs of mammalian vPRC1.6 (Fig. 1A) were missing from our Sfmbt pulldown: RYBP and MAX. We previously found that when RYBP was used as the BioTAP-XL bait, it was a strong interactor of Pho and Sfmbt (Fig. 2C), even when RYBP itself was represented by only 10 peptides (Fig. 1D). Therefore, RYBP might not be identified by the reciprocal Sfmbt pulldown because of its small size and, thus, fewer peptides for detection after cross-linking. *Drosophila* Max (CG9648) might also be undetected because of its small size (18.5 kDa). However, it seems likely that Max is not a subunit of *Drosophila* Sfmbt complexes because the bHLHZip domain for heterodimerization between MGA and MAX is absent in *Drosophila* MGA ortholog Ocm (CG3363) (*30*).

In addition to the orthologs previously implicated in mammalian vPRC1.6, we found *Drosophila*-specific interactors [Debra (CG11371), CG16711, and CG34179] and proteins with SAM domains [L(3)mbt (CG5954), CG2662, and Ph-p (CG18412)]. Fly Sfmbt itself has a SAM domain as noted above, which is absent from its mammalian counterparts (Fig. 3A). Thus, it may be expected that Sfmbt can interact with other SAM-containing proteins through homo- or heteromeric interaction, a known property of SAM domains (*31*). Sfmbt was one of the most enriched proteins in our previous analysis of BioTAP-XL Scm (*32*), and Scm mediates a trimeric complex with Sfmbt and Ph-p through its SAM domain (*33*). Consistent with these key interactions, our Sfmbt pulldown also enriched for subunits of cPRC1 (Ph-p) and PRC2 [Pcl (CG5109) and Su(z)12 (CG8013)], and previous pulldowns of cPRC1 and PRC2 identified Sfmbt and Pho (*32*, *34*). Together, our results are consistent with the ability of PhoRC to contribute to tethering PRC2, cPRC1, and potentially vPRC1.2/4 to some Polycomb response elements (PREs) (*33*).

Sfmbt complexes are linked to coactivators in embryos

We found interactions with coactivators and additional chromatin regulators in embryos that were diminished or absent in S2 cells, suggesting that Sfmbt complexes may have functions beyond PcG repression during development (Figs. 3E and 4A and fig. S3). When we examined the top 35 proteins enriched by Sfmbt pulldown in embryos, we noted that those written in red (Fig. 4A) map to a common STRING (search tool for retrieval of interacting genes/proteins) protein-protein interaction network in mammals (Fig. 4, B and C).

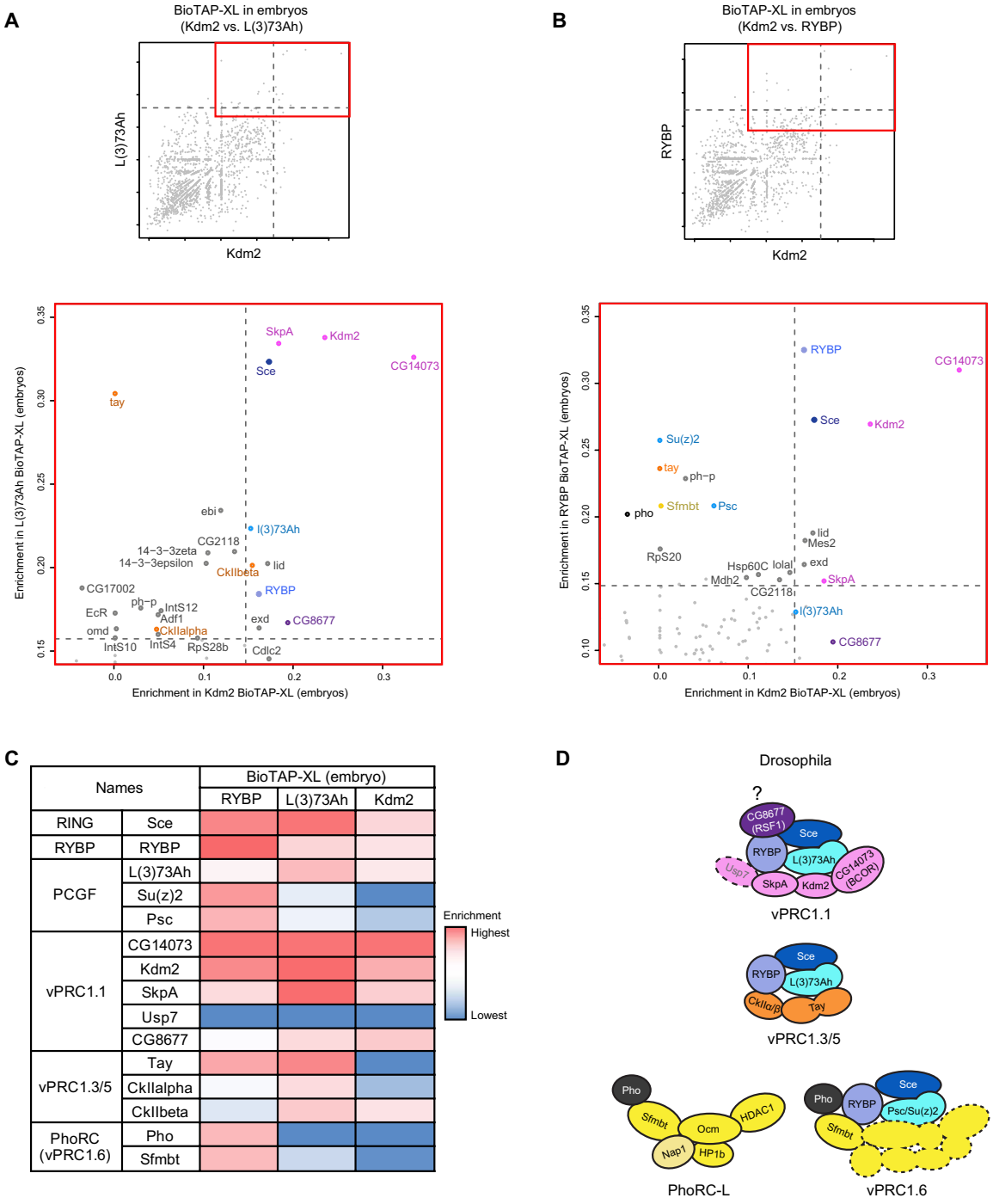
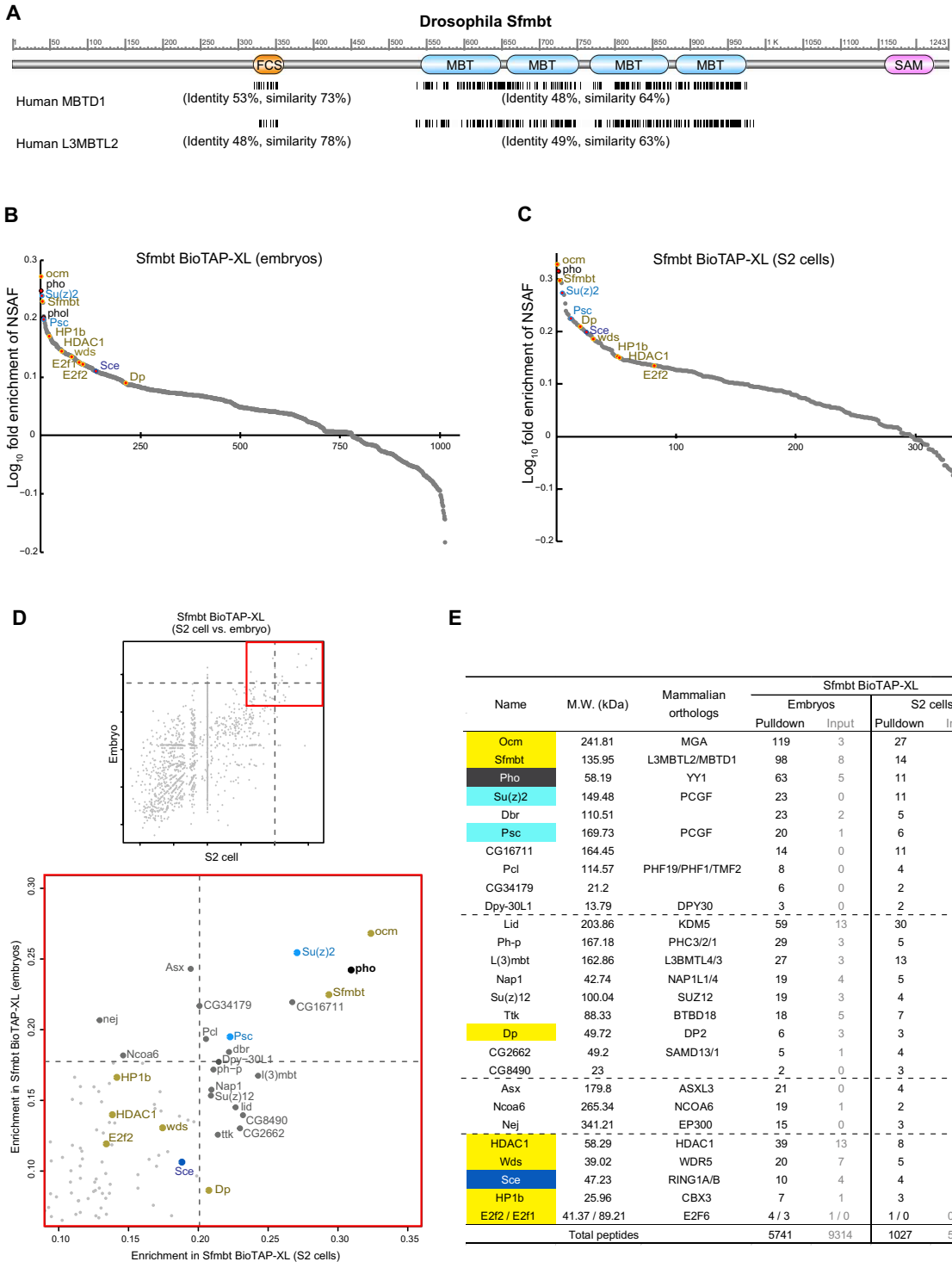


