Polytene Chromosome Structure and Somatic Genome Instability

Allan C. Spradling
Department of Embryology, Howard Hughes Medical Institute, Carnegie Institution for Science, Baltimore, Maryland 21218
Correspondence: spradling@carnegiescience.edu

Polytene chromosomes have for 80 years provided the highest resolution view of interphase genome structure in an animal cell nucleus. These chromosomes represent the normal genomic state of nearly all *Drosophila* larval and many adult cells, and a better understanding of their striking banded structure has been sought for decades. A more recently appreciated characteristic of *Drosophila* polytene cells is somatic genome instability caused by unfinished replication (UR). Repair of stalled forks generates enough deletions in polytene salivary gland cells to alter 10%–90% of the DNA strands within more than 100 UR regions comprising 20% of the euchromatic genome. We accurately map UR regions and show that most approximate large polytene bands, indicating that replication forks frequently stall near band boundaries in late S phase. Chromosome conformation capture has recently identified dense topologically associated domains (TADs) in many genomes and most UR bands are similar or slightly smaller than a cognate *Drosophila* TAD. We argue that bands serve the evolutionarily ancient function of coordinating genome replication with local gene activity. We also discuss the relatively recent evolution of polyteny and somatic instability in Diptera and propose that these processes helped propel the amazing success of two-winged flies in becoming the most ecologically diverse insect group, with 200 times the number of species as mammals.

Polyploid cells are produced during normal development when progenitors switch to a cell cycle without cytokinesis, a process that occurs at some level in most or all species. Diploid cells also may become polyploid to repair tissue damage (Losick et al. 2013, 2016). By endocycling (cycling but not dividing), polyploid cells grow in balance with genome copy number, a strategy that scarcely perturbs cellular physiology or gene regulation and explains why polyploid versions of diverse common cell types can be found across the phylogenetic spectrum (Nagl 1978; Edgar et al. 2014; Neiman et al. 2017). In contrast, polyploid cells that can return to the mitotic cycle, cells often generated by cytokinesis failure or from certain programmed endocycles (Fox et al. 2010), are susceptible to increased genomic instability and prone to oncogenesis (Fujiwara et al. 2005; Duncan et al. 2010; Schoenfelder et al. 2014). Why polyploid cells are so common has remained a matter of debate, but large cells may be mechanically advantageous and stress-resistant (Orr-Weaver 2015; Schoenfelder and Fox 2015; Neiman et al. 2017).

Dipteran polyploid cell chromosomes are termed "polytene" because they maintain replicated sister strands and homologs in exceptionally close association compared with other polyploid cells. Enhanced alignment is associated with resetting the endocycle near the end of S phase, rather than in G_2 or M phase as in most polyploid cells. Their origin from a cell cycle ending in S phase suggests that polytene chromosomes keep multiple sister chromosomes tightly aligned by stabilizing normal S phase pairing involving Cohesins (Fig. 1A,B). Polytene chromosomes have been uniquely valuable for analyzing higher-order chromatin organization (Bridges 1935; Ashburner 1970, Lefevre 1976; Zhimulev 1996; Stormo and Fox 2017). The normal functioning of diverse cell types possessing such chromosomes, which fold in a characteristic manner but lack consistent contact points between different chromosome arms (Mathog et al. 1984), sounds a cautionary note for theories postulating intricate three-dimensional interactions between distant genomic regions.

Consistent with S phase resets, polytene cells in Dipterans generally show "underreplication" of satellite-rich regions of centromeric heterochromatin, whereas *Drosophila* and some other higher Dipteran species also underreplicate-specific euchromatic regions that normally duplicate late in S phase (Gall et al. 1971; Hammond and Laird 1985; Karpen and Spradling 1990; Spradling 1993; Moshkin et al. 2001; Belyakin et al. 2005; Nordman et al. 2011; Yarosh and Spradling 2014). *Drosophila* euchromatic unfinished replication (UR) regions were originally characterized in the larval salivary gland, but most of the same major regions undergo UR in other polytene tissues examined, including larval midgut, larval fat body, and adult ovary (Nordman et al. 2011; Yarosh and Spradling 2014). UR regions in several tissues contain few expressed genes, are enriched for repressive chromatin marks, and are depleted for candidate replication origins (Sher et al. 2012). The suppressor-of-underreplication gene product (SUUR) is essential for euchromatic UR (Belyaeva et al. 1998); it binds the replication fork protein PCNA and may slow elongation (Nordman et al. 2014). Copy-number changes generated by underreplication have been proposed to result from persistent nets of unfinished

