Roles of the Brca2 and Wapl complexes with Pds5 in sister chromatid cohesion, cohesin localization, and gene expression

Ziva Misulovin,[1] Michelle Pherson,[1] Maria Gause,[1] Kathie Mihindukulasuriya,[1] and Dale Dorsett[1]*

[1]Edward A Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, Saint Louis, Missouri, USA

*Corresponding Author: dorsettd@slu.edu

Running Title: Pds5, Wapl, Brca2 and cohesin function

Key words: enhancer, promoter, Polycomb Response Element, PRE, transcription, topologically-associating domain, TAD, cohesinopathy

Abstract

The cohesin complex topologically encircles chromosomes and mediates sister chromatid cohesion to ensure accurate chromosome segregation upon cell division. Cohesin also participates in DNA repair and gene transcription. The Nipped-B protein loads cohesin onto chromosomes and the Pds5-Wapl complex removes cohesin. Pds5 is also essential for sister chromatid cohesion, indicating that it has functions beyond cohesin removal. The Brca2 DNA repair protein interacts with Pds5 but the roles of this complex are poorly understood. Here we show that Brca2 opposes Pds5 function in sister chromatid cohesion but has similar effects on gene expression. Pds5 facilitates SA cohesin subunit association with DNA replication origins and Brca2 inhibits SA binding, implying that SA levels at origins govern sister chromatid cohesion. Cohesin binding is maximal at replication origins and extends outward to occupy active genes and regulatory sequences. Pds5 and Wapl, but not Brca2, limit the distance that cohesin extends from origins, thereby determining which active genes bind cohesin. We posit that DNA replication pushes cohesin along chromosomes, and that the Pds5-Wapl complex unloads cohesin at replication forks to control the size of cohesin domains. In contrast to their opposing roles in chromatid cohesion, Pds5 and Brca2 have similar effects on gene expression. These effects overlap those of Nipped-B, cohesin and Wapl, indicating that Pds5 and Brca2 modify cohesin dynamics and function at active genes. These findings demonstrate that Brca2 regulates sister chromatid cohesion and gene expression in addition to its canonical role in DNA repair and expand the known functions of accessory proteins in cohesin function.

Author summary

The cohesin protein complex has multiple functions in eukaryotic cells. It ensures that when a cell divides, the two daughter cells receive the correct number of chromosomes. It does this by holding together the sister chromatids that are formed when chromosomes are duplicated by DNA replication. Cohesin also helps repair damaged DNA, and to regulate genes important for growth and development. Even minor deficiencies in some proteins that regulate cohesin cause significant birth defects in humans. Here we investigated in Drosophila how three proteins, Pds5, Wapl and Brca2, determine where cohesin binds to chromosomes, control cohesin's ability to hold sister chromatids together, and to regulate gene expression. We find that Pds5 and Wapl work together, likely during DNA replication, to determine which genes bind cohesin by controlling how far cohesin spreads out along chromosomes. Pds5 is required for cohesin to hold sister chromatids together, and Brca2 counteracts this function. In contrast to the opposing roles in sister chromatid cohesion, Pds5 and Brca2 work together to facilitate control of gene expression by cohesin. Brca2 plays a critical role in DNA repair, and these studies expand the known roles for Brca2 by showing that it also regulates sister chromatid cohesion and gene expression. *BRCA2* mutations in humans increase susceptibility to breast and ovarian cancer, and these findings raise the possibility that changes in chromosome segregation or gene expression might contribute to the increased cancer risk associated with these mutations.

Introduction

The cohesin complex mediates sister chromatid cohesion, which ensures accurate chromosome segregation upon cell division. Cohesin consists of the Smc1 (Flybase FBgn0040283), Smc3 (Flybase FBgn0015615), Rad21 (Vtd, Flybase FBgn0260987) and SA (Flybase FBgn0020616) proteins, which form a ring-like structure that topologically encircles chromosomes [1-4]. Smc1 and Smc3 form a heterodimer, Rad21 bridges their ATPase head domains to close the ring, and SA interacts with Rad21. In metazoan organisms, cohesin is loaded onto chromosomes by a complex of the Nipped-B (Flybase FBgn0026401) and Mau2 (Flybase FBgn0038300) proteins starting in early G1, and sister chromatid cohesion is established during S phase. Cohesin is removed from chromosome arms by a complex of the Pds5 (Flybase FBgn0260012) and Wapl (Flybase FBgn0004655) proteins upon entry into mitosis.

Although the Pds5-Wapl complex unloads cohesin from chromosomes, Pds5 and Wapl differ in their roles in sister chromatid cohesion. Mutations in the Drosophila *pds5* gene cause loss of sister cohesion, resulting in aneuploid cells prior to death [5]. In contrast, *wapl* mutations cause premature loss of sister cohesion only in pericentric heterochromatic regions and impair release of cohesion along chromosome arms [6].

Pds5 and Wapl also control cohesin chromosome binding dynamics during interphase. In vivo FRAP (fluorescence recovery after photobleaching) experiments show that partial reduction of Pds5 dosage increases the level of cohesin stably bound to chromosomes, consistent with its role in cohesin removal, while partial loss of Wapl unexpectedly decreases the level of stable cohesin [7]. Complete removal of Wapl, however, dramatically increases both the amount of cohesin on chromosomes and its average residence time [7]. Thus the collaborative roles of Pds5 and Wapl in cohesin removal alone do not easily explain their in vivo effects on cohesin dynamics and sister chromatid cohesion.

Pds5 also interacts in a 2:1 ratio with the Brca2 (Flybase FBgn0050169) DNA repair protein [8] and this complex participates in meiotic recombination and mitotic DNA repair [8-10]. The potential roles of the Pds5-Brca2 complex in cohesin dynamics, sister chromatid cohesin and gene transcription have not been explored.

We investigated the roles of Pds5, Brca2 and Wapl in cohesin binding and localization, sister chromatid cohesion, and gene regulation in cultured Drosophila ML-BG3-c2 (BG3) cells derived from larval central nervous system. We find that Brca2 and Pds5 oppose each other in sister chromatid cohesion and binding of the SA cohesin subunit to DNA replication origins, implying that SA levels at origins are a crucial determinant of sister chromatid cohesion. The data further show that cohesin extends outward from replication origins for several kilobases to bind active genes and regulatory sequences, and that Pds5 and Wapl limit the distance that cohesin spreads, determining which genes bind cohesin. We posit that DNA replication pushes cohesin along chromosomes, and that the Pds5-Wapl complex unloads cohesin at replication forks to constrain cohesin spreading. Brca2 and Pds5 have very similar effects on gene expression in striking contrast to their opposite effects on sister chromatid cohesion. They

influence expression of the same genes as Nipped-B and cohesin, indicating that they modulate cohesin activity at genes and regulatory sequences. We further find that Brca2 influences expression of many genes involved in DNA repair, transcription and morphogenesis in developing wings. Combined, the roles of Brca2 in regulation of sister chromatid cohesion and gene expression revealed in these studies suggest that the increased cancer susceptibility caused by human *BRCA2* mutations could potentially reflect changes in cell physiology beyond DNA repair deficits.

Results

Brca2 counteracts the role of Pds5 in sister chromatid cohesion

The discovery that Pds5 interacts with Brca2 in a complex that lacks detectable Wapl [8] raised the question of whether or not Brca2 plays a role in sister chromatid cohesion. RNAi-mediated depletion of Brca2 in cultured ML-DmBG3-c2 (BG3) cells indicates that it opposes the role of Pds5 in establishing or maintaining cohesion. BG3 cells are diploid male, allowing accurate measurement of sister chromatid cohesion and segregation defects in metaphase chromosome spreads [11]. Figure 1A shows that as expected, depletion of Pds5 (iPds5) causes precocious sister chromatid separation (PSCS) with ~70% of chromosomes showing partial or complete separation after 5 days of treatment, and Wapl depletion (iWapl) gives less than 10% PSCS, similar to mock-treated cells. Brca2 depletion (iBrca2) alone also does not alter PSCS frequency (Fig 1A) but if Brca2 is co-depleted with Pds5 (iPds5 iBrca2) PSCS frequency is significantly reduced relative to Pds5 depletion (Fig 1B). Western blots illustrating Pds5, Wapl and Brca2 depletion are shown in S1 Fig. Pds5 depletion reduces cell proliferation after 5 days of RNAi, while there is no discernable effect of Wapl or Brca2 depletion on cell growth. We conclude that Brca2 opposes the ability of Pds5 to support sister chromatid cohesion.

Pds5 and Brca2 chromosome binding patterns differ from cohesin and Wapl

We used chromatin immunoprecipitation with high-throughput sequencing (ChIP-seq) to map Pds5, Wapl, and Brca2 genome-wide in cultured Drosophila BG3 cells to gain further insights into their roles in regulating cohesin function. We compared their binding patterns to the Rad21 cohesin subunit, and prior ChIP-seq data [12] for the SA cohesin subunit and Nipped-B. As illustrated by the *kayak* gene region in Figure 2A, Wapl displays a broad spreading pattern very similar to that of cohesin. Nipped-B differs somewhat from cohesin and Wapl in that Nipped-B has noticeable peaks superimposed on the spreading pattern. Pds5 and Brca2 do not show significant spreading, but are largely restricted to peaks that co-localize with Nipped-B peaks (Fig 2A).

Sister chromatid cohesion factor occupancy centers at DNA replication origins

Sister chromatid cohesion is established during DNA replication. Meta-analysis of the cohesion factor ChIP-seq data shows that Nipped-B, cohesin, Pds5, Wapl and Brca2 all show the highest enrichment at DNA replication origins. The locations of early DNA replication origins in BG3

cells were reported by the modENCODE project and confirmed by mapping origin recognition complex binding sites [13]. As shown in Figure 2A, a replication origin is located within multiple active genes that bind RNA polymerase II (Rpb3) [14] in the *kayak* region. We averaged the ChIP-seq enrichment for cohesin and accessory factors in 10 kb bins from -100 kb to +100 kb outward from the centers of the 78 strongest early origins in BG3 cells (positions in S2 File). As revealed in Figure 2B, all show maximal occupancy at the meta-origin center, with the levels decreasing outward in both directions. Cohesin binds specifically to active genes and enhancers and not to inactive genes or intergenic regions [15, 16] so this meta-analysis reveals that the further a gene is from the origin, the less likely it is to bind cohesin. For instance, *kayak* under the origin center binds cohesin, and *yata* some 20 kb from the origin center binds little, even though it is transcribed (Fig 2A).

Pds5 and Wapl control cohesin localization around replication origins

We examined the effects of depleting Pds5, Wapl and Brca2 on binding of cohesin and Nipped-B to chromosomes using ChIP-seq and meta-origin analysis to determine how they influence cohesin localization and binding levels. This revealed that Pds5 and Wapl depletion increase the distance that the domains of cohesin occupancy extend outward from origins. Figure 3A shows that Pds5 depletion (iPds5) increases spreading of the Rad21 and SA cohesin subunits, Wapl, and Nipped-B in the regions flanking the *kayak* replication origin. Meta-origin analysis of the Rad21 data shows that this effect is global, with statistically significant increases in occupancy for tens of kilobases in the flanking regions (Fig 3B). Wapl depletion (iWapl) similarly increases cohesin spreading around origins as illustrated for the SA cohesin subunit at *kayak* locus in Figure 4A and meta-origin analysis in S3A Figure. Brca2 depletion (iBrca2) does not increase the size of cohesin domains (Fig 4A). Although Pds5 depletion reduces sister chromatid cohesion, Wapl depletion does not. We thus conclude that the expansion of cohesin territories caused by these depletions does not reflect loss of sister cohesion.

The cohesin domain expansion revealed by meta-origin analysis reflects association of cohesin with active genes that normally bind little or no cohesin, such as *yata* in Figure 4A. Inactive genes and intergenic regions in the extended cohesin domains do not bind cohesin. The skipping over inactive genes and intergenic regions in the extended cohesin domains is illustrated at the *string* locus in S4A Figure. Thus proteins at active genes are required for cohesin binding in the expanded domains. We posit that many active genes have the potential to bind cohesin, which is limited by their proximity to a replication origin.

Cohesin domain expansion upon Pds5 or Wapl depletion may be linked to changes in chromosome architecture. The ends of the cohesin domain at the Enhancer of split gene complex coincide with the borders of a topologically-associating domain (TAD) [17]. Pds5 or Wapl depletion causes cohesin to extend beyond these borders (S4B Fig). Thus Pds5 or Wapl depletion might alter the locations of the TAD boundaries, or alternatively, reduce the abilities of the boundaries to block cohesin spreading. DNA loop extrusion through cohesin-CTCF insulator complexes can form TADs in mammalian cells and Wapl restricts loop extrusion [18] but a

similar mechanism appears unlikely at Enhancer of split where TAD structure is independent of cohesin or insulator proteins [17].

The finding that active genes closer to origins are more likely to bind cohesin suggests that DNA replication is a key factor that determines which genes bind cohesin. DNA replisomes push topologically-bound cohesin along naked DNA in single molecule studies [19] suggesting that cohesin loaded at origins may be pushed outward by replication forks. If so, then the cohesin domain extension caused by Pds5 or Wapl depletion could reflect either a reduced rate of cohesin removal in front of replication forks, or increased speed of replication fork movement. DNA fiber assays, however, show that Pds5 depletion does not alter fork speed (S5 Fig). Brca2 depletion also does not alter replication speed, but as reported for mammalian cells [20] increases degradation of newly-replicated DNA behind stalled forks (S5 Fig). Thus increased cohesin spreading does not reflect an increase in replication fork speed, but may be caused by a reduced rate of cohesin removal at replication forks.

Pds5 and Brca2 oppositely regulate SA cohesin subunit levels at DNA replication origins

Meta-origin analysis unexpectedly revealed that the ratio of the SA cohesin subunit to the Rad21 subunit is elevated at origin centers relative to the flanking regions (Fig 4B). Moreover, Pds5 depletion reduces SA levels at origin centers (Fig 4C) and equalizes the SA to Rad21 ratio across the origin and flanking regions (Fig 4B). In contrast, Wapl depletion does not reduce origin SA levels (S3A Fig) and Brca2 depletion increases SA at origins (Fig 4D). When Pds5 and Brca2 are co-depleted, their individual effects cancel out, leaving little net change in SA levels at origin centers (Fig 4E). This makes the extension of cohesin domains more apparent with SA, giving a pattern similar to that observed with Rad21 with significant increases over several kilobases in the flanking regions (Fig 3B). These findings indicate that there is increased SA relative to cohesin at the origins, that Pds5 is required for the extra SA, and that Brca2 counteracts this Pds5 function.

The opposing roles of Pds5 and Brca2 in controlling SA occupancy of origins (Fig 4) parallels their opposing effects on sister chromatid cohesion (Fig 1B). We thus posit that Pds5 and Brca2 determine the degree of sister chromatid cohesion along chromosome arms by adjusting SA levels at origins. Although other mechanisms of sister chromatid cohesion are possible, this is consistent with the idea that SA can link two cohesin rings to establish cohesion via a handcuff mechanism [21, 22].

Pds5 depletion increases Nipped-B levels at origins and extends Nipped-B and Wapl binding domains

We considered the possibility that some effects of Pds5, Wapl, and Brca2 depletion on cohesin levels and distribution around origins could reflect how these factors influence each other's association with origins, or binding of the Nipped-B cohesin loader. As described below, the cohesin accessory factors influence each other's association with replication origins, but the changes do not result in the expected effects on cohesin levels. This leads us to postulate that

other factors, potentially the pre-replication complex (pre-RC) determine cohesin levels at origins.

The effect of Pds5 depletion on Wapl mirrors the effects it has on SA. The Wapl to Rad21 ratio, like the SA to Rad21 ratio, is higher at origin centers than in flanking regions, and Pds5 depletion equalizes this ratio across the origin and flanking regions (Fig 5A). Similar to the effect on SA, Pds5 depletion decreases Wapl at origins with moderate increases in flanking regions (Fig 5B). In contrast, Pds5 depletion has only minor effects on Brca2 levels at origins and flanking regions (Fig 5C). Because the Pds5-Wapl complex removes cohesin from chromosomes, depletion of Pds5 with the accompanying decrease in Wapl should increase cohesin levels at origins, but as shown above, SA decreases (Fig 4C) and Rad21 is not altered (Fig 3B). This indicates that Pds5 and Wapl do not effectively remove cohesin at origins.