Fig. 2. *Drosophila* complexes corresponding to mammalian vPRC1.1 and vPRC1.3/5. (**A** and **B**) Scatterplot comparing enriched proteins from Kdm2 and L(3)73Ah BioTAP-XL from 12- to 24-hour embryos (A). Scatterplot comparing enriched proteins from Kdm2 and RYBP pulldowns from 12- to 24-hour embryos (B). Dashed lines on the *x* axis denote 99th percentile threshold of enriched proteins in Kdm2 pulldown (*x* axis) and dashed lines on the *y* axis denote 99th percentile threshold of enriched proteins by L(3)73Ah and RYBP in (A) and (B), respectively. Coordinates represent $\log_{10}$ fold enrichment of NSAF of proteins in BioTAP-XL affinity purification compared to input. Red boxes indicate areas of scatterplots that are enlarged in bottom panels, and protein names are color-coded according to the vPRC1 subunit color scheme in Fig. 1A and (D). (**C**) Heatmap of enrichment of selected vPRC1 subunits copurified from RYBP, L(3)73Ah, and Kdm2 BioTAP-XL. For relative abundance comparison, the NSAF enrichment value of each protein is divided by the NSAF enrichment value of CG14073 (BCOR), which is the top hit common to all three pulldowns. (**D**) Illustration of *Drosophila* vPRC1 complexes from BioTAP-XL mass spectrometry analysis of the three bait proteins RYBP, L(3)73Ah, and Kdm2 organized similarly to their mammalian complexes in Fig. 1A. Usp7 (not copurified by three bait proteins) is depicted by a dashed line. CG8677, indicated by a question mark, is a potential *Drosophila*-specific subunit of PRC1.1.



E

Name	M.W. (kDa)	Mammalian orthologs	Sfmbt BioTAP-XL			
			Embryos		S2 cells	
			Pulldown	Input	Pulldown	Input
Ocm	241.81	MGA	119	3	27	0
Sfmbt	135.95	L3MBTL2/MBTD1	98	8	14	0
Pho	58.19	YY1	63	5	11	0
Su(z)2	149.48	PCGF	23	0	11	0
Dbr	110.51		23	2	5	0
Psc	169.73	PCGF	20	1	6	0
CG16711	164.45		14	0	11	0
Pcl	114.57	PHF19/PHF1/TMF2	8	0	4	0
CG34179	21.2		6	0	2	0
Dpy-30L1	13.79	DPY30	3	0	2	0
Lid	203.86	KDM5	59	13	30	5
Ph-p	167.18	PHC3/2/1	29	3	5	0
L(3)mbt	162.86	L3BMTL4/3	27	3	13	1
Nap1	42.74	NAP1L1/4	19	4	5	1
Su(z)12	100.04	SUZ12	19	3	4	0
Ttk	88.33	BTBD18	18	5	7	1
Dp	49.72	DP2	6	3	3	0
CG2662	49.2	SAMD13/1	5	1	4	0
CG8490	23		2	0	3	0
Asx	179.8	ASXL3	21	0	4	0
Ncoa6	265.34	NCOA6	19	1	2	0
Nej	341.21	EP300	15	0	3	1
HDAC1	58.29	HDAC1	39	13	8	6
Wds	39.02	WDR5	20	7	5	2
Sce	47.23	RING1A/B	10	4	4	1
HP1b	25.96	CBX3	7	1	3	2
E2f2 / E2f1	41.37 / 89.21	E2F6	4 / 3	1 / 0	1 / 0	0 / 0
	Total peptides		5741	9314	1027	5001

Fig. 3. Sfmbt interacts with orthologs of mammalian PRC1.6 and many additional proteins. (**A**) Conserved domains in *Drosophila* Sfmbt and sequence alignments with human MBT domain orthologs. Sequence identities between *Drosophila* Sfmbt protein sequence (NP_001137821.1) and two human orthologs, MBTD1 (XP_011523224.1) and L3MBTL2 (NP_001137821.1) are shown as black lines. Percent sequence identity and percent similarity in domains are also described in parentheses. (**B** and **C**) Log_{10} fold enrichment of relative abundance for proteins identified in Sfmbt BioTAP-XL experiments from 12- to 24-hour embryos (B) and from S2 cells (C). (**D**) Scatterplot of Sfmbt pulldown enrichments from embryos and S2 cells normalized to their inputs. Horizontal and vertical dashed lines represent the 99th percentile of proteins enriched by Sfmbt BioTAP-XL in embryos and in S2 cells, respectively. Red box indicates the area of the scatterplot enlarged in the bottom panel. (**E**) Total peptide counts of select proteins enriched from Sfmbt BioTAP-XL experiment in embryos and in S2 cells compared to peptide counts from respective inputs. Total peptide counts from pulldowns are indicated at the bottom of table. See data file S1 for the full results. (B to E) Orthologs of mammalian PRC1.6 subunits and PhoRC subunits are color-coded according to color scheme in Figs. 1A and 2D.