Published by Cold Spring Harbor Laboratory Press; doi: 10.1101/sqb.2017.82.033670

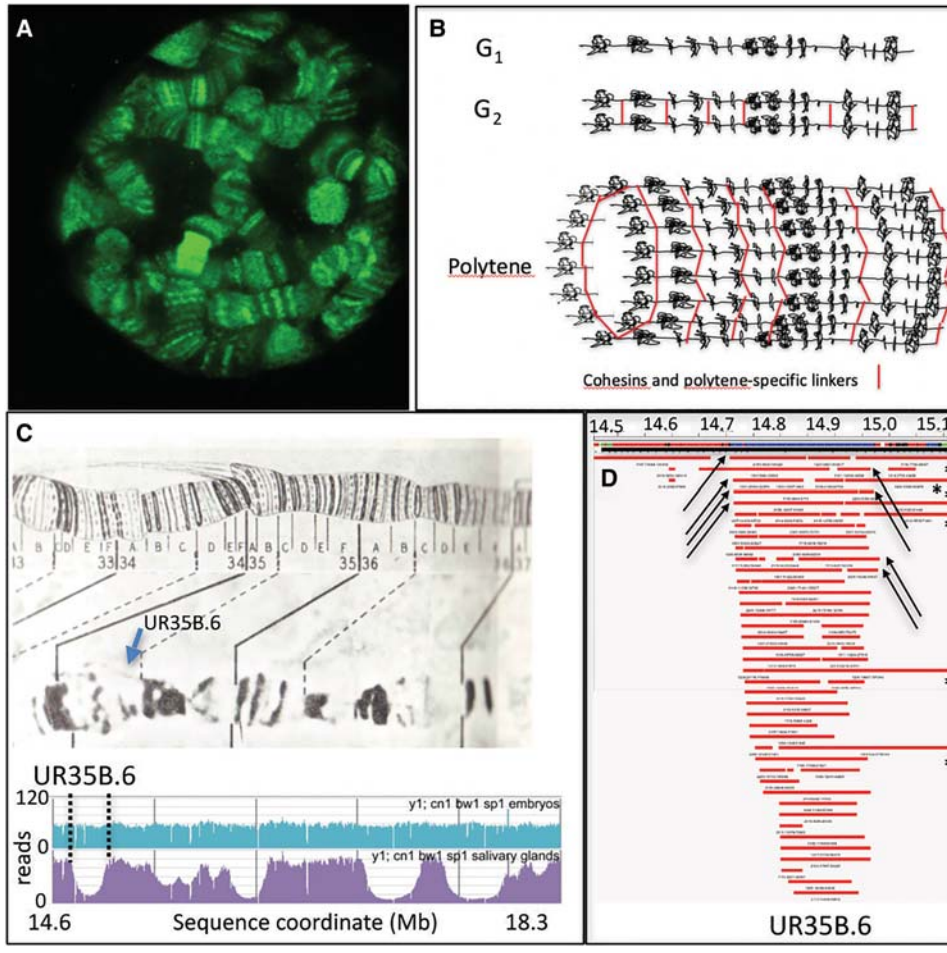

Figure 1. Polytene chromosome structure and somatic instability. (*A*) Polytene chromosome within an intact larval salivary gland nucleus, as revealed by green fluorescent protein (GFP) fluorescence of the protein trap line CC00258 (Buszczak et al. 2007). (*B*) G_1: Model of a G_1 diploid chromosome—a unit chromatin fiber containing highly folded territories separated by more extended regions. G_2: Model of a G_2 diploid chromosome—two unit fibers held together by cohesions (red). Polytene: Model of a polytene chromosome—multiple unit fibers arrayed to form a hollow cylinder and held together by cohesins and novel polytene pairing factors (red). (*C*) Bridges map (Bridges 1935) and Lefevre photograph (Lefevre 1976) of region 34–36 on chromosome 2L, the most somatically unstable region in euchromatin. Sequence read profiles from this region are shown below for diploid (*upper*) and salivary gland (*lower*) DNA. Underreplicated UR regions are revealed as smooth domains of reduced copy number; UR35B.6 is indicated. (*D*) Genome region around UR35B.6 showing chromatin domains (bars at *top*) and deletions (red bars) that are the underlying cause of sequence underrepresentation in UR regions, identified from polytene DNA sequence reads (from Yarosh and Spradling 2014). Arrows show deletion boundaries used to calculate UR boundaries (Table 1), whereas asterisks show six deletions form the region that continued to the next adjacent UR region.