The Nipped-B to Rad21 ratio is constant across the 200 kb meta-origin region, but increases at the origin upon Pds5 depletion, accompanied by an overall increase in Nipped-B levels across the 200 kb meta-origin (Fig 5D,E). We theorize that Pds5 and Nipped-B, which have similar HEAT repeat structures, compete for binding to cohesin or other proteins at origins, and loss of Pds5 allows more Nipped-B to bind. Competition between Nipped-B and Pds5 for binding to cohesin is supported by in vivo FRAP studies [7] and biochemical experiments showing that the Nipped-B and Pds5 orthologs in *C thermophilium* bind to overlapping regions in Rad21 [23]. Higher Nipped-B occupancy caused by Pds5 depletion, similar to lower Pds5 and Wapl levels, should also increase cohesin levels at origins by combining increased cohesin loading with decreased cohesin removal, but cohesin increases occur only in the flanking regions, not at origin centers.

Wapl depletion increases Pds5 (S3B Fig) and decreases Nipped-B occupancy at origins (S3C Fig). Based on the observations suggesting that Pds5 and Nipped-B compete for binding to origins, we posit that the increase in Pds5 caused by Wapl depletion could cause the decrease in Nipped-B.

The increase in Pds5 upon Wapl depletion also shows that Wapl does not recruit Pds5 to chromosomes. Instead, Brca2 is more crucial, as Brca2 depletion significantly decreases Pds5 levels (S3D Fig) but has little effect on Wapl (S3E Fig). Nipped-B and Rad21 depletion also decrease Pds5 binding (S3F,G Fig) indicating that interaction of Pds5 with cohesin is also important for Pds5 binding to origins and the flanking regions.

The picture that emerges is that in contrast to the effects on cohesin occupancy in regions flanking origins, the effects of Pds5 and Wapl depletion on cohesin levels at origins are not explained by their known roles in removing cohesin from chromosomes, or their effects on levels of the Nipped-B loading factor. We thus posit that other factors, potentially DNA replication proteins such as the origin recognition complex (ORC) or other components of the pre-RC are primary determinants of cohesin occupancy at origins.

The ratios of accessory factors to cohesin vary between different types of gene regulatory sequences

In the regions flanking replication origins cohesin and Nipped-B bind primarily to active genes, transcriptional enhancers and Polycomb Response Elements (PREs) and influence gene activation and repression [11, 12, 14-16, 18, 24]. We find that the relative levels of the different cohesin regulatory factors vary between active promoters, enhancers and PREs (Fig 6). The mean ChIP-seq enrichment for these proteins at all active promoters, extragenic enhancers and PREs in BG3 cells were calculated as previously described [12]. The Rad21 and SA cohesin subunits show the highest levels at enhancers, followed by PREs, and active promoters (Fig 6A,B). In contrast, Nipped-B has the highest levels at PREs, followed by enhancers and promoters (Fig 6C). Pds5 is highest at enhancers, followed by promoters, and PREs (Fig 6D) while Wapl shows the same pattern as cohesin, being higher at PREs than promoters (Fig 6E). Surprisingly, the Brca2 pattern is more similar to Nipped-B than it is to Pds5, being highest at PREs, followed by enhancers and promoters (Fig 6F). The variation in the relative amounts of cohesin and accessory factors at different types of regulatory sequences imply that other proteins at the regulatory sequences differentially influence cohesin and accessory factor association.

The differences in the ratios of different cohesion factors also imply that cohesin dynamics differ at promoters, enhancers, and PREs. The SA to Rad21 cohesin subunit ratio is highest at enhancers, followed by PREs and promoters (Fig 6G). In contrast, the Wapl to SA ratio is virtually identical at all regulatory sequences (Fig 6H). If as proposed above, a high SA to cohesin ratio facilitates sister chromatid cohesion, this suggests that chromatid cohesion is higher at enhancers than at promoters.

The Pds5 to Wapl ratio is substantially higher at promoters than enhancers or PREs (Fig 6I) predicting that cohesin removal is more efficient at promoters. Indeed, Pds5 depletion substantially increases Rad21 at promoters, with more moderate increases at PREs and very modest increases at enhancers (Fig 7A). All increases are statistically significant (S6 Table). However, Pds5 depletion also significantly lowers SA levels at all regulatory sequences (Fig 7B, S6 Table). The combined effect of increasing Rad21 and lowering SA reduces and equalizes the SA to Rad21 ratio at all regulatory sequences (Fig 7C) agreeing with the meta-origin analysis (Fig 4B). Pds5 depletion also slightly increases the Wapl to SA ratio at all regulatory sequences (Fig 7D). Brca2 depletion significantly increases SA levels at all regulatory sequences (Fig 7E, S6 Table). Wapl depletion slightly increases SA at promoters and PREs, with a minor statistically significant decrease at enhancers (Fig 7F, S6 Table).

Pds5 depletion significantly increases Nipped-B at all regulatory sequences (Fig 7G, S6 Table) agreeing with the meta-origin analysis, and supporting the idea that Pds5 and Nipped-B compete for binding to cohesin. Wapl depletion substantially reduces Nipped-B at enhancer and PREs (Fig 7H) consistent with the effects at origins, but unexpectedly increases Nipped-B at promoters (Fig 7H, S6 Table). Wapl is low at promoters, so this suggests that regulatory sequences with higher Wapl levels may sequester Nipped-B, and that proteins at promoters,

potentially cohesin, or the TBPH RNA-binding protein [14] recruit Nipped-B released from enhancers and PREs by Wapl depletion.

The complex interplay between cohesin and the accessory proteins described above has implications for how cohesin facilitates enhancer-promoter communication. It is widely postulated that cohesin holds enhancers and promoters together by an intra-chromosomal cohesion mechanism similar to how it holds sister chromatids together. However, the different SA to Rad21 ratios and levels of cohesin loading and removal factors at enhancers and promoters is inconsistent with this idea, which predicts that the cohesin populations at enhancers and promoters should have very similar dynamics. We cannot rule out, however, the possibility that only a small fraction of cohesin at enhancers and promoters participates in enhancer-promoter cohesion.

Pds5, Brca2 and Wapl have similar effects on gene expression as Nipped-B and cohesin

RNA-seq experiments unexpectedly show that Pds5 and Brca2 have very similar roles in gene regulation, contrasting with their opposing functions in sister chromatid cohesion. We performed RNA-seq in BG3 cells depleted for Nipped-B, Rad21, Pds5, Wapl, and Brca2, and compared their genome-wide effects on RNA levels (Fig 8). Three biological replicates were used for each depletion, and were compared to six mock control replicates. The genome-wide RNA-seq expression data is provided in S7 Data.

The Pearson correlation between the changes in RNA levels caused by Nipped-B and Rad21 depletion is 0.63 (Fig 8A,B) consistent with prior studies showing that Nipped-B and cohesin have very similar effects on gene expression and transcription [11, 16]. The correlation between Pds5 and Brca2 depletion is 0.61, similar to the correlation between Nipped-B and Rad21 (Fig 8A,C). In contrast, the correlations between the effects of Nipped-B or Rad21 depletion versus Pds5 or Brca2 depletion are lower, ranging from 0.21 to 0.43 (Fig 8A,D). The correlations between Wapl depletion with Nipped-B, Rad21, Pds5 or Brca2 depletion range from 0.24 to 0.38 (Fig 8A).

The high correlation between the effects of Pds5 and Brca2 depletion on RNA levels is unexpected because Pds5 depletion reduces sister chromatid cohesion and Brca2 depletion does not. This indicates that most effects of Pds5 depletion on gene expression are not caused by sister chromatid cohesion defects. Combined with the differences in cohesin and accessory factor ratios at enhancers and promoters described above, this further argues that cohesin is unlikely to support enhancer-promoter looping by an intra-chromosomal cohesion mechanism. Pds5 depletion reduces SA levels on promoters, enhancers and PREs, while Brca2 depletion increases SA levels. Thus the changes in gene expression caused by Pds5 or Brca2 depletion also do not reflect changes in SA occupancy.

Although the Pearson correlations between the effects of Nipped-B or Rad21 depletion with those of Pds5, Brca2 and Wapl depletion are relatively low, close examination indicates

that the same genes are affected. Most genes affected by Nipped-B depletion are altered in the same direction by Pds5 depletion (Fig 8D). Agreeing with prior studies [11] on Nipped-B and cohesin, gene ontology analysis shows that genes involved in neurogenesis and imaginal disc development are increased by all depletions, and that genes involved in protein translation are reduced (S7 Data). Moreover, there is significant overlap in the genes that increase in expression ($p \leq 0.05$) with Nipped-B or Rad21 depletion and the genes that increase with Pds5, Wapl, or Brca2 depletion, as shown in Figure 8E. All overlaps in genes that increase with depletion of any of the cohesion factors are significant by Fisher's exact test (S7 Data) and few genes show changes in the opposite directions. There is also statistically significant overlap in the genes that decrease with Nipped-B or Rad21 depletion and genes that decrease with depletion of Pds5, Wapl, or Brca2, with only a few opposite effects, primarily with Wapl depletion (Fig 8E).

Although the effects of individually depleting Pds5 and Brca2 correlate modestly with Wapl depletion (0.38, 0.28) co-depletion of Pds5 and Brca2 strongly correlates with Wapl depletion (0.73) and Wapl-Brca2 co-depletion (0.68) (Fig 8A). The Pds5-Brca2 and Wapl-Brca2 double depletions show more modest correlations with Nipped-B or Rad21 depletion (0.12 – 0.33) or with depletion of Pds5 or Brca2 alone (0.20 –0.47) than with Wapl depletion. There is still significant overlap in the genes that increase in expression with Nipped-B or Rad21 depletion, but more genes that decrease in expression with Nipped-B or Rad21 depletion show increased expression in the double depletions (Fig 8E). Because Wapl depletion has virtually the same effects as Pds5-Brca2 or Wapl-Brca2 co-depletion, we conclude that loss of Wapl is epistatic to loss of Pds5 or Brca2, suggesting that the influence of Pds5 and Brca2 on gene expression requires Wapl. Wapl depletion more frequently has an opposite effect on those genes that decrease in expression upon Nipped-B or cohesin depletion, suggesting that Wapl counteracts activation by Nipped-B and cohesin at some genes.

Brca2 influences gene expression in developing wings

The finding that the Brca2 DNA repair protein influences gene expression in BG3 cells through functional interactions with sister chromatid cohesion proteins raised the question of whether or not it also influences gene expression during in vivo development. We conducted RNA-seq in 3rd instar wing imaginal discs from different control genetic backgrounds, and two different *brca2* null mutants, which revealed that many genes increase and decrease in expression. Using a statistical threshold of $q \leq 0.05$, 208 genes increase in expression and 606 decrease, with most changes less than 2-fold (Fig 9A, S7 Data). Gene ontology analysis indicates that the increasing genes are enriched for those involved in mitotic spindle formation, and decreasing genes are involved in morphogenesis and development (S7 Data). In additional to many genes that regulate development, the genes that decrease in expression include broadly-acting transcription factors such as Mediator subunits and Pol II kinases, cell cycle control genes, and genes encoding DNA repair proteins (Fig 9B). Unlike *pds5* or *wapl* mutant flies [5, 6] *brca2* null mutant flies are viable, although females are sterile because of defective meiosis [25].

There are no overt structural mutant phenotypes in *brca2* null adult flies, but it seems likely that the many modest changes in gene expression could cause subtle growth or developmental deficits that might be revealed by close examination, or in different genetic backgrounds. Towards this end we measured the sizes of adult male and female *brca2* mutant wings, to look for possible changes in growth. Male *brca2* mutant wings are some 9% smaller than controls, and female wings are 4% smaller (S8 Fig). These reductions are similar in magnitude to the dominant effects of null mutations in the fly *myc* (*diminutive*, *dm*) and *Tor* genes encoding critical dosage-sensitive growth regulators [26]. They are also similar to the effects of heterozygous *Nipped-B* mutations [26]. We thus posit that Brca2 facilitates growth through effects on gene expression.

Discussion

The studies presented here show that the Pds5-Wapl complex limits the size of the cohesin binding domains centered around DNA replication origins, and that Pds5 and Brca2, which form a complex lacking Wapl [8], have opposing effects on sister chromatid cohesion and binding of the SA cohesin subunit to replication origins. In contrast to their opposing roles in sister cohesion, Pds5 and Brca2 have very similar roles in controlling gene expression, regulating the same genes as Nipped-B and cohesin. As outlined below, these findings have significant implications for where cohesin binds, and cohesin's roles in sister chromatid cohesion and gene regulation. Some of the key ideas are outlined in Figure 10. Our observations make predictions about mutations that could cause human developmental syndromes, and have implications for how *BRCA2* mutations increase cancer susceptibility.

Cohesin localization

The finding that maximal levels of cohesin and cohesin regulatory factors center at replication origins and decrease extending outward for many kilobases implies that DNA replication plays a critical role in positioning cohesin. This is consistent with single molecule studies showing that DNA replication causes cohesin to translocate uni-directionally [17]. The single-molecule studies also show that cohesin translocation is ATPase-dependent and suppressed by Pds5 and Wapl [17]. This agrees with our finding that depletion of Pds5 or Wapl increases cohesin levels for tens of kilobases surrounding replication origins. We thus favor the idea that pushing of cohesin by replication forks is a key determinant of cohesin localization and which genes bind cohesin (Fig 10). We further posit that reduction of Pds5 or Wapl slows the rate of cohesin removal in front of replication forks, leading to increased cohesin domain size.

In Xenopus oocyte extracts, association of the Nipped-B orthologs with chromatin and cohesin loading require replication origin licensing and formation of the complete pre-replication complex (pre-RC) containing the origin recognition complex (ORC) the minichromosome maintenance (MCM) helicase complex, Cdc6 (cell division cycle 6) and Cdt1 (cyclin-dependent transcript 1) [27-30]. This supports the idea that replication origins are major loading sites for cohesin. Studies in Drosophila cultured cells, however, indicate that origin licensing is not essential for stable association of cohesin with chromatin [13]. Xenopus oocytes do not have

active gene transcription, and we thus think it likely that a combination of both licensed replication origins and active genes determine cohesin binding in Drosophila cells (Fig 10).

We envision that DNA replication exerts global control over the extent of cohesin domains, and that within these domains, proteins binding to active genes and regulatory sequences define the detailed cohesin localization and dynamics. Once an active gene or regulatory sequence binds cohesin and accessory factors, cohesin is then continuously loaded and removed from these sites during interphase, as shown by in vivo FRAP experiments [7]. The dynamics at different sites is likely fine-tuned by the ratios of Nipped-B, Pds5, Wapl and Brca2, which as shown here, vary between promoters, enhancers and PREs. A key question this idea raises, however, is how cohesin is positioned on chromosomes in early G1 after mitosis and before DNA replication. We speculate that unknown book-marking factors remain bound to cohesin-binding genes and regulatory elements through mitosis and determine cohesin re-loading in early G1 together with the proteins binding active genes and regulatory elements. The cohesin and accessory protein levels for each gene and regulatory sequence are then refreshed during DNA replication as necessitated by duplication of the genome. This model predicts that the hypothetical book-marking factors may also be extended to additional active genes upon Pds5 or Wapl depletion as the cohesin domains increase in size.

A potential role for DNA replication in cohesin positioning in mammalian cells remains to be investigated. Recent studies show that transcription, the CTCF architectural protein, and Wapl position cohesin in mammalian cells [31]. Absence of CTCF causes cohesin to localize to transcription start sites, similar to the pattern observed in Drosophila. Unlike mammalian CTCF, fly CTCF does not interact directly with the SA cohesin subunit [32] and cohesin does not significantly co-localize with CTCF [11, 15]. Thus the pattern observed in CTCF-deficient mammalian cells is similar to the normal pattern in Drosophila. In mammalian cells that lack both CTCF and Wapl, cohesin accumulates in large islands several kilobases in size at the 3' ends of genes [31] suggesting that transcription pushes cohesin to the ends of genes, as seen in yeast [33]. Although we detect modest increases in cohesin in some intergenic regions upon Pds5 or Wapl depletion in Drosophila cells, we do not see new substantial intergenic domains, suggesting that there may be additional factors in mammalian cells that trap cohesin at the 3' ends of genes. Although it remains to be investigated, we envision that DNA replication also positions cohesin in mammalian cells, and that CTCF and transcription refine cohesin positioning, similar to our finding that the cohesin extending out from Drosophila replication origins skips over inactive genes to associate with active genes and regulatory sequences.