A

	Names	Mammalian orthologs	M.W. (kDa)	Sfmbt BioTAP-XL (embryo) Pulldown	Input
1	Ocm	MGA	241.81	119	3
2	Pho	YY1	58.19	63	5
3	Su(z)2	PCGF	149.48	23	0
4	Sfmbt	MBTD1/L3MBTL2	135.95	98	8
5	Asx	ASXL3	179.8	21	0
6	CG16711		164.45	14	0
7	Phol	YY1	73.16	10	0
8	Psc	PCGF	169.73	20	1
9	CG34179		21.2	6	0
10	Nej	EP300	341.21	15	0
11	Dbr		110.51	23	2
12	Ncoa6	NCOA6	265.34	19	1
13	Tip60	KAT5	61.2	7	0
14	Lint-1		67.86	11	1
15	Pcl	PHF19/MTF2/PHF1	114.57	8	0
16	Hth	MEIS2/MEIS1	52.75	6	0
17	Ph-p	PHC3	167.18	29	3
18	Zfh2	ZFHX4/ZFHX3	331.81	11	0
19	Usp	RXRA	55.21	6	0
20	GATAd	GATA4	93.27	7	0
21	L(3)mbt	L3MBTL4	162.86	27	3
22	HP1b	CBX3	25.96	7	1
23	E(Pc)	EPC1	228.4	14	1
24	Smt3	SUMO3	10.12	27	12
25	Esc	EED	47.96	8	1
26	Hr46	RORB	94.92	6	0
27	Su(var)3-3	KDM1A	98.33	6	0
28	Nap1	NAP1L1	42.74	19	4
29	Atac3	GABPB2	63.98	13	2
30	Ash2	ASH2L	63.19	8	1
31	MED27	MED27	33.85	4	0
32	Pa1	PAGR1	13.24	3	0
33	Egg	SETDB1	141.86	10	1
34	Pan	TCF7L2	132	6	0
35	Dpy-30L1	DPY30	13.79	3	0

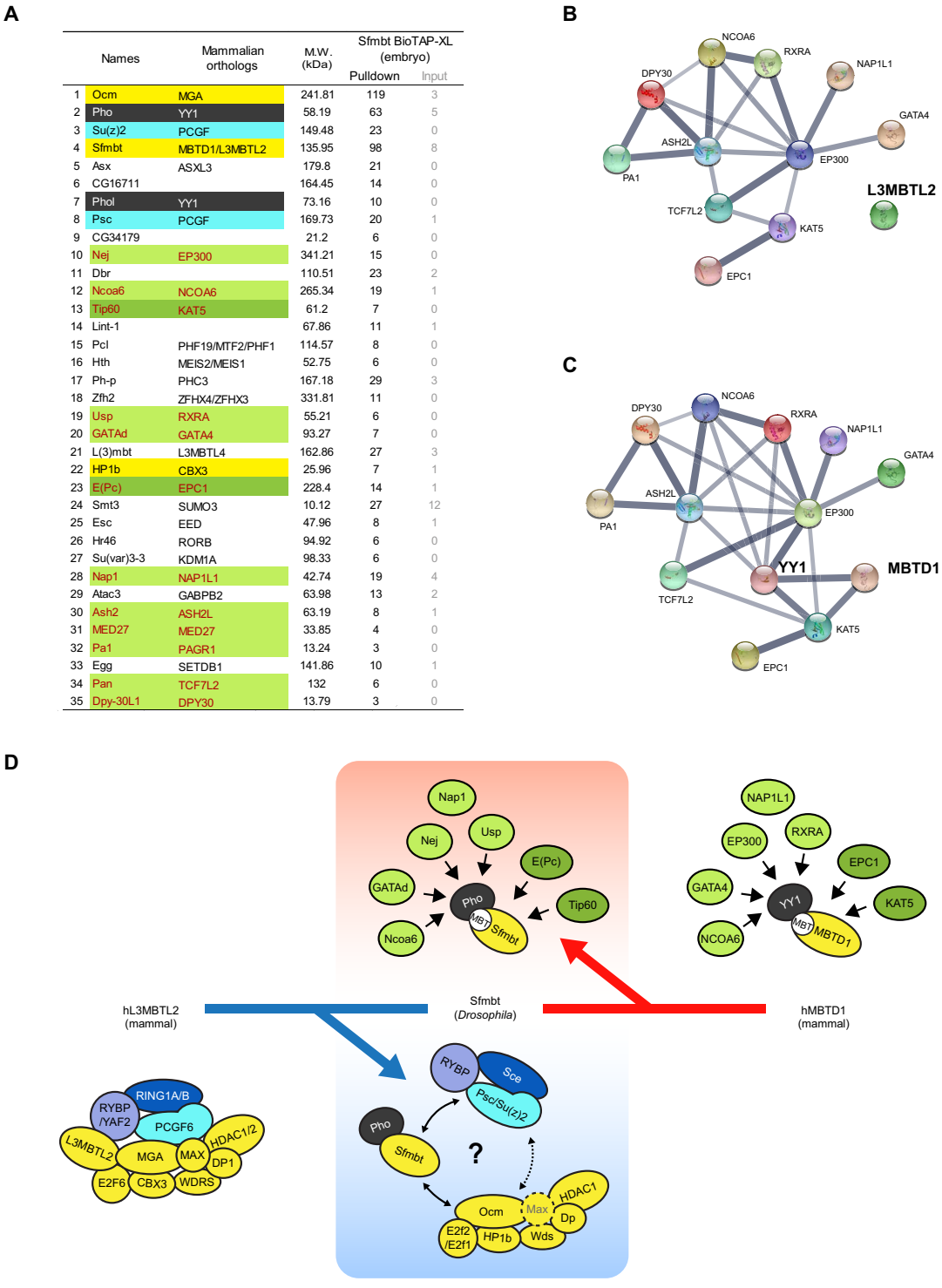


Fig. 4. Embryonic Sfmbt copurifies orthologs of coactivators linked to mammalian YY1 and MBTD1. (**A**) Total peptide counts of the top 35 proteins enriched from Sfmbt BioTAP-XL in embryos. See data file S1 for full results. Among non-PcG proteins, those with mammalian orthologs that have known interactions are written in red font and highlighted in color-coded boxes according to the color scheme in (D). (**B**) STRING protein-protein interaction network for human orthologs of proteins highlighted in (A). Only known interactions (from curated databases and experimentally determined) are used as interaction sources, and the thickness of the gray lines indicates strength of supporting data. L3MBTL2 (human ortholog of Sfmbt) is not connected with the interaction network. (**C**) Combined interaction network of MBTD1 (human ortholog of Sfmbt) and YY1 (human ortholog of Pho/Phol). (**D**) *Drosophila* Sfmbt may have properties of both L3MBTL2 and MBTD1 and may interact with binding partners of each ortholog. Furthermore, fly vPRC1.6 may not be a single entity but instead encompass multiple modules with additional combinatorial capabilities not included here.

The interaction network is not connected to L3MBTL2 (Fig. 4B). However, YY1 provides a potential bridge between the network and MBTD1 (Fig. 4C), as it can bind to a helical groove of MBTD1 in vitro, although with a relatively weak binding affinity (*9*). Together, we can speculate that these additional Sfmbt interactions may be due to an orthologous relationship to MBTD1. The MBTD1 network and our pulldown results include several prominent interactors implicated in activation and specifically enriched in embryos. These include Nej (CG15319), the ortholog of the EP300/CBP acetyltransferases, and Tip60 (CG6121) and E(Pc) (CG7776), orthologs of members of the NuA4/TIP60 acetyltransferase complex. Pho-like (Phol, CG3445) is also enriched in the embryonic pulldown and not in S2 cells (Fig. 3, A and B, and fig. S3, A and B). Phol is closely related to Pho, and *pho phol* double mutants show more severe misexpression of homeotic genes than do single *pho* mutants (*35*). However, maternal Pho and Phol are each required for fertilized egg development, strongly suggesting that they harbor some nonoverlapping functions early in development (*35*, *36*). That Pho and/or Phol could be linked to transcriptional activation would be consistent with their presence on PREs and other regulatory sequences in both repressive and activating contexts (*36*–*42*). Furthermore, the Pho/Phol mammalian ortholog, YY1, is named for its dual function in the two opposing gene regulatory states (*43*).