replication forks (Laird 1980; Nordman and Orr-Weaver 2012), but efforts to detect such structures were unsuccessful (Spierer and Spierer 1984; Glaser et al. 1992). Recently, high-throughput sequencing showed that during each endocycle, fork breakage and ligation are responsible for UR by generating deletions that covalently alter each UR region, leaving as few as 10% of their DNA strands intact (Yarosh and Spradling 2014).

The organization of metazoan genomes into discrete territories is probably an ancient and conserved property that was first revealed in the bands of giant polytene chromosomes. Only a handful of small differences in band patterns have been identified between different tissues (Hochstrasser 1987; Heino 1989; Richards 1980), suggesting that banding corresponds to a general aspect of genome organization minimally related to tissue differentiation. How bands and interbands correspond to functional genomic features has been studied in a few favorable chromosome regions (Vatolina et al. 2011; Zhimulev et al. 2014; Zielke et al. 2016). Putative interband regions are enriched in specific chromatin proteins, active histone marks, transcribed genes, replication origins, and P element insertion sites. Comparison of features such as chromatin marks and gene activity suggests that the domain structure of the *Drosophila* genome is highly similar between polytene and diploid cells (Vatolina et al. 2011; Zielke et al. 2016). Insulator proteins were localized at the junction of bands and interbands and were proposed to organize chromatin domains (Pai et al. 2004; Gerasimova et al. 2007).

Table 1. Relation of unfinished replication (UR) regions to bands using P insertions

UR or P name	*N*	L avg	SD	R avg	SD	Band	Interval
l.2.05341		595				21C7-D1	34
21D.1	12	629	10	764	39	**21D1-2**	134
l.2.01855		702				21D1-2	
l.2.04723				827		21D3-4	63
l.2.k05428		870				21D4-E1	51
21E.1 s	6	921	1	1010	14	**21E1-2**	89
l.2.k06921				1063		21E2-3	53
l.2.k00619		1159				21F1-2	70
22A.1	15	1229	24	1447	29	**22A1-2**	219
l.2.k11704				1614		22A3-4	167
l.2.k09624		1737				22B1-2	21
22B.2 s	24	1758	5	1817	2	**22B2**	59
l.2.k09932				2046		22C1-2	229
l.2.10638		2455				22F-A1-2	30
23A.1 t	14	2485	5	2704	10	**23A1-2**	219
l.2.k05909				2808		23B1-2	104
l.2.05965		3825				24C8	56
24D.1	38	3881	6	4030	16	**24D1-2**	149
l.2.k01102				4031		24D3-4	1
l.2.k08903		4390				24F1-2	165
25A.3	36	4555	6	4775	12	**25A1-2**	220
l.2.k10004				4853		25B1-2	78
l.2.03771		5327				25D4-6	66
25E.1	8	5393	13	5493	17	**25E1-2**	99
l.2.k11511				5542		25E5-6	49
l.2.k11511		5542				25E5-6	26
25F.1	15	5568	65	5713	4	**251-2**	145
l.2.k06502				5725		25F3-4	12
l.2.10642		6083				26B8-9	56
26C.1	25	6139	4	6304	8	**26C1-2**	164
l.2.k13720				6324		26C2-3	20
l.2.k06704		7040				27D1-2	248
27F.1	7	7288	3	7364	44	**27F1-2**	77
l.2.02657		7037				27F1-2	
l.2.k10113				7424		27F4-6	60
l.2.k10113		7424				27F4-6	192
28C.1 st	5	7616	23	7702	46	**28C1-2**	86
l.2.rL220				7810		28C4-6	108
l.2.03424		8528				29D4-5	6
29E.1	7	8534		8660		**29E1-2**	126
l.2.k13702		8544				29E1-2	
l.2.k04003				8687		29E3-4	27
l.2.s2978		8989				29F8-A1	24
30A.1	18	9013	*2*	9108	12	**30A1-2**	95
l.2.k05809				9176		30A3-6	68
l.2.k10307		10,517				31F4-5	37
32A.1	28	10,554	2	10,707	7	**32A1-3**	153
l.2.k13206				10,767		32A4-5	60
l.2.k02807		11,221				32E1-2	100
32F.1	19	11,321	17	11,445	29	**32F1**	124
l.2.03602				11,446		32F1-2	1
l.2.03602		11,446				32F1-2	96
32F.3	58	11,542	6	11,778	3	**32F3-4**	235
l.2.04418				11,805		33A1-2	27
l.2.06470		11,808				33A2-3	40
33B.1 s	3	11,848	27	11,908	60	**33B1-4**	61
l.2.01810				12,028		33B8-12	120