Sister chromatid cohesion

Why Pds5 is required for sister chromatid cohesion is poorly understood [33]. A recent study in budding yeast links its role in cohesion to DNA replication by showing that mutations affecting the Elg1 protein that unloads the PCNA replication clamp, or alternatively overexpression of PCNA, suppress cohesion defects of *pds5* mutants [35]. Here we find that Pds5 is required for a high SA to Rad21 cohesin subunit ratio at replication origins, and that Brca2, which opposes the function of Pds5 in cohesion, has a negative effect on SA levels at origins. Core cohesin

subunits, including Smc1, Smc3 and Rad21 interact with themselves in a SA-dependent manner, leading to a handcuff model for sister chromatid cohesion in which SA links two cohesin rings via interaction with Rad21 [22]. In other cohesion models, such as the embrace model in which one cohesin ring encircles both sisters, SA may inhibit random opening of the cohesin ring. Cohesin rings lacking SA and Pds5 maintain topological association with DNA in the absence of sister chromatid cohesion [36] so it is feasible that the ratio of SA to cohesin at the origin determines the fraction of cohesin rings that participate in sister cohesion.

The Pds5-Brca2 complex is important for mitotic DNA repair and homologous recombination during meiosis, corresponding with the known role of Brca2 in DNA repair [8-10]. The finding that Brca2 can oppose Pds5 function in sister chromatid cohesion shows that Brca2 plays an additional role in regulating genome stability even in the absence of DNA damage. Although increased SA at origins is apparent with Brca2 depletion alone, *brca2* null mutant flies are viable, and the anti-cohesion effects of Brca2 are not apparent unless Pds5 is reduced. The anti-cohesion role of Brca2 is also opposite to its role in DNA repair in the sense that reduced cohesion decreases genome stability and DNA repair increases stability. It remains to be seen if hypermorphic *brca2* mutations or overexpression of Brca2 can reduce sister chromatid cohesion sufficiently to cause chromosome missegregation, but this is potentially relevant to the increased cancer susceptibility in some individuals with *BRCA2* missense mutations. It is also of interest that the human *BRCA2* gene neighbors *PDS5B*, with some cancer-associated mutations likely altering both genes [37]. This raises the possibility that some mutations could alter the BRCA2-PDS5B ratio in a way that disfavors sister chromatid cohesion.

Gene regulation

Since the original discovery that *Nipped-B* mutations alter enhancer-dependent gene expression [38] an attractive concept has been that cohesin holds enhancers and promoters together by a topological mechanism similar to how it holds sister chromatids together. This idea has been expanded in loop extrusion models in which chromatin is threaded through topologically bound cohesin to form intra-chromosomal loops important for topological domain formation and gene regulation [39]. These models rely on the handcuff or embrace models to hold together two regions of the same chromosome.

Not all TADs in Drosophila cells, however, require cohesin for their formation [17] and data presented here argue against the intra-chromosomal cohesion model for facilitating enhancer-promoter loops. The Pds5 to Wapl ratio is substantially higher at promoters than at enhancers, which predicts that cohesin binding is less stable at promoters (Fig 10). Indeed, Pds5 depletion substantially increases cohesin levels at promoters but has much less effect at enhancers. The intra-chromosomal cohesion model predicts that the cohesin dynamics will be same at enhancers and promoters, either because they are the same molecules at both sequences in the embrace model, or tightly linked to each other in the handcuff model. The caveat to this argument is that if there are several cohesin rings at these regulatory sequences, as few as one or two may mediate intra-chromosomal cohesion, while the majority differ in their binding dynamics.

The finding that Pds5 and Brca2 depletion have nearly identical genome-wide effects on gene expression, however, provides further evidence against intra-chromosomal cohesion because Pds5 depletion strongly reduces sister cohesion and Brca2 depletion does not. It could be that only a small fraction of the effects on gene expression involve looping deficits, and that changes in cohesin's other roles in gene regulation are responsible for most effects on gene expression. For instance, cohesin directly influences transcription of active genes by recruiting the PRC1 Polycomb repressive complex to the promoter region, which also indirectly controls gene silencing by regulating PRC1 availability [14, 24]. Nipped-B and cohesin bind to essentially all enhancers, however, and there are many instances where Nipped-B or cohesin depletion reduces transcription of genes activated by known enhancers [16]. We thus currently prefer the idea that interactions between Nipped-B or cohesin with other proteins, such as the Mediator complex [40] facilitate enhancer-promoter looping.

The finding that Pds5 and Brca2 influence expression of the same genes as Nipped-B and cohesin supports the idea that they mediate their effects on gene expression in large part through their effects on cohesin binding and dynamics. It is counterintuitive, however, that their effects on gene expression are in the same direction as Nipped-B or cohesin because Pds5 depletion increases cohesin at promoters and Nipped-B at all regulatory sequences. It is also unexpected that Pds5 and Brca2 depletion, which have opposite effects on sister chromatid cohesion and SA levels at gene regulatory sequences, have the same effect on gene expression. These findings thus argue that cohesin dynamics are more important than absolute cohesin levels for gene expression. Another possibility to be considered is that the level of Pds5, and not Nipped-B or cohesin, is the most crucial for gene regulation. Nipped-B and Brca2 depletion reduce Pds5 at all regulatory sequences except PREs, and thus have similar effects on Pds5 binding levels as depletion of Pds5.

The finding that Brca2 has significant effects on gene expression in both cultured cells and in vivo has implications for the role of *BRCA2* mutations in cancer susceptibility. Many of the genes whose expression is altered in *brca2* mutant wing discs are involved in development, cell cycle control and DNA repair, or are broadly active transcriptional regulators. Thus effects on gene expression caused by *BRCA2* mutations could contribute to cancer etiology even in the absence of DNA damage.

Can *PDS5A*, *PDS5B*, *WAPL* and *BRCA2* mutations cause cohesinopathy-like syndromes?

Heterozygous loss-of-function mutations in *NIPBL*, the human Nipped-B ortholog cause Cornelia de Lange syndrome (CdLS) which displays significant deficits in physical and intellectual growth and development, and structural abnormalities in the face, limbs and organs [41, 42]. Typically milder forms of CdLS are caused by dominant missense mutations in the *SMC1A* and *SMC3* cohesin subunit genes [43, 44]. Dominant loss of function mutations in *HDAC8* encoding the deacetylase that recycles acetylated SMC3 cause CdLS similar in severity to *NIPBL* mutations [45] and deficiencies in *RAD21* cause a developmental syndrome with

developmental deficits overlapping CdLS [46]. The individuals with these cohesinopathy mutations do not display overt sister chromatid cohesion or chromosome segregation phenotypes, and thus the leading idea is that the developmental deficits arise from changes in gene expression [47, 48]. The finding that Pds5, Brca2 and Wapl depletion have similar effects on gene expression as Nipped-B and cohesin in BG3 cells, and that *brca2* null mutant wing discs show significant changes in the expression of hundreds of genes that influence growth and development raises the question of whether the human orthologs could also be cohesinopathy genes.

There are two Pds5 orthologs in mammals, and mice homozygous mutant for *Pds5A* or *Pds5B* show severe lethal developmental phenotypes [49, 50] overlapping but not identical to those caused by heterozygous *Nipbl* loss-of-function mutations [51]. The individual *Pds5A* and *Pds5B* heterozygotes do not display overt phenotypes. It can be predicted, therefore, that similar mutations in humans would not cause a dominant syndrome, and would be recessive lethal.

Drosophila *wapl* hemizygous null mutations are lethal, but there are no overt adult phenotypes in heterozygous females, despite a measurable effect on cohesin chromosome-binding dynamics [7]. Strikingly, a dominant negative *wapl* mutant allele that produces a truncated Wapl protein stabilizes and increases cohesin binding and causes a Polycomb mutant phenotype reflecting decreased epigenetic silencing of homeotic genes [52]. *Nipped-B* and cohesin mutations dominantly suppress the dominant transformation phenotypes of *Pc* mutants [23, 53, 54] and thus it is possible that a *WAPL* dominant-negative mutation could cause a human developmental syndrome that would differ significantly from CdLS and other cohesinopathies, and which may be lethal in utero.

Although fly *brca2* null mutants are viable [25] mouse *Brca2* mutations are homozygous lethal during embryogenesis [55] and thus are also likely early recessive lethal in humans. Interestingly, some mice with *Brca2* hypomorphic mutations survive to adulthood, and are smaller than littermates, with multiple developmental abnormalities [56]. The sources of these developmental deficits are unknown, but our results raise the likelihood that they reflect multiple changes in gene expression. The mouse studies also raise the possibility that *BRCA2* hypomorphic mutations could cause a recessive developmental syndrome in humans.

Materials and methods

Cell culture, RNAi treatment and metaphase spreads

BG3 cells were cultured and RNAi treatment and metaphase spreads were performed as previously described [11].

Pds5 and Brca2 antibodies

Full length Pds5 (residues 1 – 1218) was expressed as a His6 fusion in *E. coli*, purified by nickel chromatography, and used to immunize a guinea pig at Pocono Rabbit Farm and Laboratory, Inc. Brca2 residues 1 to 404 of were expressed as a His6 fusion peptide in *E. coli*, purified and used to immunize a rabbit and a guinea pig at Josman, LLC.

ChIP-seq

ChIP-seq was performed as described before [12, 57]. The Nipped-B, Rad21, SA and Wapl antibodies and their validation were described previously [15, 52]. The Pds5 and Brca2 antibodies were made for this study as described above and validated by RNAi westerns (S1 Fig). At least two independent biological replicate experiments were performed for each experimental ChIP-seq group, sequencing to at least 10x genome coverage per replicate, and normalizing to input chromatin sequencing greater than 45x genome coverage. Mean enrichment at active promoters, enhancers and PREs was calculated as described before [12, 14] using custom R [58] scripts. Meta-origin analysis was performed using custom R scripts and the locations of early DNA replication origins in BG3 cells [13] (S2 File). Meta-origin plots were made using Microsoft Excel. Violin plots were made using the wvioplot.R package (https://github.com/cran/wvioplot). The Integrated Genome Browser (IGB) [59] was used to inspect ChIP-seq data and prepare figures.

DNA fiber analysis

DNA fibers were spread on microscope slides by adapting a published procedure [60]. Mock or RNAi-depleted BG3 cells were labeled with 20 micromolar IdU (ThermoFisher, cat no. 10035701) in the media for 20 min at 25°, washed twice with phosphate buffered saline (PBS) and then labeled for 20 min with 200 micromolar CldU in the culture media (Sigma Aldrich, cat no. C6891) at 25°. Labeled cells were washed twice with PBS, trypsinized, washed twice with PBS and brought to 7,000 cells per microliter in PBS and placed on ice. For experiments with hydroxyurea (HU) treatment, cells were washed twice with PBS after CldU incorporation and incubated at 25° with 4 mM HU for 4 hours prior to cell collection. Labeled cells were diluted with unlabeled control BG3 cells in a 1:5 or 1:10 ratio before spreading to decrease the number of overlapping labeled fibers [60]. Two microliters of diluted cells were mixed with eight microliters of lysis buffer (200 mM Tris·HCl pH 7.5, 50 mM EDTA, 0.5% SDS) on top of a microscope slide (VWR, cat no. 48311-703) for eight min. After lysis, the slides were tilted at a 20 to 40° angle to allow the DNA to spread. Slides were air-dried at room temperature for 1 hour and fixed in 3:1 v/v methanol:acetic acid solution for eight min. Fixed slides were dried and stored at 4° before immunostaining.

For immunostaining, fixed slides were washed twice for five min with PBS and treated with 2.5 N HCl for one hour. Acid-treated slides were washed three times for five min with PBS, blocked with 5% (w/v) bovine serum albumin (BSA) in PBS for one hour at 37° and rinsed in PBST (0.05% Tween 20 in PBS) for five to ten sec just prior to adding primary antibodies diluted in PBST containing 1% BSA. The primary antibodies used were 1:20 mouse anti-BrdU B44 (BD

Biosciences, cat no. 347580) which recognizes IdU and 1:100 rat anti-BrdU monoclonal antibody BU1/75 (ThermoFisher, cat no. MA1-82088) which recognizes CldU. Forty microliters of the primary antibody mix was added to each slide, covered with a coverslip, and the slides were incubated in a dark humid chamber for 1.5 hours at room temperature. Coverslips were removed in PBS, and the slides washed three times for five min with PBST and kept in PBS until addition of the secondary antibodies. The secondary antibodies were each diluted 1:125 in PBST containing 1% BSA. The secondary antibodies used were Alexa Fluor 546 goat anti-mouse IgG1 (Invitrogen, cat no. A21123) and Alexa Fluor 488 chicken anti-rat IgG (Invitrogen, cat no. A21470). Forty microliters of diluted secondary antibody mix was added to each slide, covered with a coverslip, and the slides were incubated for one hour in a dark humid chamber at room temperature. Coverslips were removed in PBS, and the slides washed three times for five min in PBST and left in PBS before mounting. For mounting, slides were air-dried at room temperature in the dark, and mounted with 20 microliters of ProLong Gold antifade reagent (Invitrogen, cat no. P36930) under a coverslip. Slides were dried at room temperature in the dark and stored at 4° before imaging. Fluorescent micrographs were digitally captured using a Leica SP5 laser scanning confocal microscope with a 60x objective, and the lengths of connected red and green fibers were measured using NIH ImageJ software. Statistical analysis was conducted using R and violin plots generated using wvioplot.R.

RNA-seq

Total RNA-seq seq with BG3 cells was performed using ribosomal depletion and processed as described previously [26] using three independent biological replicates for each depletion and six previously reported mock control samples [14]. Gene expression data in mean normalized nucleotide coverage per gene, and gene ontology analysis is provided in S7 Data. Total RNA-seq with dissected late third instar wing discs from female larvae was conducted as described before using ribosomal depletion [26]. To minimize genetic background effects, five wild-type controls consisted of two Canton S samples and three *y w* samples from independent stocks. The five *brca2 -/-* replicates consisted of two isolations of GFP negative larvae from a $brca2^{KO}$ / *CyO, Kr-GFP* stock and three isolations from GFP negative larvae from a $brca2^{56e}$ / *CyO, Kr-GFP* stock. Both *brca2* null alleles were the gift of Trudi Schüpbach. The expression data in mean normalized nucleotide coverage per gene, and gene ontology analysis is provided in S7 Data.

Wing area measurements

Wings were collected from adult flies from crosses conducted at 25°, mounted, photographed and measured as previously described [26].

Acknowledgements

We thank Judy Kassis for providing Wapl antibody, Trudi Schüpbach for the $brca2^{KO}$ and $brca2^{56e}$ alleles, David MacAlpine for the processed BG3 early DNA replication data, members of the Alessandro Vindigni laboratory for advice on DNA fiber analysis, and Michael Rauchman

for comments on the manuscript. This work was supported by grants from the National Institutes of Health (R01 GM108714) and the Alvin J Siteman Cancer Center, Washington University School of Medicine (8074-88) to DD.

References

1. Gligoris T, Löwe J. Structural insights into ring formation of cohesin and related Smc complexes. Trends Cell Biol. 2016;26: 680-93. doi: 10.1016/j.tcb.2016.04.002.

2. Ouyang Z, Yu H. Releasing the cohesin ring: A rigid scaffold model for opening the DNA exit gate by Pds5 and Wapl. Bioessays. 2017;39. doi: 10.1002/bies.201600207.

3. Rankin S, Dawson DS. Recent advances in cohesin biology. F1000Res. 2016;5. pii: F1000 Faculty Rev-1909. doi: 10.12688/f1000research.8881.1.

4. Uhlmann F. SMC complexes: from DNA to chromosomes. Nat Rev Mol Cell Biol. 2016;17: 399-412. doi: 10.1038/nrm.2016.30.

5. Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, McKim K. Effects of sister chromatid cohesion proteins on *cut* gene expression during wing development in Drosophila. Development. 2005;132: 4743-53. doi: 10.1242/dev.02064.