In summary, we provide strong evidence that *Drosophila* has counterparts for vPRC1 complexes found in mammals, with most subunits conserved. Notably, PCGF proteins do not have a strict one-to-one correspondence between flies and mammals, leaving some ambiguity regarding the configurations of vPRC1.2/4 and PRC1.6 in *Drosophila*. Furthermore, when using Sfmbt as the bait protein in affinity purifications, we detected interactions that could be related to Sfmbt linking PRC1 and PRC2 complexes at PREs, and/or playing the orthologous roles of either L3MBTL2 or MBTD1. We favor a model in which a hypothetical fly vPRC1.6 is not a single entity but instead encompasses modules with distinct functions that will require additional molecular analysis and dissection in the future (Fig. 4D).

Conservation of modular PRC2.1 and PRC2.2 complexes

Mammals express a core PRC2 complex, with two main alternatives incorporating mutually exclusive accessory proteins (*3*, *6*). PRC2.1 and PRC2.2 appear redundant in mammalian embryonic stem cells (ESCs), but their functions diverge during differentiation (*44*, *45*). The presence of mutually exclusive PRC2.1 and PRC2.2 accessory subunits in *Drosophila* has been inferred from several previous studies (*46*–*48*). Our previous BioTAP-XL analysis of *Drosophila* E(z) also recovered the four PRC2 core subunits, as well as Scm, Pcl, Jarid2, and Jing (*32*). The latter three proteins are the orthologs of mammalian accessory subunits PHF1/PHF19/MTF2, JARID2, and AEBP2. To confirm whether these accessory subunits form orthologous PRC2.1 and PRC2.2 complexes in *Drosophila*, we prepared transgenic fly lines expressing BioTAP-tagged Pcl, Jarid2, and Jing. The tagged transgenes rescued their corresponding homozygous lethal mutants to viability, demonstrating that they each expressed functional proteins (fig. S2A). Although pulldown of epitope-tagged Jing failed (data file S1), Pcl and Scm copurified each other but not Jarid2 and Jing, and Jarid2 only copurified Jing but not Pcl and Scm (Fig. 5A). The graphed comparison of Pcl and Jarid2 BioTAP-XL enrichments reveal the common core subunits on the diagonal and the mutually exclusive accessory subunits as top 1% interacting partners in their respective pulldowns (Fig. 5B).

Although BioTAP-XL using Jing as the bait was unsuccessful, we knew that Jing was clearly detected in our original E(z) pulldown (Fig. 5A). Therefore, we designed a sequential purification approach to further validate the existence of PRC2.2 in *Drosophila*. We split the dual BioTAP tag between two bait proteins (Fig. 5C), tagging E(z) with protein A and either Jarid2 or Jing with the biotinylation target sequence. The sequential tandem tag purifications of either ProteinA-E(z)/Bio-Jarid2 or ProteinA-E(z)/Bio-Jing resulted in relatively low numbers of total peptides but confirmed Jarid2 and Jing as genuine PRC2.2 subunits, which were mutually exclusive with Pcl and Scm (Fig. 5D).

To investigate whether the conservation of PRC2.1 and PRC2.2 in mammals and flies extends to potentially divergent functions during differentiation (*45*), we compared the genomic occupancy of BioTAP-Pcl and BioTAP-Jarid2 in 12- to 24-hour embryos, mapping the protein A epitope in each case. As expected, we found that occupancy of *Drosophila* BioTAP-Pcl correlates with stable repression and H3K27me3 (Fig. 5E and fig. S4C). In contrast, occupancy of genetically functional BioTAP-Jarid2 was of lower amplitude and dispersed across the genome (Fig. 5E and fig. S4B). Comparison to the occupancy of functional Jarid2–green fluorescent protein from the ENCODE modERN project validated the dispersed, potentially heterogenous binding pattern (fig. S4, D to G). While this unexpected result awaits further analysis, it could be consistent with a transitional rather than stable repressive role for PRC2.2, as proposed for the mammalian complex during development (*45*, *49*).

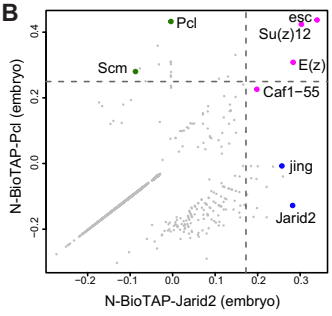
On the basis of the importance of a balance between PRC2.1 and PRC2.2 in mammalian cells (*50*), we asked whether overexpressing Jing or Jarid2 would affect normal development. We found that overexpression of *UAS-Jarid2* driven by *engrailed-GAL4* was viable with no detectable phenotype but that overexpression of *UAS-jing* yielded no adult progeny (fig. S4H). *Jarid2* partial RNA interference (RNAi) knockdown suppressed the lethality caused by overexpression of Jing (fig. S4H). As Jing protein is normally in limiting amounts compared to Jarid2 by mid-to-late embryogenesis (*51*), our result is consistent with PRC2.2 as a functional unit whose proper levels are influenced by subunit abundance and critical during development.

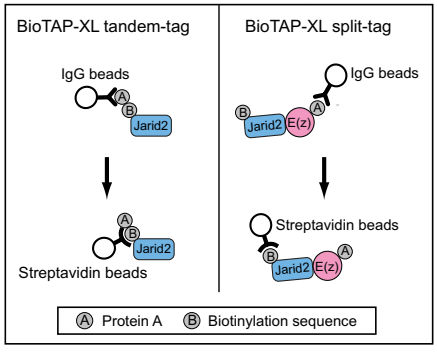
As shown in Fig. 5F, our mass spectrometry data demonstrate similar compositions of PRC2.1 and PRC2.2 in flies and mammals, with the notable substitution of vertebrate-specific PALI1 or EPOP for Scm in *Drosophila*. In previous proteomic analyses, histone methyltransferase G9a was one of the top hits from Scm BioTAP-XL purification in *Drosophila* embryos (*32*), and mammalian orthologs G9A and GLP were strongly copurified by pulldowns of BioTAP or FLAG-tagged C10ORF12 and PALI1 (*52*, *53*). Thus, as they both have strong associations with orthologous G9A methyltransferases, Scm and PALI1 may play functionally conserved roles without sequence homology. In addition, SCML2, a potential mammalian ortholog of Scm, is known to regulate PRC2 activity during spermatogenesis in mice, suggesting that a common primordial Scm-PRC2 interaction may occur in the mammalian germline (*54*).

In summary, our data provide strong support for an ancient functional diversity of PcG complexes with remarkable conservation between mammals and *Drosophila*. Historically, this was not evident, as many of the cognate mutants failed to display the classical homeotic phenotype first associated with defects in Polycomb repression in *Drosophila*. In the future, it will be interesting to dissect the germline functions and potential pleiotropy of these undoubtedly important regulators of cellular identity during development.

A

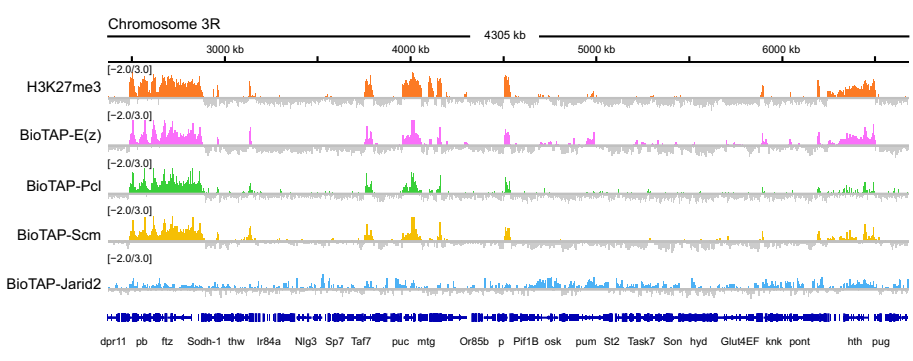
PRC2 subunits		BioTAP tag (Baits) E(z)	Pcl	Scm	Jarid2	Input
Core	Su(z)12	83	10	18	8	1
	Caf1-55	32	3	10	6	8
	Esc	39	7	6	8	1
	E(z)	23	3	7	6	1
Additional	Pcl	33	9	12	0	0
	Scm	18	4	28	0	3
	Jarid2	33	0	1	29	8
	Jing	30	0	0	4	0
Copurified total peptide #		1732	185	1760	537	10,490