Continued

Table 1. *Continued*

UR or P name	*N*	L avg	SD	R avg	SD	Band	Interval
l.2.08323		12,108				33D1-2	100
33E.1 s	40	12,208	12	12,318	18	**33E1**	110
l.2.k06909				12,435		33E5-7	117
l.2.k00612		12,545				33F1-2	12
33F.1	8	12,557	19	12,662	37	**33F1-2**	105
l.2.k05448				12,704		33F1-2	42
l.2.k05448		12,704				33F1-2	52
34A.1	31	12,756	6	12,948	14	**34A1-2**	192
l.2.k09035		12,822				34A1-2	
l.2.01510				12,975		34A1-2	27
l.2.06646		13,878				34E1-2	70
34F.1 s	21	13,948	16	14,071	53	**34F1**	123
l.2.k11509				14,233		34F3-4	162
l.2.k13218		14,689				35B3-5	23
35B.6	73	14,712	3	14,972	11	**35B6**	260
l.2.k08808				15,008		35B6-10	36
l.2.05441		15,111				35C1-2	36
35C.3	8	15,147	8	15,267	111	**35C3-4**	120
l.2.06430				15,271		35D1-4	4
l.2.06430		15,271				35D1-4	-4
35D.1	44	15,267	21	15,472	98	**35D1-2**	204
13 inserts		15,333		15,338		35D1-2	
l.2.k05305				15,496		35D3-4	24
l.2.k05305		15,496				35D3-4	19
35D.3 s	84	15,515	2	15,659	7	**35D3-4**	144
l.2.05206				15,746		35D3-4	87
l.2.k09033		15,763				35D6-7	17
35E.1	52	15,780	4	15,914	52	**35E1-2**	134
l.2.k07829				16,287		35F1-2	373
l.2.k09033		15,763				35D6-7	171
35E.2	43	15,934	14	16,194	44	**35E3-4**	260
l.2.k07829				16,287		35F1-2	93
l.2.k04216		16,352				35F11-12	39
36A.1	9	16,391	11	16,467	16	**36A1**	76
l.2.k07510				16,493		36A2-3	26
1.2.k16215		16,526				36A4-5	26
36A.6	6	16,552	32	16,684	101	**36A6-7**	131
l.2.k15102				16,691		36A10-11	7
l.2.k03902		16,825				36B1-2	100
36C.1	69	16,925	3	17,349	23	**36C1-2**	425
l.2.k10816				17,450		36D1-3	101
l.2.k10816		17,450				36D1-3	61
36E.1	94	17,511	2	17,953	11	**36E1**	441
l.2.k09927				18,320		36E3-4	367
l.2.k10816		17,450				36D1-3	514
36E.2	20	17,964	4	18,173	37	**36E2**	209
l.2.k09927				18,320		36E3-4	147
l.2.k10816		17,450				36D1-3	715
36E.3 st	15	18,165	2	18,269	4	**36E3**	103
l.2.k09927				18,320		36E3-4	51
l.2.k06028		19,166				37C6-7	46
37D.1 s	10	19,212	3	19,292	22	**37D1-2**	80
l.2.01068				19,575		37F1-2	283
l.2.03552		20,085				38B4-6	19
38C.1	23	20,104	28	20,255	30	**38C1-2**	151
l.2.01820				20,243		38C1-2	
l.2.k07219				20,382		38C5-6	127