6. Vernì F, Gandhi R, Goldberg ML, Gatti M. Genetic and molecular analysis of *wings apart-like* (*wapl*), a gene controlling heterochromatin organization in *Drosophila melanogaster*. Genetics. 2000;154: 1693-710.

7. Gause M, Misulovin Z, Bilyeu A, Dorsett D. Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5, and Wapl. Mol Cell Biol. 2010;30: 4940-51. doi: 10.1128/MCB.00642-10.

8. Kusch T. Brca2-Pds5 complexes mobilize persistent meiotic recombination sites to the nuclear envelope. J Cell Sci. 2015;128: 717-27. doi: 10.1242/jcs.159988.

9. Brough R, Bajrami I, Vatcheva R, Natrajan R, Reis-Filho JS, Lord CJ, Ashworth A. APRIN is a cell cycle specific BRCA2-interacting protein required for genome integrity and a predictor of outcome after chemotherapy in breast cancer. EMBO J. 2012;31: 1160-76. doi: 10.1038/emboj.2011.490.

10. Couturier AM, Fleury H, Patenaude AM, Bentley VL, Rodrigue A, Coulombe Y, et al. Roles for APRIN (PDS5B) in homologous recombination and in ovarian cancer prediction. Nucleic Acids Res. 2016;44: 10879-97. doi: 10.1093/nar/gkw921.

11. Schaaf CA, Misulovin Z, Sahota G, Siddiqui AM, Schwartz YB, Kahn TG, et al. Regulation of the Drosophila Enhancer of split and *invected-engrailed* gene complexes by sister chromatid cohesion proteins. PLoS One 2009;4: e6202. doi: 10.1371/journal.pone.0006202.

12. Swain A, Misulovin Z, Pherson M, Gause M, Mihindukulasuriya K, Rickels RA, et al. Drosophila TDP-43 RNA-Binding protein facilitates association of sister chromatid cohesion proteins with genes, enhancers and Polycomb Response Elements. PLoS Genet. 2016;12:e1006331. doi: 10.1371/journal.pgen.1006331.

13. MacAlpine HK, Gordân R, Powell SK, Hartemink AJ, MacAlpine DM. Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. Genome Res. 2010;20: 201-11. doi: 10.1101/gr.097873.109.

14. Pherson M, Misulovin Z, Gause M, Mihindukulasuriya K, Swain A, Dorsett D. Polycomb Repressive Complex 1 modifies transcription of active genes. Sci Adv. 2017; in press.

15. Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, MacArthur S, et al. Association of cohesin and Nipped-B with transcriptionally active regions of the *Drosophila melanogaster* genome. Chromosoma. 2008;117: 89-102. doi: 10.1007/s00412-007-0129-1.

16. Schaaf CA, Kwak H, Koenig A, Misulovin Z, Gohara DW, Watson A, et al. Genome-wide control of RNA polymerase II activity by cohesin. PLoS Genet. 2013;9: e1003382. doi: 10.1371/journal.pgen.1003382.

17. Schaaf CA, Misulovin Z, Gause M, Koenig A, Dorsett D. The Drosophila enhancer of split gene complex: architecture and coordinate regulation by notch, cohesin, and polycomb group proteins. G3 (Bethesda). 2013;3: 1785-94. doi: 10.1534/g3.113.007534.

18. Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, et al. The cohesin release factor WAPL restricts chromatin loop extension. Cell. 2017;169: 693-707.e14. doi: 10.1016/j.cell.2017.04.013.

19. Kanke M, Tahara E, Huis In't Veld PJ, Nishiyama T. Cohesin acetylation and Wapl-Pds5 oppositely regulate translocation of cohesin along DNA. EMBO J. 2016;35: 2686-2698. doi: 10.15252/embj.201695756.

20. Schlacher K, Christ N, Siaud N, Egashira A, Wu H, Jasin M. Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. Cell. 2011;145: 529-42. doi: 10.1016/j.cell.2011.03.041.

21. Zhang N, Kuznetsov SG, Sharan SK, Li K, Rao PH, Pati D. A handcuff model for the cohesin complex. J Cell Biol. 2008;183: 1019-31. doi: 10.1083/jcb.200801157.

22. Zhang N, Pati D. Handcuff for sisters: a new model for sister chromatid cohesion. Cell Cycle. 2009;8: 399-402. doi: 10.4161/cc.8.3.7586.

23. Kikuchi S, Borek DM, Otwinowski Z, Tomchick DR, Yu H. Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy. Proc Natl Acad Sci U S A. 2016;113: 12444-9. doi: 10.1073/pnas.1611333113.

24. Schaaf CA, Misulovin Z, Gause M, Koenig A, Gohara DW, Watson A, Dorsett D. 2013. Cohesin and polycomb proteins functionally interact to control transcription at silenced and active genes. PLoS Genet. 2013;9: e1003560. doi: 10.1371/journal.pgen.1003560.

25. Klovstad M, Abdu U, Schüpbach T. Drosophila brca2 is required for mitotic and meiotic DNA repair and efficient activation of the meiotic recombination checkpoint. PLoS Genet. 2008;4: e31. doi: 10.1371/journal.pgen.0040031.

26. Wu Y, Gause M, Xu D, Misulovin Z, Schaaf CA, Mosarla RC, et al. Drosophila *Nipped-B* mutants model Cornelia de Lange syndrome in growth and behavior. PLoS Genet. 2015;11: e1005655. doi: 10.1371/journal.pgen.1005655.

27. Gillespie PJ, Hirano T. Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. Curr Biol. 2004;14: 1598-603. doi: 10.1016/j.cub.2004.07.053.

28. Takahashi TS, Yiu P, Chou MF, Gygi S, Walter JC. Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex. Nat Cell Biol. 2004;6: 991-6. doi: 10.1038/ncb1177.

29. Bermudez VP, Farina A, Higashi TL, Du F, Tappin I, Takahashi TS, Hurwitz J. In vitro loading of human cohesin on DNA by the human Scc2-Scc4 loader complex. Proc Natl Acad Sci U S A. 2012;109: 9366-71. doi: 10.1073/pnas.1206840109.

30. Takahashi TS, Basu A, Bermudez V, Hurwitz J, Walter JC. Cdc7-Drf1 kinase links chromosome cohesion to the initiation of DNA replication in Xenopus egg extracts. Genes Dev. 2008;22: 1894-905. doi: 10.1101/gad.1683308.

31. Busslinger GA, Stocsits RR, van der Lelij P, Axelsson E, Tedeschi A, Galjart N, Peters JM. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature. 2017;544: 503-7. doi: 10.1038/nature22063.

32. Xiao T, Wallace J, Felsenfeld G. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. Mol Cell Biol. 2011;31: 2174-83. doi: 10.1128/MCB.05093-11.

33. Ocampo-Hafalla M, Muñoz S, Samora CP, Uhlmann F. Evidence for cohesin sliding along budding yeast chromosomes. Open Biol. 2016;6. pii: 150178. doi: 10.1098/rsob.150178.

34. Skibbens RV. Establishment of sister chromatid cohesion. Curr Biol. 2009;19: R1126-32. doi: 10.1016/j.cub.2009.10.067.

35. Tong K, Skibbens RV. Pds5 regulators segregate cohesion and condensation pathways in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A. 2015;112: 7021-6. doi: 10.1073/pnas.1501369112.

36. Kulemzina I, Schumacher MR, Verma V, Reiter J, Metzler J, Failla AV, et al. Cohesin rings devoid of Scc3 and Pds5 maintain their stable association with the DNA. PLoS Genet. 2012;8: e1002856. doi: 10.1371/journal.pgen.

37. Geck P, Sonnenschein C, Soto AM. The D13S171 marker, misannotated to BRCA2, links the AS3 gene to various cancers. Am J Hum Genet. 2001;69: 461-3. doi: 10.1086/321968.

38. Rollins RA, Morcillo P, Dorsett D. Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the *cut* and *Ultrabithorax* genes. Genetics. 1999;152: 577-93.

39. Barrington C, Finn R, Hadjur S. Cohesin biology meets the loop extrusion model. Chromosome Res. 2017;25: 51-60. doi: 10.1007/s10577-017-9550-3.

40. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, et al. Mediator and cohesin connect gene expression and chromatin architecture. Nature. 2010 23;467: 430-5. doi: 10.1038/nature09380.

41. Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, et al. Cornelia de Lange syndrome is caused by mutations in *NIPBL*, the human homolog of *Drosophila melanogaster Nipped-B*. Nat Genet. 2004;36: 631-5. doi: 10.1038/ng1364.

42. Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. *NIPBL*, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet. 2004;36: 636-41. doi: 10.1038/ng1363.

43. Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, et al. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of Cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet. 2007;80: 485-94. doi: 10.1086/511888.

44. Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, et al. X-linked Cornelia de Lange syndrome owing to *SMC1L1* mutations. Nat Genet. 2006;38: 528-30. doi: 10.1038/ng1779.

45. Deardorff MA, Bando M, Nakato R, Watrin E, Itoh T, Minamino M, et al. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. Nature. 2012;489: 313-7. doi: 10.1038/nature11316.

46. Deardorff MA, Wilde JJ, Albrecht M, Dickinson E, Tennstedt S, Braunholz D, et al. RAD21 mutations cause a human cohesinopathy. Am J Hum Genet. 2012;90: 1014-27. doi: 10.1016/j.ajhg.2012.04.019.

47. Dorsett D, Krantz ID. On the molecular etiology of Cornelia de Lange syndrome. Ann N Y Acad Sci. 2009;1151: 22-37. doi: 10.1111/j.1749-6632.2008.03450.x.

48. Liu J, Krantz ID. Cornelia de Lange syndrome, cohesin, and beyond. Clin Genet. 2009;76:303-14. doi: 10.1111/j.1399-0004.2009.01271.x.

49. Zhang B, Jain S, Song H, Fu M, Heuckeroth RO, Erlich JM, et al. Mice lacking sister chromatid cohesion protein PDS5B exhibit developmental abnormalities reminiscent of Cornelia de Lange syndrome. Development. 2007;134: 3191-201. doi: 10.1242/dev.005884.

50. Zhang B, Chang J, Fu M, Huang J, Kashyap R, Salavaggione E, et al. Dosage effects of cohesin regulatory factor PDS5 on mammalian development: implications for cohesinopathies. PLoS One. 2009;4: e5232. doi: 10.1371/journal.pone.0005232.

51. Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, et al. Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome. PLoS Genet. 2009;5: e1000650. doi: 10.1371/journal.pgen.1000650.

52. Cunningham MD, Gause M, Cheng Y, Noyes A, Dorsett D, Kennison JA, Kassis JA. Wapl antagonizes cohesin binding and promotes Polycomb-group silencing in Drosophila. Development. 2012;139: 4172-9. doi: 10.1242/dev.084566.

53. Kennison JA, Tamkun JW. Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila. Proc Natl Acad Sci U S A. 1988;85: 8136-40.

54. Hallson G, Syrzycka M, Beck SA, Kennison JA, Dorsett D, Page SL, et al. The Drosophila cohesin subunit Rad21 is a trithorax group (trxG) protein. Proc Natl Acad Sci U S A. 2008;105: 12405-10. doi: 10.1073/pnas.0801698105.

55. Hakem R, de la Pompa JL, Mak TW. Developmental studies of *Brca1* and *Brca2* knock-out mice. J Mammary Gland Biol Neoplasia. 1998;3: 431-45.

56. Connor F, Bertwistle D, Mee PJ, Ross GM, Swift S, Grigorieva E, et al. Tumorigenesis and a DNA repair defect in mice with a truncating Brca2 mutation. Nat Genet. 1997;17: 423-30. doi: 10.1038/ng1297-423

57. Dorsett D, Misulovin Z. Measuring sister chromatid cohesion protein genome occupancy in *Drosophila melanogaster* by ChIP-seq. Methods Mol Biol. 2017;1515: 125-139. doi: 10.1007/978-1-4939-6545-8_8.

58. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

59. Freese NH, Norris DC, Loraine AE. Integrated genome browser: visual analytics platform for genomics. Bioinformatics. 2016;32: 2089-95. doi: 10.1093/bioinformatics/btw069.

60. Jackson DA, Pombo A. Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol. 1998;140: 1285-95.

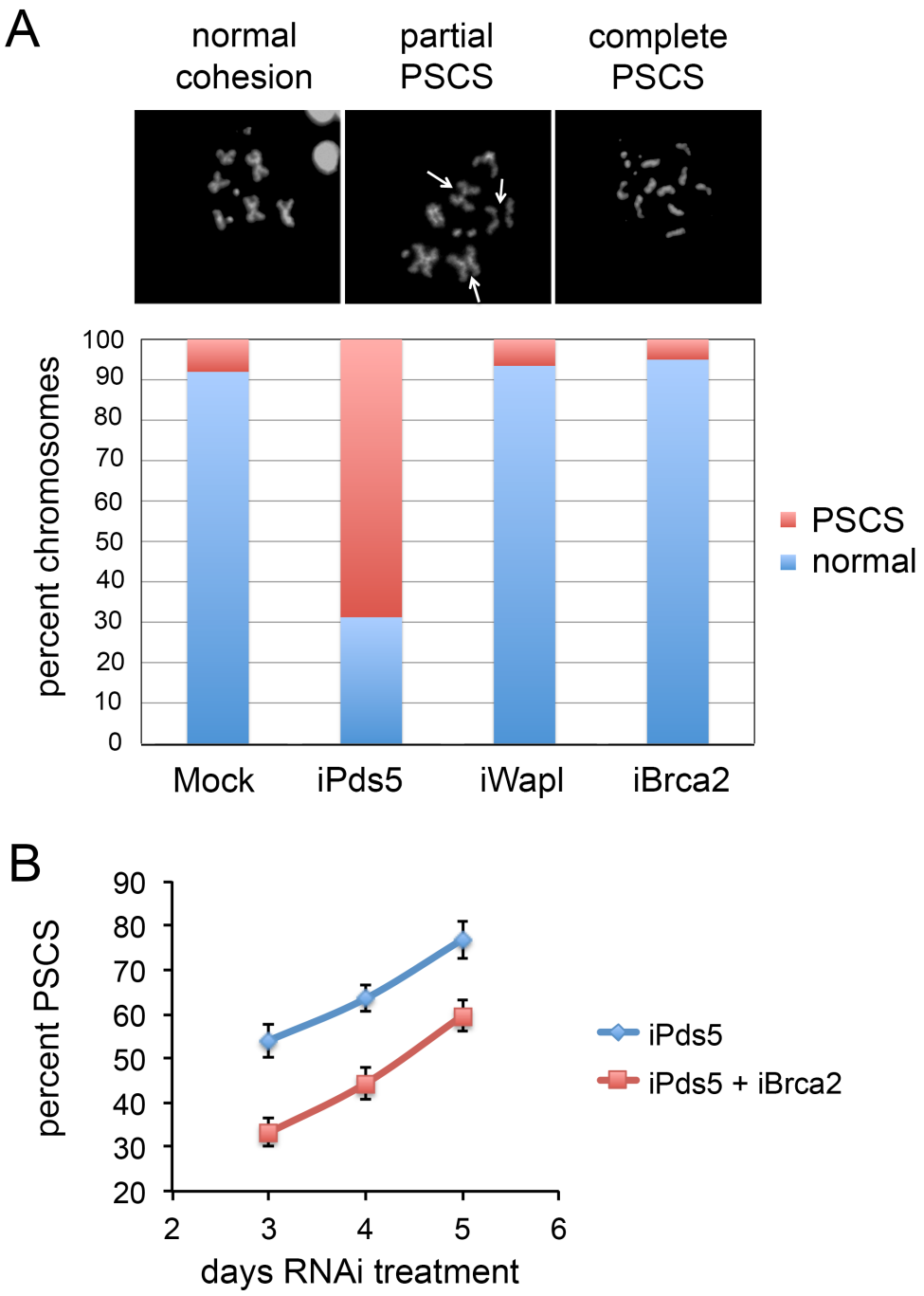


Fig 1. Pds5 and Brca2 oppose each other in sister chromatid cohesion in BG3 cells.
(**A**) The three micrographs show examples of normal metaphase chromosomes and of precocious sister chromatid cohesion (PSCS). The bar graph shows the percent of chromosomes showing normal cohesion (blue) or partial or complete PSCS (red) in mock-treated (Mock), and Pds5-depleted (iPds5), Wapl-depleted (iWapl) or Brca2-depleted (iBrca2) BG3 cells after 5 days of RNAi treatment. (**B**) The graphs show the mean percentage of chromosomes showing PSCS in individual cells after three to five days of RNAi treatment for Pds5 only (iPds5, blue) or for both Pds5 and Brca2 (iPds5 + iBrca2, red). Error bars are standard errors of the mean. Similar results were obtained in two additional experiments.