D

PRC2 subunits		BioTAP-XL split-tag E(z)-Jarid2	E(z)-Jing
Core	Su(z)12	23	5
	Caf1-55	6	2
	Esc	5	6
	E(z)	5	1
Additional	Pcl	0	0
	Scm	0	0
	Jarid2	26	1
	Jing	11	4
Copurified total peptide #		581	630



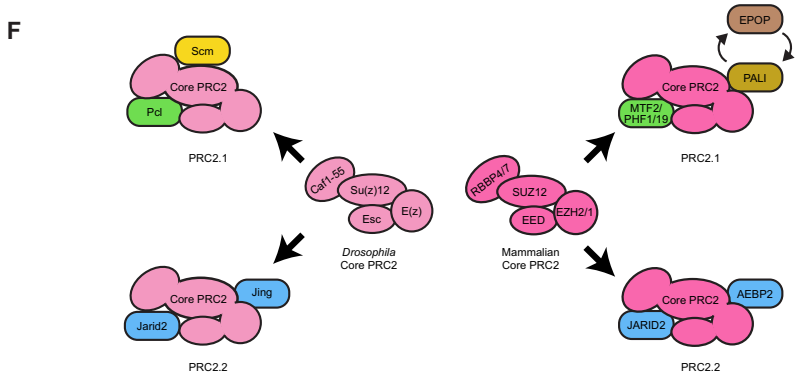


Fig. 5. *Drosophila* forms mutually exclusive PRC2.1 and PRC2.2 complexes. (**A**) Peptide counts of PRC2 subunits recovered from BioTAP-XL pulldowns of bait proteins in embryos compared with peptide counts of input. Total peptides are indicated at the bottom of the table. See data file S1 for the full range of results. (**B**) Scatterplot showing Jarid2 and Pcl pulldown enrichment (normalized to total embryonic chromatin input). PRC2 core subunits are highlighted in pink, and additional subunits of PRC2.1 and PRC2.2 are highlighted in green and blue, respectively. (**C**) Schematic of copurification strategies of BioTAP-XL tandem tag and split tag. (**D**) Peptide counts of PRC2 subunits copurified by BioTAP-XL split tag pulldowns. (**E**) Genome browser view of a 4.3-Mb region of chromosome 3R showing colocalization of Pcl and Scm with E(z) and H3K27me3 in embryos, while Jarid2 is widely dispersed across the genome. The normalized ChIP/input ratio is presented on a $\log_2$ scale. (**F**) Cartoon showing conservation of PRC2 complexes between *Drosophila* and mammals. Scm and PALI are not conserved orthologs but may have functional similarity via their separate interaction with G9A methyltransferase.

MATERIALS AND METHODS

Fly genetics

Transgenic fly lines

The pFly vector was used as the backbone to construct all transgenes (*55*). For expression of Sce, RYBP, Sfmbt, Kdm2, Pcl, E(Z), and Jing, cDNA fragments were cloned under control of the α-tubulin 1 promoter. *Jarid2* and *l(3)73Ah* were prepared from a ~13-kb genomic fragment amplified from BAC CH321-48A6 and a ~3.7-kb region amplified from *Drosophila* genomic DNA, respectively. The genomic fragments encompassed the promoter, coding, and upstream and downstream regions. In most cases, we prepared both N-terminal and C-terminal BioTAP (protein A + biotinylation target sequences) transgenic lines for each bait protein. Neither N- or C-BioTAP Sce could rescue mutant lethality. N-BioTAP–tagged RYBP, Sfmbt, Jarid2, and Pcl and C-BioTAP–tagged L(3)73Ah and Kdm2 were used for BioTAP-XL affinity purification in this study. For split tag experiments, N-terminal protein A was added to an *E(z)* transgene [N-ProteinA-E(z)], and biotinylation target sequences were added to the N terminus of either *jing* or *Jarid2* transgenes (N-Biotin-Jing and N-Biotin-Jarid2). Injections of all transgenes were performed at Bestgene Inc. (www.thebestgene.com), with the transgenes integrated into site-specific attP-docking sites by PhiC31 integrase–mediated transgenesis systems [N-BioTAP-RYBP, N-BioTAP-Sfmbt, N-BioTAP-Pcl, N-BioTAP-Jing, and N-Biotin-Jing transgenes; Bloomington Drosophila Stock Center (BDSC) no. 9732 (76A2), N- and C-BioTAP-Sce, C-BioTAP-L(3)73Ah, C-BioTAP-Kdm2, N-BioTAP-Jarid2, and N-ProteinA-E(z); BDSC no. 9736 (53B2) and N-Biotin-Jarid2; BDSC no. 9748 (62E1)] (table S1). For split tag purification, transgenic fly lines stably expressing both ProteinA-E(z) and Biotin-Jarid2 (or Biotin-Jing) were made by crossing the N-ProteinA-E(z) transgenic fly line to N-Biotin-Jarid2 or N-Biotin-Jing fly lines, followed by establishment of homozygous stocks.

Rescue tests and mutant strains

The viability rescue tests of BioTAP transgenics were described in fig. S2A. Viability was assessed by the absence of balancer markers in adult flies expressing the mini-*white* marker linked to the transgenes. Mutant fly stocks were provided by J. Müller [$Sfmbt^{1}$ and Df(2L) BSC30], H. M. Herz ($Jarid2^{e03131}$ and $Jarid2^{MB00996}$), W.W. Bender [In(2R)Pcl^{11}, Df(2R)Pcl7B, and Pcl^{R5}], and D. J. Montell ($jing^{47H6}$ and $jing^{22F3}$). Other fly stocks were obtained from the BDSC at Indiana University (table S1). Two new *l(3)73Ah* null alleles were generated using CRISPR-Cas9–mediated genome editing as described by Kondo and Ueda (*56*). Briefly, the y^{1}, w^{67c23}; *pWattB-L3-Dual-CRISPR* flies, which had a transgene expressing two single-guide RNAs [*l(3)73Ah*-gRNA1 and *l(3)73Ah*-gRNA2; sequences in table S1] from ubiquitous U6 promoter integrated at *M{3×P3-RFP.attP}ZH-51D* landing site, were crossed to y^{2},cho^{2},v^{1}; attP40{*nos-Cas9*}/*CyO* strain. Resulting *attP40{nos-Cas9}/pWattB-L3-Dual-CRISPR*; +/+ females were crossed to w^{1}; *If/CyO*; *MKRS/TM6,Tb* and the progeny individually screened for editing events by polymerase chain reaction with *l(3)73Ah*-upstr1 and *l(3)73Ah*-dwnstr1 primers (sequences in table S1), which amplify 1479–base pair (bp) fragments from unedited chromosomes and 427-bp fragments from chromosomes with expected precise deletion. Two null alleles $l(3)73Ah^{612}$ (deletion breakpoints: chr3L: 16588680-16589732) and $l(3)73Ah^{1012}$ (deletion breakpoints: chr3L: 16588681-16589549) that remove the start codon and almost the entire open reading frame were selected for further experiments.

Overexpression and RNAi knockdown

The *UAS-jing* transgenic fly line used for overexpression of Jing was a gift from J. Culi. Transgenic RNAi Project HMS02022 line and *P{Mae-UAS.6.11} Jarid2*^{LA00681} line were used for Jarid2 knockdown and overexpression, respectively. These misexpression or RNAi knockdown lines were crossed with the *engrailed (en)-GAL4* flies (gift from N. Perrimon).