Continued

Table 1. *Continued*

UR or P name	*N*	L avg	SD	R avg	SD	Band	Interval
l.2.k07219		20,382				38C5-6	103
38D.1 s	15	20,485	5	20,608	15	**38D1-2**	123
l.2.k02501				20,639		38D1-2	31
l.2.02074		21,659				39F1-2	62
40D.1	55	21,721	78	22,105	2	**40A1-4**	383
l.2.k16406		21,828				40A1-4	
l.2.04319		21,829				40A1-4	

The table shows the mapped UR regions by name (bold), the number of associated salivary gland deletions in reads analyzed here (*N*), the average left (L) and right (R) boundaries and standard deviations (SDs) in R6 coordinates (kb) usually determined from four deletion end points, and the associated band (bold) from this study (band). An "s" following the name indicates a UR that may be smaller than its associated band, whereas a "t" indicates that less than half of Hi-C studies predicted a similar topologically associated domain. Also shown are relevant P element insertions (P name), their R6 insertion site in L or R, and their cytogenetic position assigned by Laverty (band). Insertions in bold are internal to the UR. The size of the UR (Interval) or the distance (Interval) between the L and R flanking insertions and the start or end of the UR are shown.

This emerging picture of how metazoan genomes are arranged into specific territories has received strong support from studies using chromosome conformation capture methods including Hi-C (see Ghirlando and Felsenfeld 2016; Rowley and Corces 2016). Regional units with a greater probability of interaction known as topologically associated domains (TADs) represent a common feature of animal cell chromatin, including *Drosophila* (Hou et al. 2012; Sexton et al. 2012; Eagen et al. 2015; Ullanov et al. 2016). TADs were nearly identical in polytene and diploid DNA, confirming that polytene chromosomes are good models of diploid genome organization (Eagen et al. 2015) and suggesting that TADs are bands (Eagen et al. 2015; Ullanov et al. 2016). Subsequent studies at higher resolution have revealed up to several thousand TAD boundaries (Eagen et al. 2017; Hug et al. 2017; Ramirez et al. 2017; Stadler et al. 2017). The idea that TADs correspond generally to bands remains attractive, but the boundaries identified by different groups often differ, and some recent studies may be resolving structures smaller than bands.

MAPPING UR REGIONS AND POLYTENE BANDS USING HIGH-RESOLUTION IN SITU HYBRIDIZATION

A major limitation of understanding bands has been the difficulty of mapping them precisely onto the genome sequence (see Zielke et al. 2016). Most in situ hybridization mapping as summarized in FlyBase (Marygold et al. 2016) has been performed at a resolution far lower than that of individual bands and with sometimes-contradictory results due to the difficulty of high-resolution polytene mapping. To circumvent this problem, we investigated the cytogenetic location of UR domains, bands, and TADs using a collection of P element insertions mapped by in situ hybridization with exceptional accuracy and consistency. The ~1100 insertions in the collection were all localized in situ by Todd Laverty using polytene chromosomes stretched to provide single-band resolution and photographically recorded as part of the Drosophila Genome Project (Laverty and Rubin 2000) before their insertion points on the genome sequence or relevance to this project were known.

UR DOMAINS CORRESPOND TO MAJOR BANDS

Before mapping, we reexamined the number and exact location of UR regions determined previously using copy-number data (Yarosh and Spradling 2014). UR region boundaries are imperfectly defined using copy-number profiles, because boundaries represent the point where the UR signal goes to zero (Fig. 1C). We reasoned that UR boundaries could be measured more accurately using the deletions that give rise to UR (Fig. 1D). Although deletion end points are not confined to the edges of UR regions, frequently a series of similar breakpoints are located near UR boundaries (Fig. 1D, arrows). Consequently, using a collection of 4459 deletions identified from polytene larval salivary gland sequence reads essentially as described (Yarosh and Spradling 2014), we averaged the end coordinates of boundary-associated deletions for each UR region's beginning and end point and determined their average values. We found that using clusters of polytene-specific deletions as the criterion for identifying a UR region improved sensitivity and increased the total number of identified URs from 115 (Yarosh and Spradling 2014) to 203. We then compared these end points with the most closely flanking and any internal P element insertions from the Laverty collection.