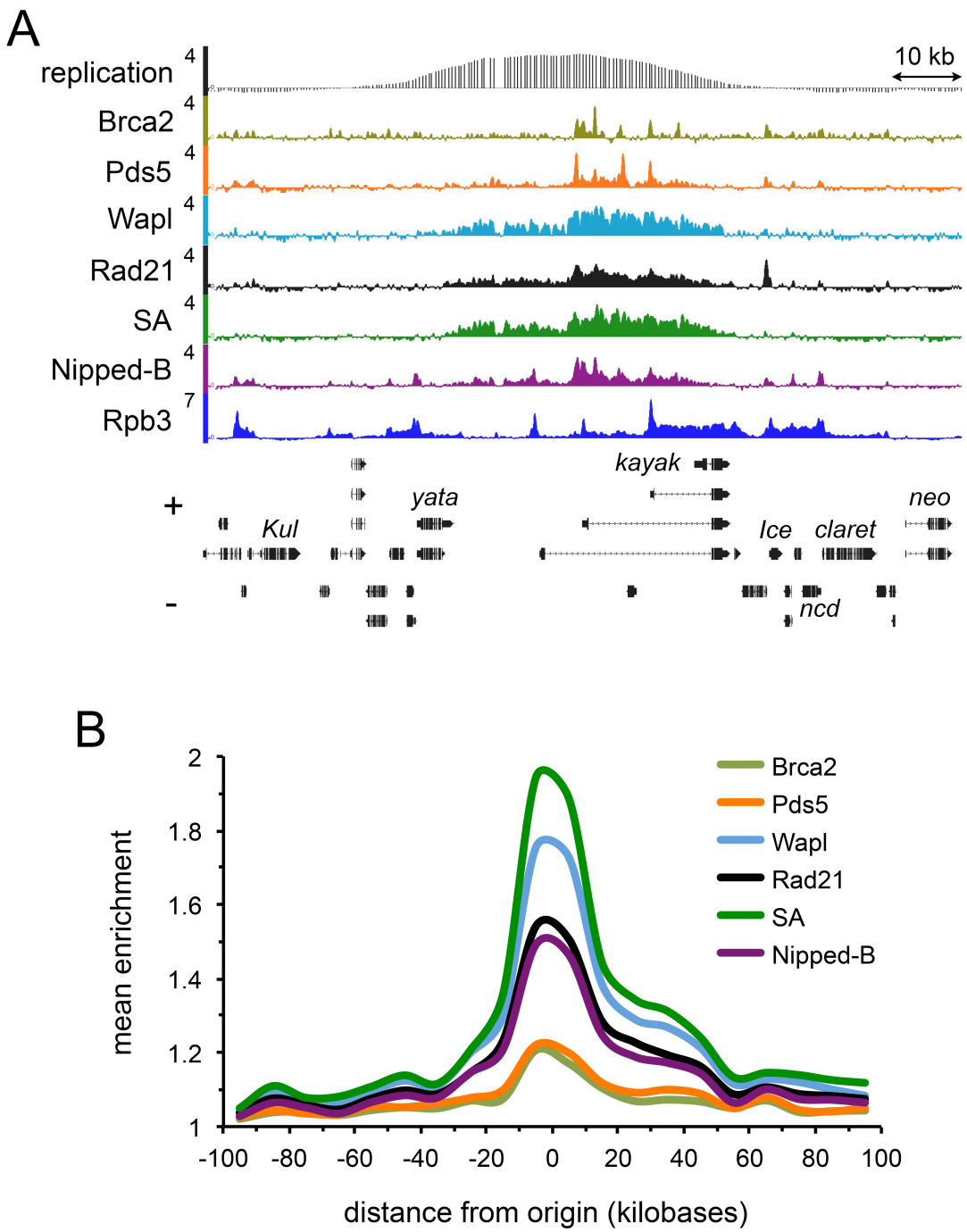


Fig 2. Cohesin and the accessory factors localize to DNA replication origins in BG3 cells. (**A**) Genome browser view of Brca2, Pds5, Wapl, Rad21, SA, Nipped-B and Rpb3 ChIP-seq at the *kayak* locus containing an early DNA replication origin. The scales are log2 enrichment. The processed BG3 early DNA replication data was provided by David MacAlpine (GEO accession GSE17287) and the scale is MA2C score. The SA, Nipped-B, and Rpb3 ChIP-seq data are published elsewhere [12, 14]. (**B**) Meta-origin analysis of the ChIP-seq data using the 78 strongest early DNA replication origins (positions in S2 File). The mean enrichment in 10 kb bins was calculated from -100 kb to +100 kb from the origin centers.

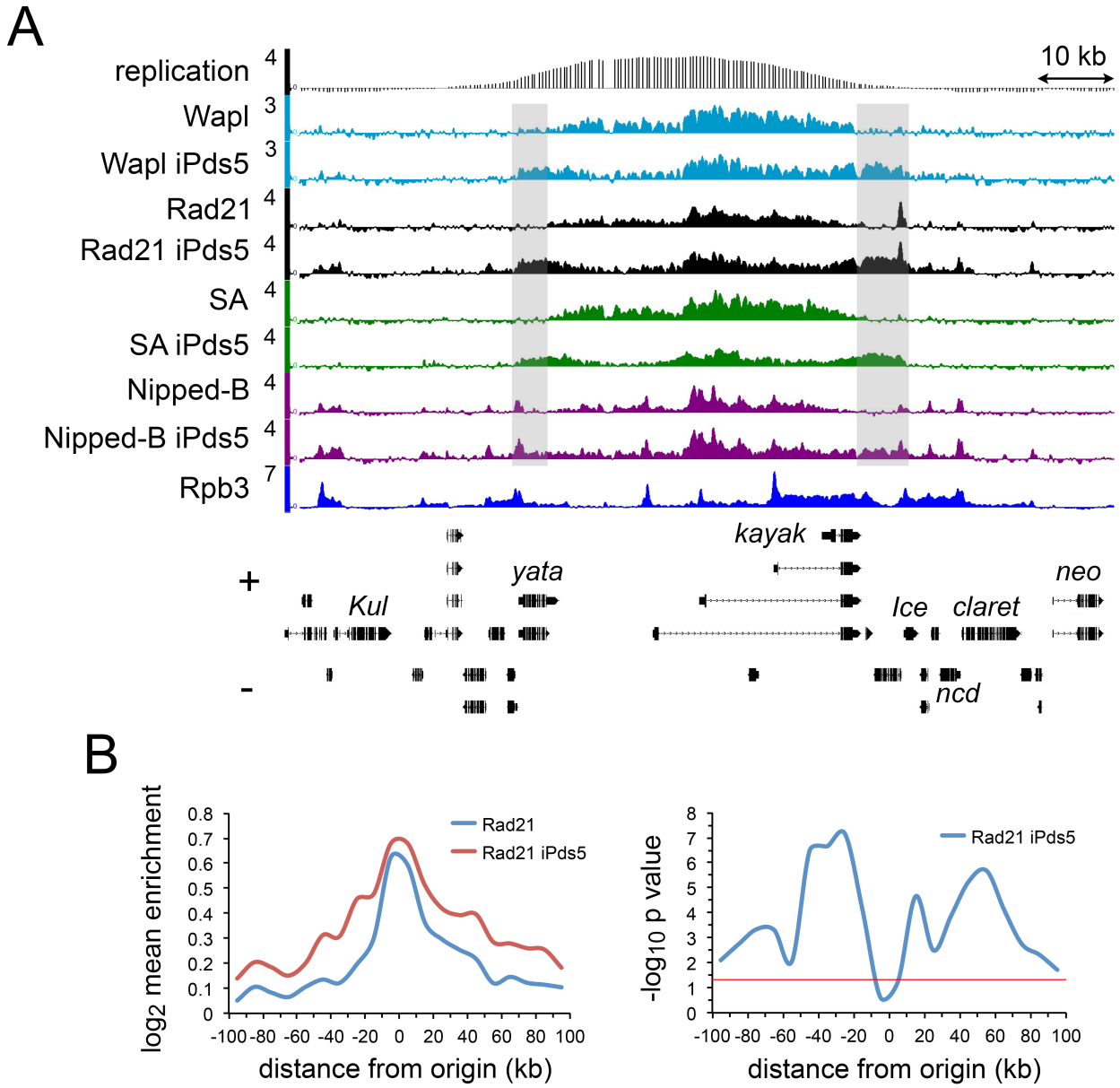


Fig 3. Depletion of Pds5 causes extension of cohesin and accessory factor binding domains surrounding replication origins.

(**A**) Genome browser view of the *kayak* origin region, showing ChIP-seq for Wapl, Rad21, SA, and Nipped-B in control cells and cells depleted for Pds5 (iPds5). Shaded areas show regions with increased binding of cohesin, Wapl and Nipped-B in Pds5-depleted cells. (**B**) The left panel shows the Rad21 meta-origin analysis in mock control cells (Rad21, blue) and cells depleted for Pds5 (Rad21 iPds5, red). The right panel shows the –log10 p values for differences in Rad21 enrichment in each bin used for the meta-origin analysis. P values were calculated using the Wilcoxon signed rank test.

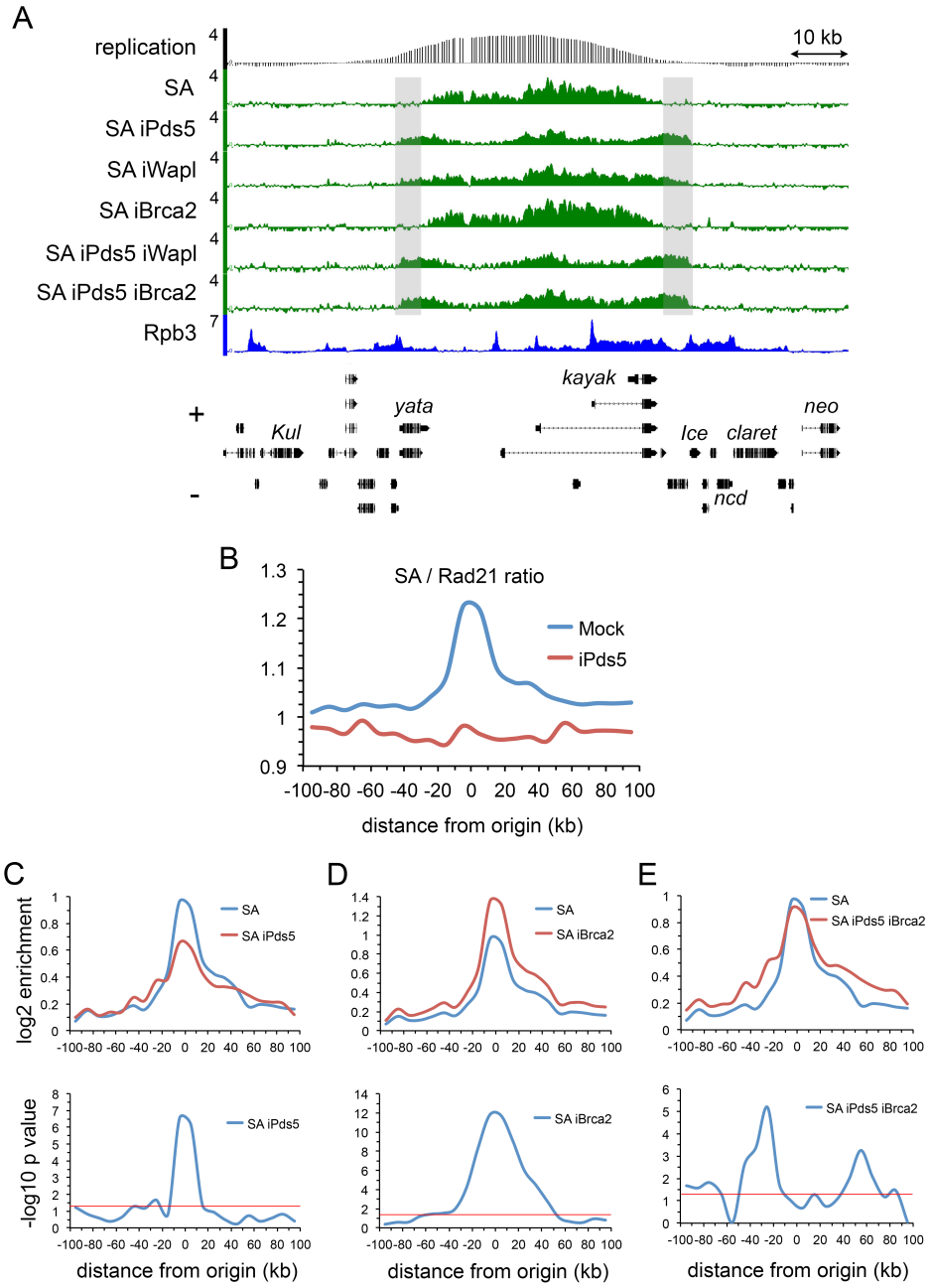


Fig 4. Pds5 and Brca2 have opposing effects on SA cohesin subunit levels at DNA replication origins.

(**A**) Genome browser view of SA ChIP-seq at the *kayak* locus in mock-treated BG3 cells (SA) and BG3 cells RNAi depleted for Pds5 (iPds5), Wapl (iWapl), Brca2 (iBrca2), Pds5 and Wapl (iPds5 iBrca2) and Pds5 and Brca2 (iPds5 iBrca2). (**B**) Meta-origin analysis of the SA to Rad21 ChIP-seq enrichment ratio in mock control cells (Mock, blue) and cells depleted for Pds5 (iPds5, red). (**C**) The top panel is the meta-origin plot of SA enrichment in control (SA, blue) and Pds5-depleted cells (SA iPds5, red). The bottom panel is the meta-origin plot of –log10 p values for the difference in SA enrichment calculated using the Wilcoxon signed rank test. (**D**) Same as C except for cells depleted for Brca2 (iBrca2). (**E**) Same as C with cells depleted for both Pds5 and Brca2 (iPds5 iBrca2).