Polytene chromosome immunostaining

Salivary gland polytene chromosomes from third instar larvae were prepared by first fixation for 1 min with 1% Triton X-100, 4% formaldehyde in phosphate-buffered saline, followed by second fixation for 2 min with 50% acetic acid and 4% formaldehyde before squashing to spread the polytene chromosomes. Rabbit peroxidase-antiperoxidase antibody (1:100 dilution; Sigma-Aldrich, catalog no. P1291) was used for detection of BioTAP-Sfmbt, and donkey anti-rabbit Alexa Fluor 594 (1:500 dilution; Thermo Fisher Scientific, catalog no. A32754) was used as secondary antibody.

BioTAP-XL

All BioTAP transgenic fly stocks used for BioTAP-XL experiments except for N-BioTAP-Sfmbt and C-BioTAP-Kdm2 flies had mutant backgrounds of endogenous counterpart genes. Embryos (12- to 24-hour-old) were collected on molasses agar plates in embryo collection cage (Genesee Scientific, catalog no. 59-101) and stored for up to 3 days at 4°C. The BioTAP-XL protocol was performed as described by Alekseyenko *et. al* (*57*). Briefly, 12- to 24-hour-old embryos were cross-linked with 3% formaldehyde. Nuclear extracts were prepared as described, snap-frozen in liquid nitrogen, and then stored at −80°C. These steps were repeated until nuclear extracts from 30- to 40-g embryos were pooled. After sonication of extracts, the first purification step of interaction between immunoglobulin G–agarose beads and protein A was followed by the second step binding between streptativdin-conjugated beads and biotin to purify the tagged bait proteins along with their protein interaction partners and associated genomic DNA. Bound protein complexes were trypsinized on bead, and peptides were desalted with C18-STAGE tips (3M) as described previously (*58*) for liquid chromatography–mass spectrometry analysis. Most liquid chromatography–mass spectrometry files were searched using the SEQUEST algorithm with precursor mass tolerance of 20 parts per million and fragment ion tolerance of 0.9 Da. Peptide identifications were filtered with XCorr $\geq$2 for $z = 2$ and ΔCorr $\geq$0.1. In the case of L(3)73Ah pulldown and its input, we used 0.03 Da for a fragment ion tolerance. Genomic localization of BioTAP-Pcl and -Jarid2 proteins was determined by high-throughput DNA sequencing of biological replicates generated from BioTAP-XL–purified genomic DNA using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, catalog no. E7645) and NEBNext Multiplex Oligos for Illumina (New England Biolabs, catalog no. E7335). The library samples were sequenced using an Illumina HiSeq 2500.

Proteomic analysis

To identify the proteins enriched by the BioTAP-tagged fusion proteins, a modification of the normalized spectral abundance factor was used to calculate enrichment ratios (pulldown/input) for each identified protein (data file S1) (*59*). The total number of identified peptides for a given protein were divided by the protein molecular weight (kilodalton) to control for protein size. To allow the calculation for proteins that were not recovered in the input sample, a pseudocount of 0.5 peptides was substituted for zero peptides. Then, the weight-normalized counts for each protein were divided by the

sum of weight-normalized counts across all proteins in the sample. After natural log transformation of the value, the immunoprecipitation pulldown values were divided by the input values, $\log_{10}$-normalized, and multiplied by −1 to yield enrichments of interactors for the BioTAP-bait proteins.

High-throughput DNA sequencing analysis

Raw fastq files were preprocessed with Trimmomatic (*60*) to remove Illumina adapters and quality-filtered using the fastq_quality_filter function from FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit) to retain reads with a quality score of ≥20 for at least 80% of bases. Reads were then aligned to the *Drosophila* genome (dm3) using Bowtie2 (*61*) with default parameters. SAM files were processed to BAM files, and reads with mapping quality ≥20 were retained using Samtools (*62*). BAM files were read normalized by reads per kilobase per million mapped reads, and $\log_2$ chromatin immunoprecipitation (ChIP)/input ratio binding profiles for all ChIP sequencing (ChIP-seq) replicates were generated using deepTools (*63*) bamCompare function for visualization using Integrative Genome Viewer (*64*). Correlation between biological replicates was assessed with deepTools multiBigwigSummary bins and plotCorrelation functions to compute Pearson correlation coefficient. Peaks were called using MACS2 (*65*) for each experimental sample with matched input as control and a significance cutoff of $-q$ 0.05. Overlapping peak regions were determined using the BEDTools software suite (*66*), and Venn diagrams were generated using the VennDiagram R package (*67*).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.add0103