The results for chromosome 2L (Table 1) show that 26 of 37 UR regions mappable by in situ hybridization correspond closely to individual, generally large, dense bands (usually denoted as doublets by Bridges). For example (Fig. 2A), the 164-kb-long UR26C.1 is flanked 56 kb on the left by insertion l(2)10642 at 26B8-9. Fifty-six kilobases is a reasonable amount of DNA to encompass the remaining sequences in region 26B, suggesting that the UR corresponds to the dark band 26C1-2. Just 20 kb beyond the UR lies insertion l(2)k13720 which was mapped just past 26C1-2, at "26C2-3." Although the precise start and end points of the UR cannot be determined from two flanking sites, the great majority of the 164-kb

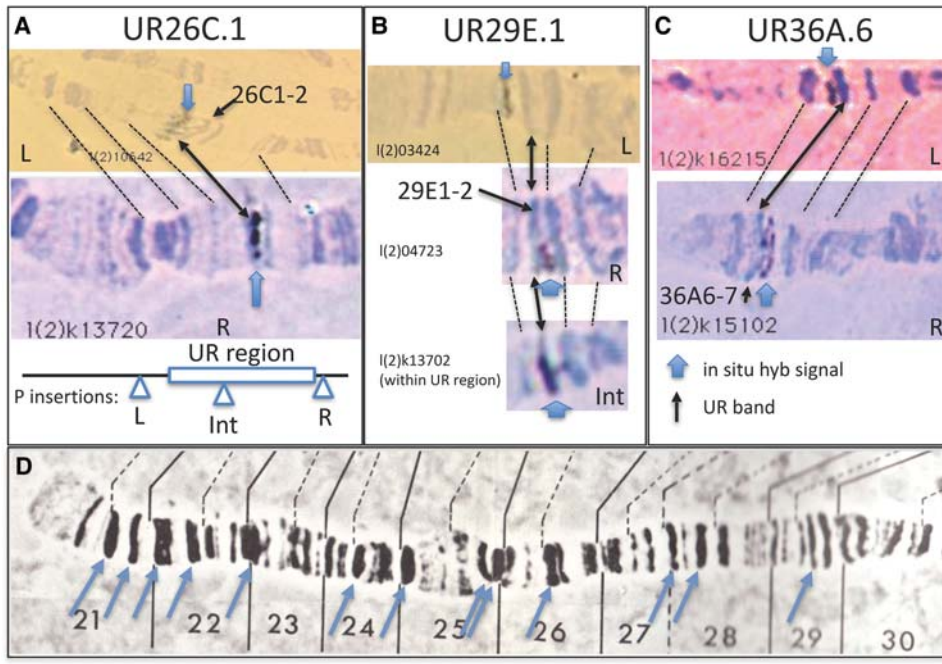

Figure 2. UR regions correspond to single dense bands. (*A*) Mapping UR26C.1 to band 26C1-2 using in situ hybridizations of the indicated flanking P element insertions (as diagrammed below) lying closest to the left (L) and (R) boundaries (Table 1). Dashed arrows indicate equivalent bands in the two panels, whereas the heavy double arrow points to band 25C1-2 that is flanked by in situ hybridization signals (blue arrows). (*B*) Mapping of UR29E.1 to band 29E1-2 as in *A*. This UR containing an internal (int) P insertion, l(2)k13702, whose localization to band 29E1-2 is shown. (*C*) Mapping of UR36A.6 to band 36A6-7. (*D*) A summary photograph of the distal half of chromosome 2L with arrows showing the strong bands mapped here to UR regions. In situ hybridizations are from the Laverty collection (Laverty and Rubin 2000).

UR must reside in the dark 26C1-2 band. It is plausible to assume that replication forks proceeding from outside (Sher et al. 2012) usually stall at or just inside 26C1-2 boundaries to generate the observed deletions and UR profile. Likewise, UR29E.1 (Fig. 2B) is flanked by even closer insertions that link it to the band 29E1-2, and UR36A.6 (Fig. 2C) maps to band 36A6-7. When a UR, such as 29E.1 (Fig. 2B), also has an internal insertion (l(2)01855), it invariably maps to the expected large band, whereas the flanking insertions map to either side of that band (Table 1).