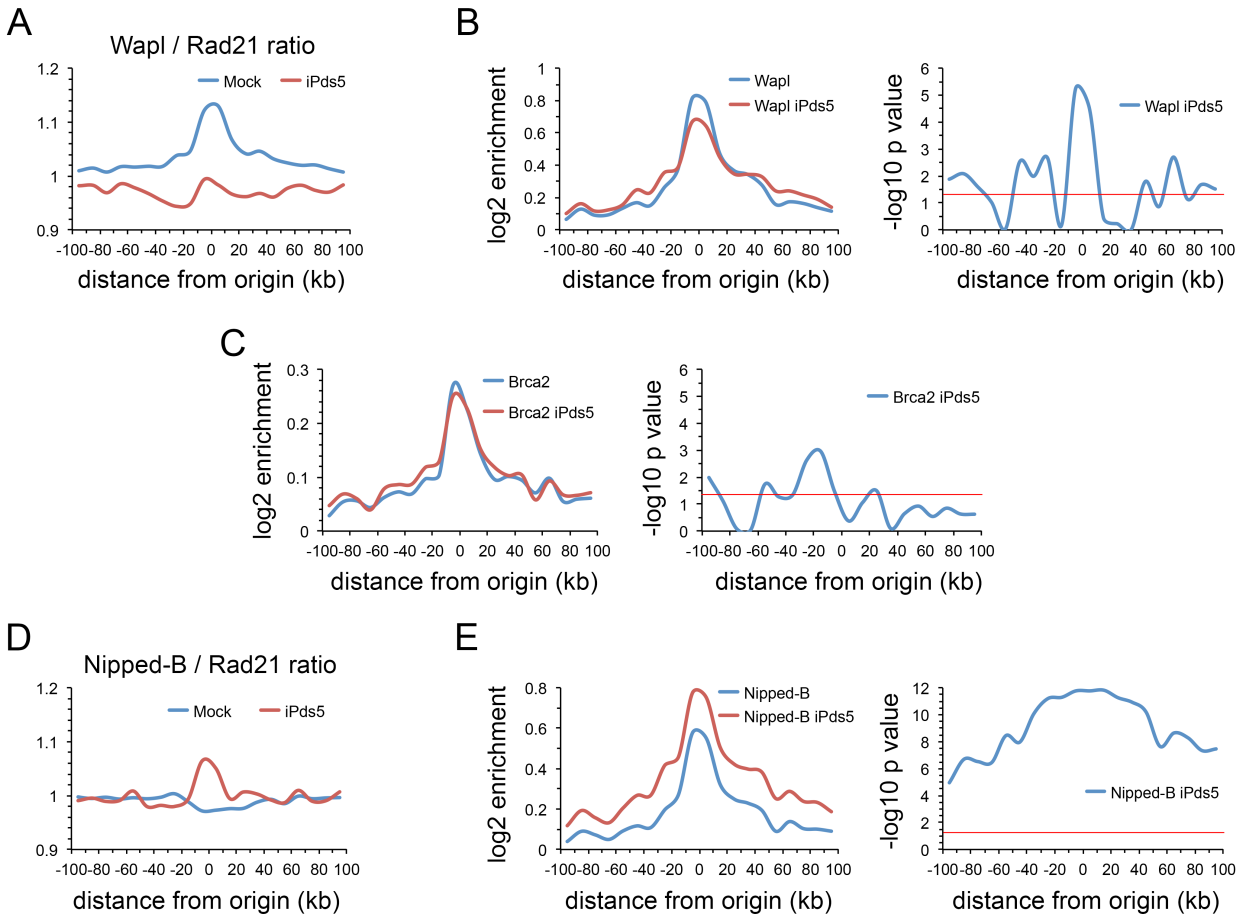


Fig 5. Pds5 influences Wapl and Nipped-B binding at DNA replication origins with little effect on Brca2.

(**A**) Meta-origin plot of Wapl to Rad21 ratio in control cells (Mock, blue) and cells depleted for Pds5 (iPds5, red). (**B**) Left panel is meta-origin analysis of Wapl ChIP-seq enrichment in control cells (Wapl, blue) and cells depleted for Pds5 (Wapl iPds5, red). Right panel is the plot of –log10 p values for the differences in Wapl enrichment in the meta-origin bins calculated using the Wilcoxon signed rank test. (**C**) Same as B for Brca2 enrichment. (**D**) Same as A for the Nipped-B to Rad21 ratio. (**E**) Same as B for Nipped-B ChIP-seq enrichment.

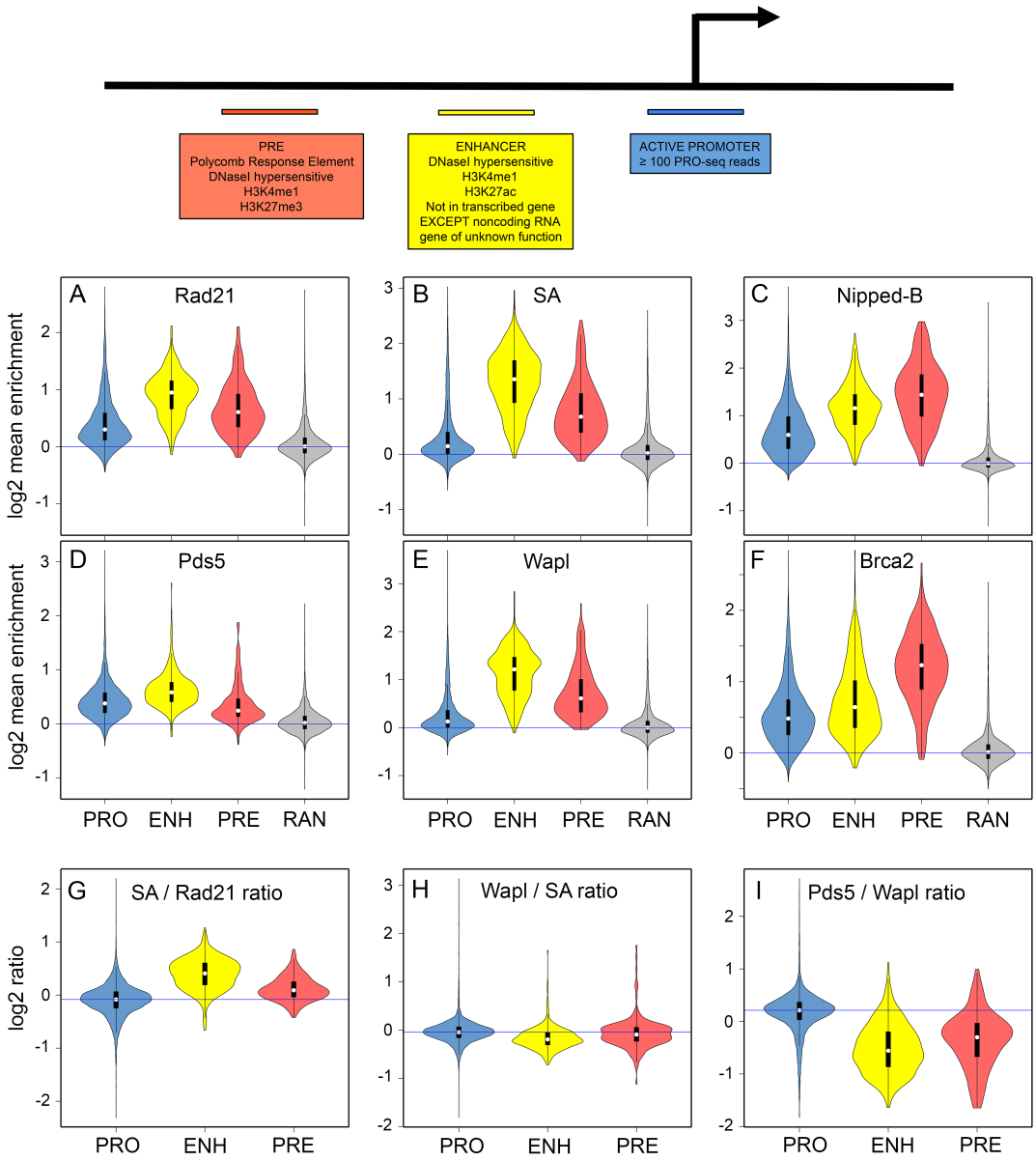


Fig 6. The relative levels of cohesin, Nipped-B, Pds5, Wapl and Brca2 vary between active promoters, enhancers and Polycomb Response Elements (PREs).

The top diagram summarizes how active promoters (blue) extragenic enhancers (yellow) and PREs (orange) are defined as 500 bp sequences as described elsewhere [12, 16, 23]. There are 7,389 non-heterochromatic active promoters, 523 extragenic enhancers and 195 PREs. There are over 2,500 total active enhancers in BG3 cells but intragenic enhancers are excluded to avoid effects caused by changes in transcription. (**A**) Violin plots of the distribution of Rad21 ChIP-seq enrichment values (mean enrichment in each 500 bp element) for promoters (PRO, blue), extragenic enhancers (ENH, yellow) and PREs (PRE, orange), and 6,892 random 500 bp sequences as a negative control. White dots show the median values. (**B**) Same as A for SA. (**C**) Same as A for Nipped-B. (**D**) Same as A for Pds5. (**E**) Same as A for Wapl. (**F**) Same as A for Brca2. (**G**) Distribution of SA to Rad21 ChIP-seq enrichment ratios for promoters (PRO) enhancers (ENH) and PREs (PRE). (**H**) Same as G for the Wapl to SA ratio. (**I**) Same as G for the Pds5 to Wapl ratio.

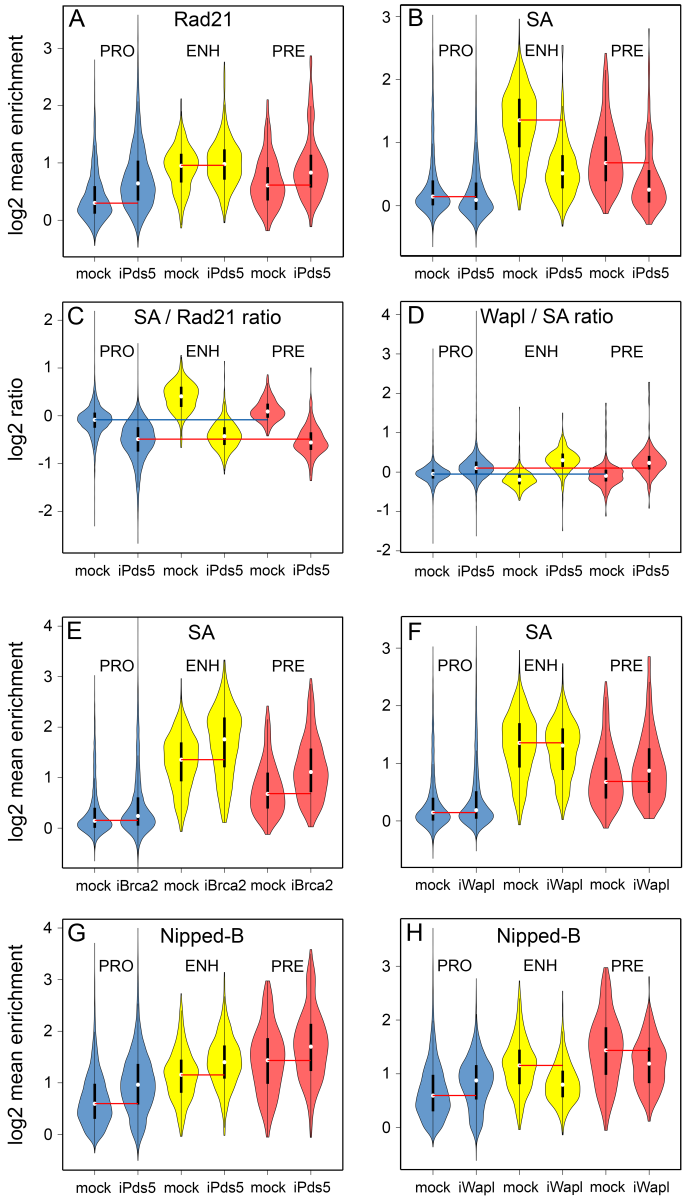


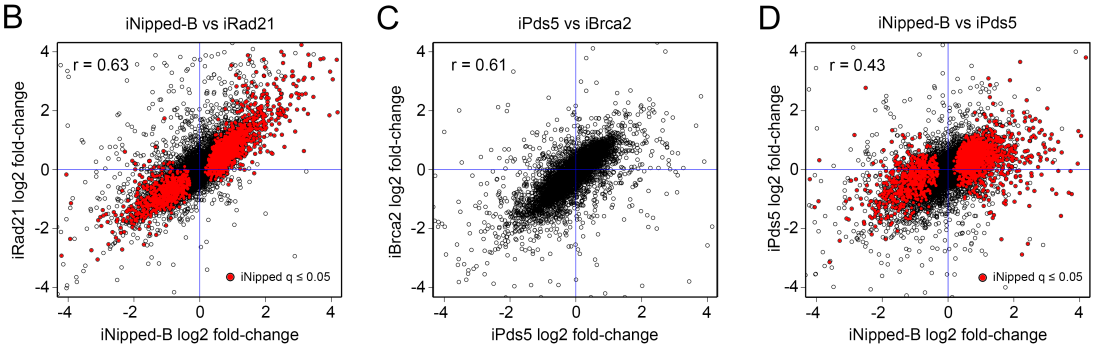
Fig 7. Pds5, Brca2 and Wapl differentially influence the levels of cohesin subunits and Nipped-B at gene regulatory sequences.

(**A**) The violin plots show the distributions of Rad21 ChIP-seq enrichment values at promoters (PRO, blue) enhancers (ENH, yellow) and PREs (PRE, orange) in mock-treated control BG3 cells (mock) and cells depleted for Pds5 (iPds5). Red lines indicate the median values for each type of regulatory sequence in mock control cells. (**B**) Same as A for SA enrichment. (**C**) Distributions of the SA to Rad21 ratios at promoters (PRO) enhancers (ENH) and PREs (PRE) in control (mock) and Pds5-depleted (iPds5) BG3 cells. The blue line indicates the median ratio at promoters in mock control cells and the red line indicates the median ratio at promoters in Pds5-depleted cells. (**D**) Same as C for the Wapl to SA ratio. (**E**) Same as A for SA in control and Brca2-depleted (iBrca2) cells. (**F**) Same as A for SA in control and Wapl-depleted (iWapl) cells. (**G**) Same as A for Nipped-B enrichment. (**H**) Same as A for Nipped-B in mock control and Wapl-depleted (iWapl) cells. Statistical tests of the differences in the distributions of ChIP-seq enrichment after protein depletions in panels A, B, E, F, G and H are provided in S6 Table.

A

Pearson correlation coefficients

	iRad21	iPds5	iWapl	iBrca2	iBrca2 iPds5	iBrca2 iWapl
iNipped-B	**0.63**	0.43	0.24	0.44	0.12	0.33
iRad21		0.28	0.25	0.21	0.27	0.25
iPds5			0.38	**0.61**	0.43	0.47
iWapl				0.28	**0.73**	**0.68**
iBrca2					0.20	0.44
iBrca2 iPds5						**0.68**



E

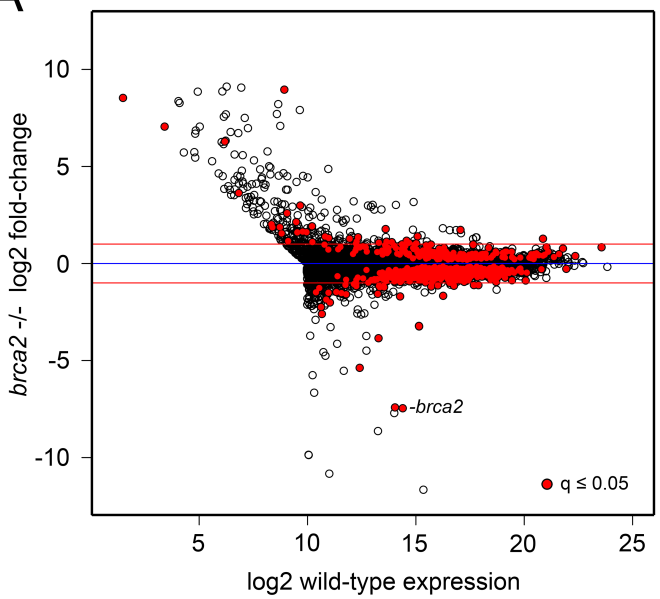
Overlaps in genes that increase and decrease in expression (p ≤ 0.05)

		iRad21		iPds5		iWapl		iBrca2		iBrca2 iPds5		iBrca2 iWapl	
		1475	604	1030	292	1011	318	1111	420	598	240	991	379
iNipped-B	1830	1047	1	609	12	626	36	558	8	280	49	523	51
	696	1	401	28	133	14	115	39	172	36	53	42	106
iRad21	1475			342	12	386	45	298	16	242	49	333	57
	604			37	70	27	83	55	100	35	50	47	83
iPds5	1030					503	1	581	1	287	0	414	4
	292					1	83	0	168	3	42	2	71
iWapl	1011							529	0	346	0	616	1
	318							0	129	1	162	0	213
iBrca2	1111									207	1	476	2
	420									3	65	2	117
iBrca2 iPds5	598											333	2
	240											0	193

Fig 8. Pds5 and Brca2 have similar effects on gene expression in BG3 cells that overlap the effects of Nipped-B and cohesin.