REFERENCES AND NOTES

1. S. Aranda, G. Mas, L. Di Croce, Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* **1**, e1500737 (2015).
2. J. Gil, A. O'Loghlen, PRC1 complex diversity: Where is it taking us? *Trends Cell Biol.* **24**, 632–641 (2014).
3. D. Holoch, R. Margueron, Mechanisms regulating PRC2 recruitment and enzymatic activity. *Trends Biochem. Sci.* **42**, 531–542 (2017).
4. A. Piunti, A. Shilatifard, The roles of Polycomb repressive complexes in mammalian development and cancer. *Nat. Rev. Mol. Cell Biol.* **22**, 326–345 (2021).
5. S. A. Turner, A. P. Bracken, A "complex" issue: Deciphering the role of variant PRC1 in ESCs. *Cell Stem Cell* **12**, 145–146 (2013).
6. G. van Mierlo, G. J. C. Veenstra, M. Vermeulen, H. Marks, The complexity of PRC2 subcomplexes. *Trends Cell Biol.* **29**, 660–671 (2019).
7. M. I. Kuroda, H. Kang, S. De, J. A. Kassis, Dynamic competition of polycomb and trithorax in transcriptional programming. *Annu. Rev. Biochem.* **89**, 235–253 (2020).
8. J. M. Gahan, F. Rentzsch, C. E. Schnitzler, The genetic basis for PRC1 complex diversity emerged early in animal evolution. *Proc. Natl. Acad. Sci. U.S. A.* **117**, 22880–22889 (2020).
9. C. Alfieri, M. C. Gambetta, R. Matos, S. Glatt, P. Sehr, S. Fraterman, M. Wilm, J. Muller, C. W. Muller, Structural basis for targeting the chromatin repressor Sfmbt to polycomb response elements. *Genes Dev.* **27**, 2367–2379 (2013).
10. A. Lagarou, A. Mohd-Sarip, Y. M. Moshkin, G. E. Chalkley, K. Bezstarosti, J. A. Demmers, C. P. Verrijzer, dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes Dev.* **22**, 2799–2810 (2008).
11. I. Cohen, D. Zhao, C. Bar, V. J. Valdes, K. L. Dauber-Decker, M. B. Nguyen, M. Nakayama, M. Rendl, W. A. Bickmore, H. Koseki, D. Zheng, E. Ezhkova, PRC1 fine-tunes gene repression and activation to safeguard skin development and stem cell specification. *Cell Stem Cell* **22**, 726–739.e7 (2018).
12. N. A. Fursova, N. P. Blackledge, M. Nakayama, S. Ito, Y. Koseki, A. M. Farcas, H. W. King, H. Koseki, R. J. Klose, Synergy between variant PRC1 complexes defines polycomb-mediated gene repression. *Mol. Cell* **74**, 1020–1036.e8 (2019).
13. Z. Gao, P. Lee, J. M. Stafford, M. von Schimmelmann, A. Schaefer, D. Reinberg, An AUTS2-Polycomb complex activates gene expression in the CNS. *Nature* **516**, 349–354 (2014).
14. Z. Gao, J. Zhang, R. Bonasio, F. Strino, A. Sawai, F. Parisi, Y. Kluger, D. Reinberg, PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45**, 344–356 (2012).
15. A. Scelfo, D. Fernandez-Perez, S. Tamburri, M. Zanotti, E. Lavarone, M. Soldi, T. Bonaldi, K. J. Ferrari, D. Pasini, Functional landscape of PCGF proteins reveals both RING1A/B-dependent-and RING1A/B-independent-specific activities. *Mol. Cell* **74**, 1037–1052.e7 (2019).
16. W. Zhao, Y. Huang, J. Zhang, M. Liu, H. Ji, C. Wang, N. Cao, C. Li, Y. Xia, Q. Jiang, J. Qin, Polycomb group RING finger proteins 3/5 activate transcription via an interaction with the pluripotency factor Tex10 in embryonic stem cells. *J. Biol. Chem.* **292**, 21527–21537 (2017).
17. R. Wang, A. B. Taylor, B. Z. Leal, L. V. Chadwell, U. Ilangovan, A. K. Robinson, V. Schirf, P. J. Hart, E. M. Lafer, B. Demeler, A. P. Hinck, D. G. McEwen, C. A. Kim, Polycomb group targeting through different binding partners of RING1B C-terminal domain. *Structure* **18**, 966–975 (2010).
18. L. Tavares, E. Dimitrova, D. Oxley, J. Webster, R. Poot, J. Demmers, K. Bezstarosti, S. Taylor, H. Ura, H. Koide, A. Wutz, M. Vidal, S. Elderkin, N. Brockdorff, RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148**, 664–678 (2012).
19. S. Fereres, R. Simon, A. Mohd-Sarip, C. P. Verrijzer, A. Busturia, dRYBP counteracts chromatin-dependent activation and repression of transcription. *PLOS ONE* **9**, e113255 (2014).
20. D. Beuchle, G. Struhl, J. Muller, Polycomb group proteins and heritable silencing of Drosophila Hox genes. *Development* **128**, 993–1004 (2001).
21. J. A. Kassis, J. A. Kennison, J. W. Tamkun, Polycomb and trithorax group genes in *Drosophila*. *Genetics* **206**, 1699–1725 (2017).
22. I. F. King, R. B. Emmons, N. J. Francis, B. Wild, J. Muller, R. E. Kingston, C. T. Wu, Analysis of a polycomb group protein defines regions that link repressive activity on nucleosomal templates to in vivo function. *Mol. Cell. Biol.* **25**, 6578–6591 (2005).
23. S. M. Lo, N. K. Ahuja, N. J. Francis, Polycomb group protein suppressor 2 of zeste is a functional homolog of posterior sex combs. *Mol. Cell. Biol.* **29**, 515–525 (2009).
24. H. G. Lee, T. G. Kahn, A. Simcox, Y. B. Schwartz, V. Pirrotta, Genome-wide activities of Polycomb complexes control pervasive transcription. *Genome Res.* **25**, 1170–1181 (2015).
25. Y. Zheng, Y. Xue, X. Ren, M. Liu, X. Li, Y. Jia, Y. Niu, J. Q. Ni, Y. Zhang, J. Y. Ji, The lysine demethylase dKDM2 is non-essential for viability, but regulates circadian rhythms in drosophila. *Front. Genet.* **9**, 354 (2018).
26. B. Poeck, T. Triphan, K. Neuser, R. Strauss, Locomotor control by the central complex in drosophila-an analysis of the tay bridge mutant. *Dev. Neurobiol.* **68**, 1046–1058 (2008).
27. N. Oksenberg, N. Ahituv, The role of AUTS2 in neurodevelopment and human evolution. *Trends Genet.* **29**, 600–608 (2013).
28. Z. Zhang, A. E. Jones, W. Wu, J. Kim, Y. Kang, X. Bi, Y. Gu, I. K. Popov, M. B. Renfrow, M. N. Vassylyeva, D. G. Vassylyev, K. E. Giles, D. Chen, A. Kumar, Y. Fan, Y. Tong, C. F. Liu, W. An, C. Chang, J. Luo, L. T. Chow, H. Wang, Role of remodeling and spacing factor 1 in histone H2A ubiquitination-mediated gene silencing. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7949–E7958 (2017).
29. T. Klymenko, B. Papp, W. Fischle, T. Kocher, M. Schelder, C. Fritsch, B. Wild, M. Wilm, J. Muller, A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev.* **20**, 1110–1122 (2006).
30. P. J. Hurlin, E. Steingrimsson, N. G. Copeland, N. A. Jenkins, R. N. Eisenman, Mga, a dual-specificity transcription factor that interacts with Max and contains a T-domain DNA-binding motif. *EMBO J.* **18**, 7019–7028 (1999).
31. C. A. Kim, J. U. Bowie, SAM domains: Uniform structure, diversity of function. *Trends Biochem. Sci.* **28**, 625–628 (2003).
32. H. Kang, K. A. McElroy, Y. L. Jung, A. A. Alekseyenko, B. M. Zee, P. J. Park, M. I. Kuroda, Sex comb on midleg (Scm) is a functional link between PcG-repressive complexes in Drosophila. *Genes Dev.* **29**, 1136–1150 (2015).
33. F. Frey, T. Sheahan, K. Finkl, G. Stoehr, M. Mann, C. Benda, J. Muller, Molecular basis of PRC1 targeting to Polycomb response elements by PhoRC. *Genes Dev.* **30**, 1116–1127 (2016).
34. G. Strubbe, C. Popp, A. Schmidt, A. Pauli, L. Ringrose, C. Beisel, R. Paro, Polycomb purification by in vivo biotinylation tagging reveals cohesin and Trithorax group proteins as interaction partners. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5572–5577 (2011).
35. J. L. Brown, C. Fritsch, J. Mueller, J. A. Kassis, The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. *Development* **130**, 285–294 (2003).
36. J. A. Kassis, J. L. Brown, Polycomb group response elements in Drosophila and vertebrates. *Adv. Genet.* **81**, 83–118 (2013).
37. K. K. Langlais, J. L. Brown, J. A. Kassis, Polycomb group proteins bind an engrailed PRE in both the "ON" and "OFF" transcriptional states of engrailed. *PLOS ONE* **7**, e48765 (2012).
38. M. Fujioka, G. L. Yusibova, J. Zhou, J. B. Jaynes, The DNA-binding Polycomb-group protein Pleiohomeotic maintains both active and repressed transcriptional states through a single site. *Development* **135**, 4131–4139 (2008).