(**A**) Genome-wide Pearson correlation coefficients for the log2 fold-changes in RNA levels caused by depletion of Nipped-B (iNipped-B), Rad21 (iRad21), Pds5 (iPds5), Wapl (iWapl), Brca2 (iBrca2), Brca2 and Pds5 (iBrca2 iPds5) and Brca2 and Wapl (iBrca2 iWapl). Gene expression values used for the analysis are in S7 Data. (**B**) Dot plot of log2 fold-changes in RNA levels caused by Nipped-B depletion versus the changes caused by Rad21 depletion. Red dots show statistically significant changes in gene expression caused by Nipped-B depletion (q ≤ 0.05). (**C**) Dot plot of log2 fold-changes in RNA levels caused by Pds5 depletion versus changes caused by Brca2 depletion. (**D**) Dot plot of log2 fold-changes in RNA levels caused by Nipped-B depletion versus the changes caused by Pds5 depletion. Red dots show genes significantly altered by Nipped-B depletion (q ≤ 0.05). (**E**) Overlap in the genes that increase and decrease in expression with the indicated depletions at p ≤ 0.05. P values were used instead of the more stringent q values to obtain larger groups of genes. Numbers in red indicate genes that increase in expression and numbers in blue are genes that decrease. The numbers in the overlap boxes show the number that change with both depletion treatments. Red indicates genes that increase with both and blue indicate genes that decrease with both. Brown indicates genes that increase with one treatment, and decrease with the other. All overlaps in genes that increase or decrease in expression are statistically significant by Fisher's exact test (S7 Data).

A



B

gene	fold-change	p value	q value	description
General Transcription Factors				
Bx	0.73	0.008	0.045	LIM only protein
Cdk12	0.83	0.004	0.036	Pol II kinase
CtBP	0.85	0.008	0.044	co-repressor
MED14	0.84	0.010	0.047	Mediator complex subunit
MED25	0.76	0.006	0.040	Mediator complex subunit
ptip	0.63	0.003	0.032	Trr histone methyltransferase complex
Set1	0.73	0.005	0.037	histone methyltransferase
Spt5	0.74	0.000	0.010	pausing elongation factor
Su(Tpl)	0.69	0.003	0.030	Super Elongation Complex subunit
Su(z)2	0.58	0.002	0.029	Polycomb Repressive Complex 1 subunit
Trf2	0.67	0.003	0.031	TATA binding protein
Trl	0.84	0.003	0.031	GAGA factor
Utx	0.78	0.002	0.028	histone demethylase
Cell Cycle / Chromosome Structure				
Bub1	0.72	0.005	0.038	mitotic spindle checkpoint kinase
BubR1	0.67	0.003	0.031	Bub1-related kinase
Cap-D3	0.59	0.003	0.031	Condensin II subunit
Cap-H2	0.51	0.007	0.042	Condensin II subunit
cdc16	0.70	0.003	0.032	Anaphase Promoting Complex subunit
cdc2c	0.75	0.010	0.048	Cdk1 kinase
Cen	0.68	0.010	0.048	centrocortin centrosome component
Cenp-C	0.62	0.009	0.047	centromere protein
CycE	0.67	0.000	0.013	S phase cyclin
CycG	0.79	0.009	0.046	cyclin
JIL-1	0.72	0.006	0.040	histone kinase
DNA Repair				
FANCI	0.68	0.006	0.039	Fanconi anemia group I
Lig4	0.78	0.000	0.008	DNA repair ligase
mus308	0.69	0.005	0.038	DNA polymerase theta
mus81	0.66	0.008	0.045	DNA repair endonuclease
RecQ5	0.61	0.005	0.038	DNA helicase
Sister Chromatid Cohesion				
dmt	0.76	0.007	0.043	Dalmatian sororin-shugoshin
pds5	0.82	0.000	0.006	

Fig 9. Brca2 influences gene expression in developing wings.
(**A**) The log2 fold-change in gene expression in *brca2* null mutant 3rd instar wing imaginal discs is plotted versus the log2 expression level in control wing discs for all active genes. Active genes are defined as those that are expressed at or above the median level in control discs plus those expressed at or above the control median level in *brca2* mutant discs. Red dots indicate statistically significant changes in gene expression ($q \leq 0.05$). The dot representing the *brca2* gene is labeled. The blue line indicates no change, and the two red lines indicate 2-fold increases or decreases. (**B**) Examples of genes down-regulated in *brca2* mutant discs in the indicated categories. This is not a comprehensive list, which can be generated from the expression data provided in S7 Data.

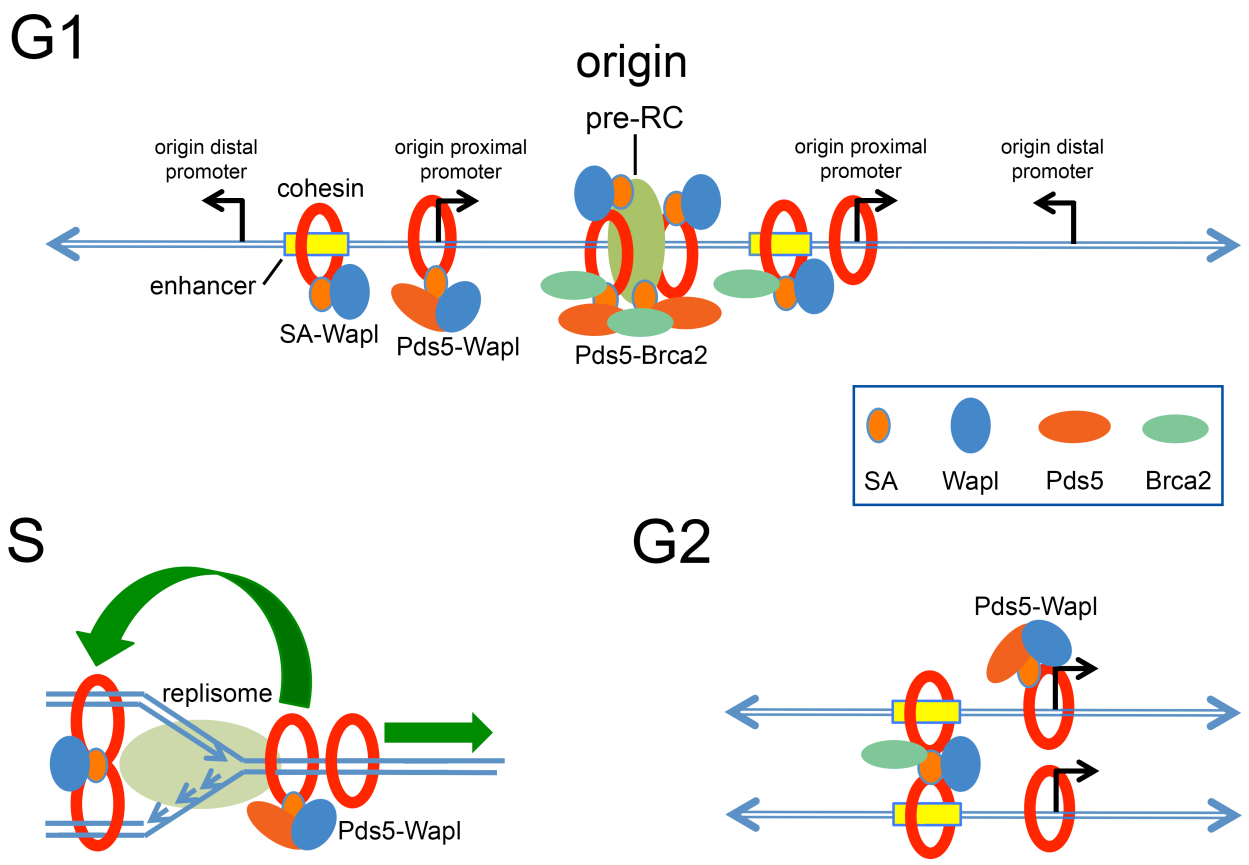
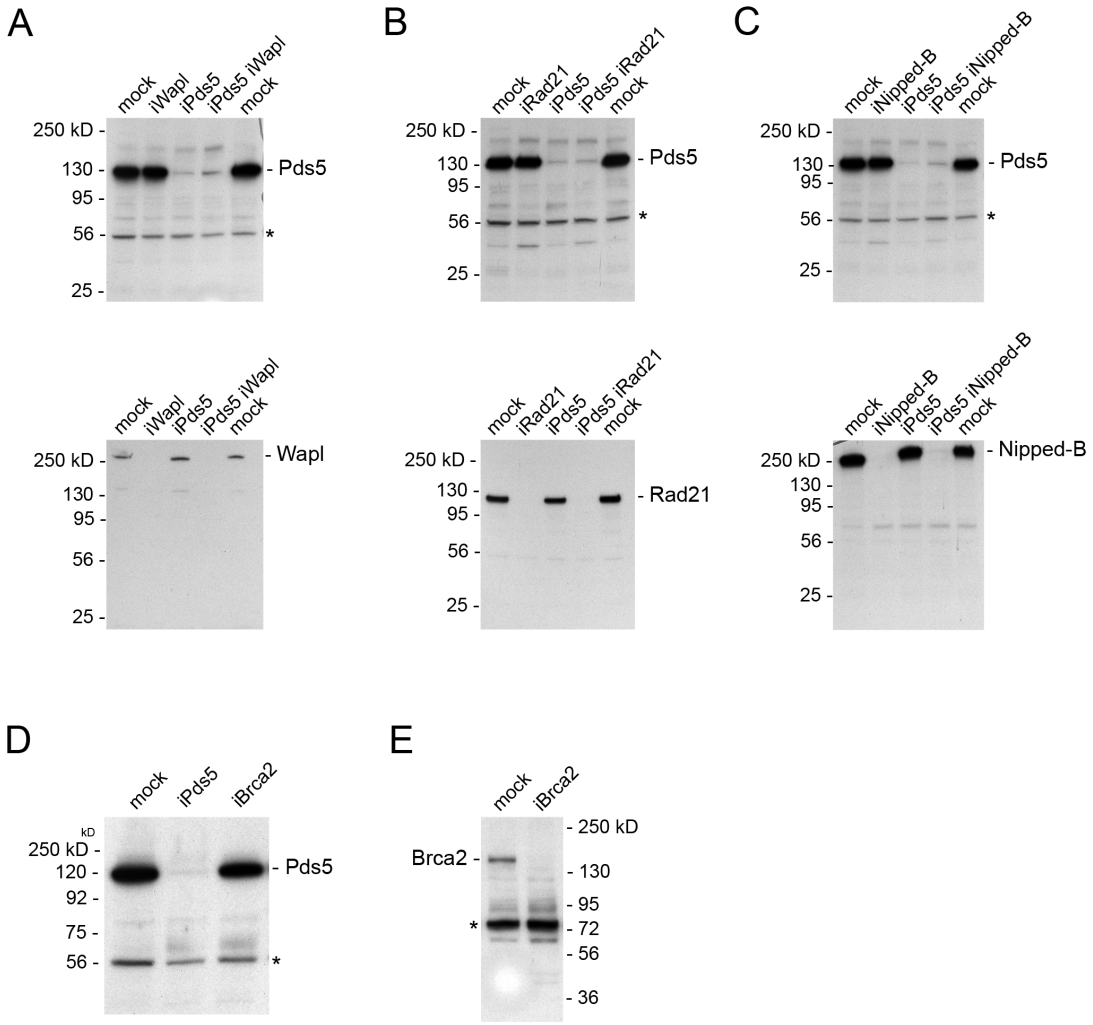


Fig 10. Models for the roles of Pds5, Wapl and Brca2 in regulating cohesin dynamics and function in Drosophila cells.

At the top, in the G1 phase of the cell cycle, active promoters (angled arrows) and enhancers (yellow boxes) proximal to DNA replication origins bind cohesin (red rings) while origin-distal promoters and inactive genes do not. For simplicity, the Nipped-B cohesin-loading factor is not depicted, but it is present wherever there is cohesin. We posit that book-marking proteins (not depicted) at promoters and enhancers remain bound through mitosis to recruit Nipped-B and enable cohesin loading after cell division. Enhancers have a relatively high level of the SA cohesin subunit (small orange oval) compared to promoters. Wapl (blue oval) is stoichiometric with SA, and unlike Pds5 (large orange oval) binds everywhere with cohesin. Promoters have a high Pds5 to Wapl ratio, and the Pds5-Wapl cohesin removal complex keeps cohesin levels comparatively low. The pre-replication complex (large light green oval) containing the origin recognition complex (ORC) and the MCM helicase complex licenses DNA replication origins (origin) in early G1 and recruits Nipped-B resulting in cohesin loading and binding of Pds5, Wapl, and Brca2. Pds5 aids and Brca2 inhibits SA binding at origins to titrate the fraction of cohesin complexes that mediate sister cohesion during G2. At origins, Pds5 and Wapl do not remove cohesin. Wapl inhibits Pds5 binding, which requires cohesin and Brca2. Pds5 restricts binding of Nipped-B via competition for cohesin. During S phase (lower left) the replisome pushes cohesin ahead of the replication fork. The Pds5-Wapl complex unloads cohesin in front of the fork, limiting cohesin spreading. Cohesin is reloaded behind the fork to establish sister chromatid cohesion, which requires SA. A handcuff model is shown, but cohesion mechanisms with single cohesin rings are possible. During G2 (lower right) enhancers have a high SA to cohesin ratio, and a low Pds5 to Wapl ratio, making them sister chromatid cohesion sites. Cohesion is low at promoters, which have a low SA to cohesin ratio, and a high Pds5-Wapl ratio. We envision that the unrestrained promoters loop independently to the double enhancer complex held

together by sister cohesion, aided by interactions between Nipped-B, cohesin and the Mediator complex (not depicted). Pds5, Brca2 and Wapl control cohesin binding dynamics and cohesin-dependent looping to influence gene expression.