39. E. Ghotbi, P. Ye, T. Ervin, A. Kum, J. Benes, R. S. Jones, Polycomb-group recruitment to a *Drosophila* target gene is the default state that is inhibited by a transcriptional activator. *Sci. Adv.* **7**, (2021).
40. T. G. Kahn, P. Stenberg, V. Pirrotta, Y. B. Schwartz, Combinatorial interactions are required for the efficient recruitment of pho repressive complex (PhoRC) to polycomb response elements. *PLOS Genet.* **10**, e1004495 (2014).
41. V. Loubiere, G. L. Papadopoulos, Q. Szabo, A. M. Martinez, G. Cavalli, Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping. *Sci. Adv.* **6**, eaax4001 (2020).
42. B. Schuettengruber, M. Ganapathi, B. Leblanc, M. Portoso, R. Jaschek, B. Tolhuis, M. van Lohuizen, A. Tanay, G. Cavalli, Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. *PLoS Biol.* **7**, e13 (2009).
43. Y. Shi, E. Seto, L. S. Chang, T. Shenk, Transcriptional repression by YY1, a human GLI-Krüippel-related protein, and relief of repression by adenovirus E1A protein. *Cell* **67**, 377–388 (1991).
44. E. Healy, M. Mucha, E. Glancy, D. J. Fitzpatrick, E. Conway, H. K. Neikes, C. Monger, G. Van Mierlo, M. P. Baltissen, Y. Koseki, M. Vermeulen, H. Koseki, A. P. Bracken, PRC2.1 and PRC2.2 synergize to coordinate H3K27 trimethylation. *Mol. Cell* **76**, 437–452.e6 (2019).
45. A. Petracovici, R. Bonasio, Distinct PRC2 subunits regulate maintenance and establishment of Polycomb repression during differentiation. *Mol. Cell* **81**, 2625–2639.e5 (2021).
46. H. M. Herz, M. Mohan, A. S. Garrett, C. Miller, D. Casto, Y. Zhang, C. Seidel, J. S. Haug, L. Florens, M. P. Washburn, M. Yamaguchi, R. Shiekhattar, A. Shilatifard, Polycomb repressive complex 2-dependent and -independent functions of Jarid2 in transcriptional regulation in Drosophila. *Mol. Cell. Biol.* **32**, 1683–1693 (2012).
47. R. Kalb, S. Latwiel, H. I. Baymaz, P. W. Jansen, C. W. Muller, M. Vermeulen, J. Muller, Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* **21**, 569–571 (2014).
48. M. Nekrasov, T. Klymenko, S. Fraterman, B. Papp, K. Oktaba, T. Kocher, A. Cohen, H. G. Stunnenberg, M. Wilm, J. Muller, Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO J.* **26**, 4078–4088 (2007).
49. V. Kasinath, C. Beck, P. Sauer, S. Poepsel, J. Kosmatka, M. Faini, D. Toso, R. Aebersold, E. Nogales, JARID2 and AEBP2 regulate PRC2 in the presence of H2AK119ub1 and other histone modifications. *Science* **371**, eabc3393 (2021).
50. D. T. Youmans, A. R. Gooding, R. D. Dowell, T. R. Cech, Competition between PRC2.1 and 2.2 subcomplexes regulates PRC2 chromatin occupancy in human stem cells. *Mol. Cell* **81**, 488–501.e9 (2021).
51. J. Bonnet, R. G. H. Lindeboom, D. Pokrovsky, G. Stricker, M. H. Celik, R. A. W. Rupp, J. Gagneur, M. Vermeulen, A. Imhof, J. Muller, Quantification of proteins and histone marks in drosophila embryos reveals stoichiometric relationships impacting chromatin regulation. *Dev. Cell* **51**, 632–644.e6 (2019).
52. A. A. Alekseyenko, A. A. Gorchakov, P. V. Kharchenko, M. I. Kuroda, Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2488–2493 (2014).
53. E. Conway, E. Jerman, E. Healy, S. Ito, D. Holoch, G. Oliviero, O. Deevy, E. Glancy, D. J. Fitzpatrick, M. Mucha, A. Watson, A. M. Rice, P. Chammas, C. Huang, I. Pratt-Kelly, Y. Koseki, M. Nakayama, T. Ishikura, G. Streubel, K. Wynne, K. Hokamp, A. McLysaght, C. Ciferri, L. Di Croce, G. Cagney, R. Margueron, H. Koseki, A. P. Bracken, A family of vertebrate-specific polycombs encoded by the LCOR/LCORL genes balance PRC2 subtype activities. *Mol. Cell* **70**, 408–421.e8 (2018).
54. S. Maezawa, K. Hasegawa, M. Yukawa, N. Kubo, A. Sakashita, K. G. Alavattam, H. S. Sin, A. V. Kartashov, H. Sasaki, A. Barski, S. H. Namekawa, Polycomb protein SCML2 facilitates H3K27me3 to establish bivalent domains in the male germline. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4957–4962 (2018).
55. C. I. Wang, A. A. Alekseyenko, G. LeRoy, A. E. Elia, A. A. Gorchakov, L. M. Britton, S. J. Elledge, P. V. Kharchenko, B. A. Garcia, M. I. Kuroda, Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila. *Nat. Struct. Mol. Biol.* **20**, 202–209 (2013).
56. S. Kondo, R. Ueda, Highly improved gene targeting by germline-specific Cas9 expression in Drosophila. *Genetics* **195**, 715–721 (2013).
57. A. A. Alekseyenko, K. A. McElroy, H. Kang, B. M. Zee, P. V. Kharchenko, M. I. Kuroda, BioTAP-XL: Cross-linking/tandem affinity purification to study DNA targets, RNA, and protein components of chromatin-associated complexes. *Curr. Protoc. Mol. Biol.* **109**, 213021–213032 (2015).
58. B. M. Zee, A. A. Alekseyenko, K. A. McElroy, M. I. Kuroda, Streamlined discovery of cross-linked chromatin complexes and associated histone modifications by mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1784–1789 (2016).
59. B. Zybailov, A. L. Mosley, M. E. Sardiu, M. K. Coleman, L. Florens, M. P. Washburn, Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J. Proteome Res.* **5**, 2339–2347 (2006).
60. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
61. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
62. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, S. Genome, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
63. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
64. J. T. Robinson, H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
65. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
66. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
67. H. Chen, P. C. Boutros, VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35 (2011).
68. E. W. Deutsch, N. Bandeira, V. Sharma, Y. Perez-Riverol, J. J. Carver, D. J. Kundu, D. Garcia-Seisdedos, A. F. Jarnuczak, S. Hewapathirana, B. S. Pullman, J. Wertz, Z. Sun, S. Kawano, S. Okuda, Y. Watanabe, H. Hermjakob, B. MacLean, M. J. MacCoss, Y. Zhu, Y. Ishihama, J. A. Vizcaino, The ProteomeXchange consortium in 2020: Enabling 'big data' approaches in proteomics. *Nucleic Acids Res.* **48**, D1145–D1152 (2020).
69. Y. Perez-Riverol, J. Bai, C. Bandla, D. Garcia-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D. J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S. Wang, A. Brazma, J. A. Vizcaino, The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).

Acknowledgments: We thank R. Tomaino for expert help with protein samples (Taplin Mass Spectrometry Facility, Harvard Medical School); J. Alhaj Abed for initial work on *jing* fly crosses; and J. Müller, H. M. Herz, W. Bender, J. Culi, and D. Montell for sharing cDNA and/or fly stocks. We are grateful to J. Kassis, J. Simon, S. DeLuca, and J. Müller for many helpful discussions. We thank the TRiP at Harvard Medical School (NIH/National Institute of General Medical Sciences R01-GM084947) and the BDSC at Indiana University (NIH P40OD018537) for providing fly stocks used in this study. **Funding:** This work was supported by an NIH grant to M.I.K. (R35-GM126944) and a Swedish Research Council grant to Y.B.S. (2021-04435). **Author contributions:** Conceptualization: H.K., J.R.C., H.A.K., and M.I.K. Methodology: H.K., J.R.C., and M.I.K. Investigation: H.K., J.R.C., B.M.Z., H.A.K., J.M.J., M.B.H., A.E.B., and A.G. Visualization: H.K. Supervision: H.K., J.R.C., Y.B.S., and M.I.K. Writing (original draft): H.K. and M.I.K. Writing (review and editing): H.K., H.A.K., B.M.Z., Y.B.S., and M.I.K. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The mass spectrometry proteomics data of BioTAP fusion proteins have been deposited to the ProteomeXchange Consortium (*68*) via the PRIDE (*69*) partner repository with the dataset identifier PXD033615 and 10.6019/PXD033615. High-throughput DNA sequencing data for BioTAP-Jarid2 and BioTAP-Pcl have been submitted to the NCBI Gene Expression Omnibus (GEO) repository under the accession number GSE201842. High-throughput DNA sequencing data for BioTAP-E(z) and BioTAP-Scm were generated in our previous study (*32*); raw sequencing files are deposited in GEO (accession number GSE66183). ChIP-seq files for H3K27me3, H3K27me2, and H3K27ac were obtained from GEO (GSE47230, GSE47277, and GSE47237, respectively). ChIP-seq files for Jarid2–green fluorescent protein were obtained from http://encodeproject.org (accession numbers: ENCFF090YBK, ENCFF064HTW, ENCFF649NYX, and ENCSR388YUZ). Drosophila stocks newly created in this study are available upon request.

Submitted 14 May 2022
Accepted 22 July 2022
Published 7 September 2022
10.1126/sciadv.add0103