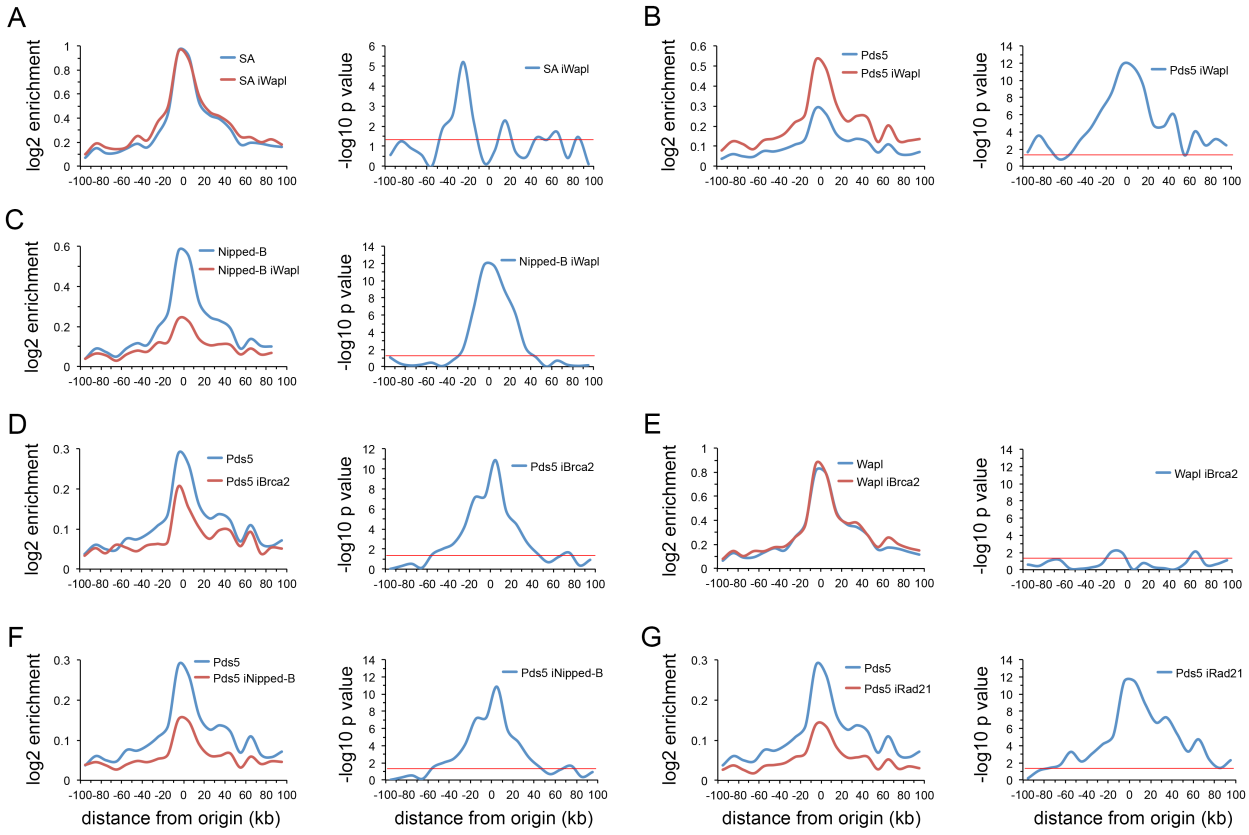
Supporting information



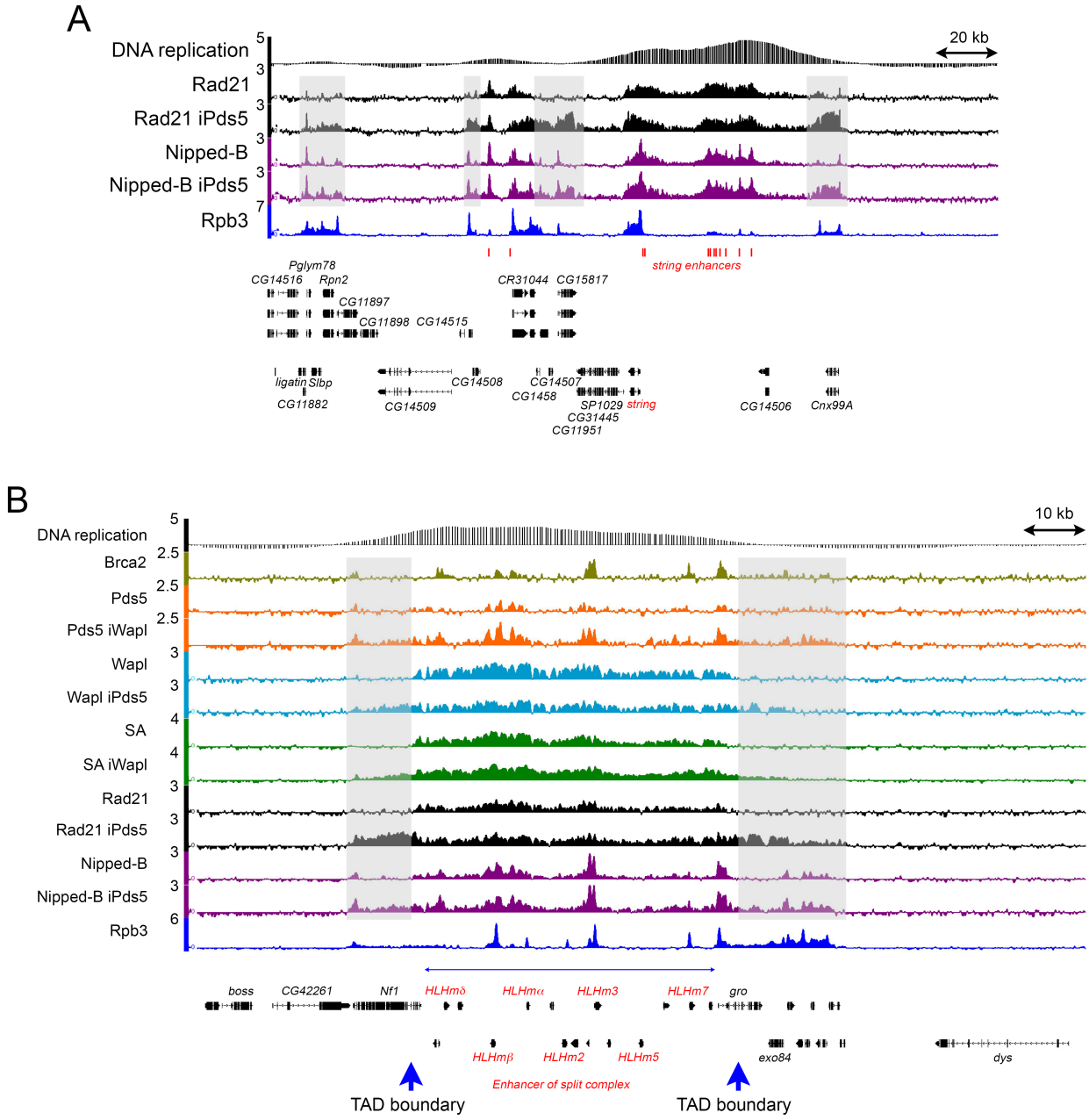
S1 Fig. Example western blots showing RNAi depletion of Pds5, Wapl, Brca2, Nipped-B and Rad21 in BG3 cells.

(**A**) The top and bottom panels are matched western blots of whole cell extracts of cells after the indicated single and double RNAi treatments for 4 to 5 days. The top panel is probed with anti-Pds5 antibody and the bottom panel is probed with anti-Wapl. The asterisk (*) indicates a non-specific band recognized by the Pds5 antibody used as a loading control. (**B**) Matched western blots with extracts from cells with the indicated single and double RNAi treatments. The top panel was probed with anti-Pds5 and the bottom panel was probed with anti-Rad21. (**C**) Matched western blots with extracts from cells with the indicated RNAi treatments. The top was probed with anti-Pds5 and the bottom was probed with anti-Nipped-B. (**D**) Western of extracts of cells after the indicated RNAi treatments probed with anti-Pds5. (**E**) Western of extract of cells after the indicated RNAi treated probed with anti-Brca2. The asterisk (*) indicates a non-specific band recognized by anti-Brca2.

S2 File. Locations of strongest early DNA replication origins in BG3 cells.

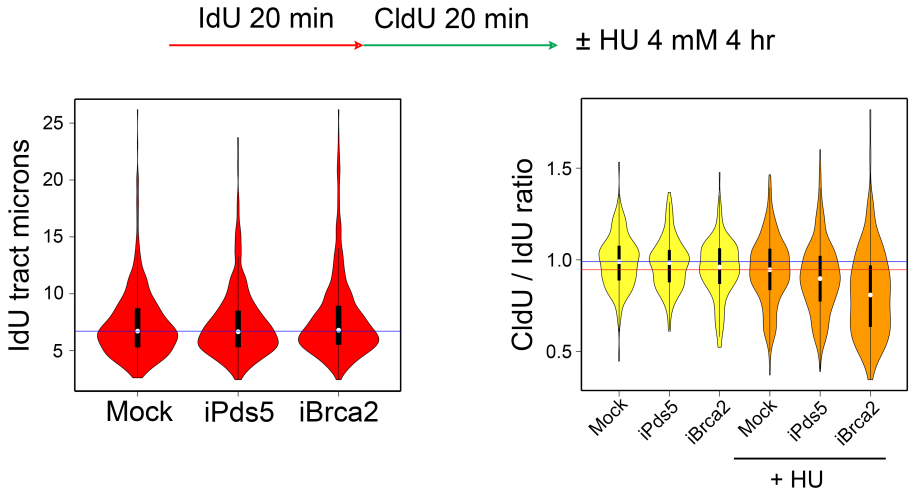


S3 Fig. Meta-origin analyses in BG3 cells after Wapl, Brca2, Nipped-B and Rad21 depletion. (**A**) Left panel is the SA distribution in mock-treated control cells (blue, SA) and cells depleted for Wapl (red, SA iWapl). Right panel is the -log10 p values of each bin for the difference in control versus the depletion calculated using the Wilcoxon signed rank test. (**B**) Same as A for the Pds5 distribution. (**C**) Same as A for the Nipped-B distribution. (**D**) Left panel is the Pds5 meta-origin distribution in control (blue, Pds5) cells and cells depleted for Brca2 (red, Pds5 iBrca2). Right panel shows the –log10 p values. (**E**) Same as D for Wapl distribution. (**F**) Left panel is the Pds5 distribution in control (blue, Pds5) cells, and cells depleted for Nipped-B (red, Pds5 iNipped-B). Right panel shows the –log10 p values. (**G**) Same as F except for Rad21 depleted cells.



S4 Fig. Examples of extended cohesin domains upon Pds5 or Wapl depletion.
(**A**) Genome browser view of the *string* region showing skipping of spreading cohesin over inactive genes and intergenic regions. The *string* (*Cdc25*) gene and its active enhancers are labeled in red type. The top track is the early DNA replication pattern. The ChIP-seq tracks show the Rad21 (black) and Nipped-B (purple) binding patterns in control cells and cells depleted for Pds5 (iPds5). The bottom track shows the Rpb3 RNA polymerase pattern (blue). Shaded areas indicate regions of increased Rad21 and Nipped-B occupancy in Pds5-depleted cells. Pds5 depletion increases Rad21 and Nipped-B occupancy at the active gene cluster containing *Pglym78* some 100 kb centromere-proximal (left) to the origin, but not with inactive genes such as *CG14509* in the intervening region. Similarly, cohesin and Nipped-B increase substantially at the active *Cxn99A* gene to the right of the origin, but show only a modest increase in the intervening intergenic region. *CG15817* is active but binds much less cohesin than some active genes located further from the origin. Pds5 depletion increases cohesin and Nipped-B at *CG15817*. (**B**)

Enhancer of split gene complex showing expansion of cohesin domains beyond TAD (topologically associating domain) boundaries upon Pds5 or Wapl depletion. Enhancer of split genes are labeled in red, and the thin blue horizontal arrow just above the genes shows the extent of the gene complex. Wide vertical blue arrows below the genes indicate the TAD boundaries determined by 3C (chromosome conformation capture) [17]. The top track shows the early DNA replication pattern. The ChIP-seq tracks show the Pds5 (orange) pattern in control cells and cells depleted for Wapl (iWapl), the Wapl pattern (light blue) in control cells and cells depleted for Pds5 (iPds5), the SA (green) binding in control cells and cells depleted for Wapl (iWapl), the Rad21 (black) binding in control cells and cells depleted for Pds5 (iPds5), the Nipped-B (purple) pattern in control cells and cells depleted for Pds5 (iPds5), and the Rpb3 (blue) RNA polymerase binding in control cells. Shaded areas indicate regions with increased occupancy by cohesion proteins upon Pds5 or Wapl depletion.



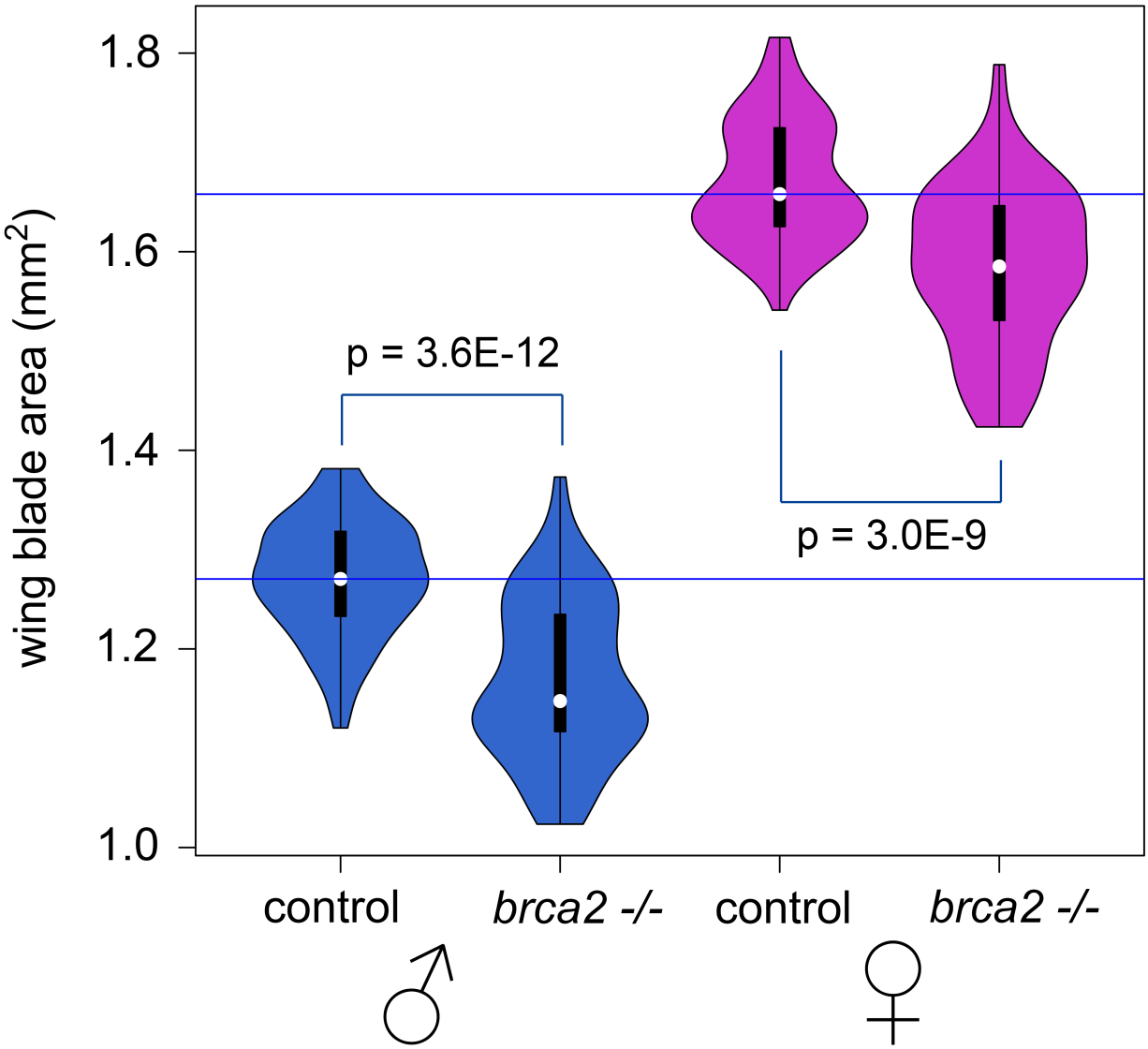
	N	median	p value vs Mock
Mock	601	6.72	
iPds5	459	6.67	9.91E-01
iBrca2	434	6.80	9.77E-02

	Control			HU		
	N	median	p value vs Mock	N	median	p value vs Mock HU
Mock	320	0.99		281	0.95	
iPds5	173	0.98	3.49E-01	286	0.90	6.40E-03
iBrca2	152	0.96	3.90E-02	282	0.81	2.84E-12

S5 Fig. DNA fiber replication analysis in BG3 cells depleted for Pds5 and Brca2.
The top diagram shows the labeling scheme and the hydroxyurea (HU) treatment after CldU incorporation. The left panel shows the lengths of the IdU tracts in Mock control cells, and cells depleted for Pds5 (iPds5) or Brca2 (iBrca2). The blue line indicates the median tract length in Mock control cells. The IdU tracts were combined for the HU-treated and untreated cells, and only IdU tracts continuous with CldU tracts were measured. The table beneath the left panel gives the median tract lengths in microns for each group and the p values using the Wilcoxon test. The right panel shows the ratio of lengths of the CldU tracts to the connected IdU tracts. The blue line indicates the median ratio in Mock control cells not treated with HU, and the red line indicates the median ratio in Mock control cells treated with HU. The table beneath the right panel gives the median ratio for all groups and the p values for the indicated comparisons using the Wilcoxon test. Similar results were obtained in two other independent experiments.

S6 Table. Statistical tests of changes in cohesin and Nipped-B ChIP-seq enrichment at promoters, enhancers and PREs upon Pds5, Brca2 or Wapl depletion in BG3 cells.

S7 Data. RNA-seq data in BG3 cells depleted for Nipped-B, Rad21, Pds5, Wapl, and Brca2, and control and brca2 mutant female 3rd instar wing imaginal discs.



S8 Fig. Adult wing areas in *brca2* mutant and control flies.
Violin plots show the distributions of adult wing blade areas from the indicated groups of flies grown at 25°. The *brca2* -/- mutant wings are from ~20 homozygous $brca2^{KO}$ and ~40 $brca2^{KO}$ / $brca2^{56e}$ flies and the control wings are from ~20 each from *cn bw*, Oregon R, $brca2^{KO}$ / *cn bw* and $brca2^{KO}$ / *cn bw* flies. The various mutant and control genotypes were combined to minimize the effects of genetic background differences that can influence growth. P values were calculated using the Wilcoxon test. Horizontal blue lines indicate the median wing areas for male and female control flies.