The other 11 URs show a similar pattern with insertions flanking a relatively large band positioned to contain the UR. However, either the UR size or its location relative to flanking insertions suggests that the UR comprises only part of the large band. For example, the 192-kb UR34A.1 corresponds to band 34A1-2 but insertion l(2)01510, located 27 kb past the end of the UR, still localized to 34A1-2. Across the genome as a whole there are also a few atypical URs that span several bands, such as those comprising the two major homeotic gene clusters on chromosome 3, ANT-C and BX-C.

As can be seen from the summary for distal chromosome arm 2L (Fig. 2D), ~33% of the large dense bands on the chromosome comprise UR regions. A substantial number of euchromatic regions defined by preferential breakage, late replication, etc., have been termed "intercalary heterochromatin" (Kaufmann 1939; Zhimulev and Belyaeva 2003) and ~25% of these correspond to UR regions (Belyakin et al. 2005). Consistent with this, the mapping presented here shows that a region's chromatin state is not a good predictor of whether it will underreplicate. The great majority of genomic Polycomb domains are not URs, and only ~12% of URs correspond to Polycomb domains, most quite weak in their effects. Black chromatin is widespread, and many dense bands with black chromatin are not URs, but the majority of UR bands contain mostly "black" chromatin. Domains enriched in H3K9me3 outside the centric regions would be expected to act like intercalary heterochromatin, and these relatively rare zones include some of the strongest URs such as 36C.1 and 36E.1. However, most URs are not enriched in H3K9me3, and some H3K9me3-rich zones, despite perhaps replicating fairly late in S, are not URs. Thus, the only consistent feature that defines a region as a UR is failing to complete replication during a significant number of endocycles.

The observation (Yarosh and Spradling 2014) that UR regions contain many large genes was further investigated by calculating the abundance of genes of various sizes in URs versus in the genome as a whole. Genes (protein-coding or all annotated genes) are progressively enriched up to fourfold as a function of size in UR regions relative to their frequency in the genome as a whole (Table 2). URs were also reported to be enriched in genes encoding IgG superfamily and other cell surface proteins in both *Drosophila* and mammals (Hannibal et al. 2014; Yarosh and Spradling 2014). The nature of the 1939 UR genes as a

Table 2. Unfinished replication (UR) enrichment versus gene size

Gene size	Total	Euch	UR genes	UR euch	UR/total	UR euch/euch
Protein-coding genes						
Any size	13,762	13,563	1404	1328	0.10	0.10
>10 kb	1948	1874	239	212	0.12	0.11
>50 k	305	277	83	73	0.27	0.26
>100 kb	86	68	30	24	0.35	0.35
>150 kb	25	17	9	7	0.36	0.41
Protein and RNA genes						
Any size	16,106	15,837	1940	1826	0.12	0.12
>10 kb	1994	1917	259	229	0.13	0.12
>50 kb	317	287	91	79	0.29	0.28
>100 kb	91	72	36	28	0.40	0.39
>150 kb	28	19	12	9	0.43	0.47

The table shows the number of protein coding (*upper*) or protein and RNA (*lower*) genes based on R6 genome annotation within the indicated size class, either throughout the genome (total), within euchromatin (euch), within UR regions as defined by Yarosh and Spradling (2014) and updated here (UR genes), or within euchromatic UR regions (UR euch). The ratios of genes in the indicated size classes within UR regions to total genes (UR/total), and the ratios of genes in euchromatic UR regions to total euchromatic genes (UR euch/euch) are shown. Enrichment of genes of a given size class in UR regions is estimated by comparing the ratios of that size class to "Any size."

group was further investigated using gene ontology (Huang et al. 2009). The results (Table 3) confirm the large enrichment in IgG superfamily genes, many of which have functions in neural development and pathfinding. Additionally, UR regions are significantly overrepresented with genes encoding membrane proteins and with genes that sense odorants and function in olfaction.

BANDS SHOWING UR USUALLY CORRESPOND TO TADs

The relationship between *Drosophila* TADs and bands was similarly analyzed using the Laverty collection and the TAD boundary values reported in recent publications. URs defined by deletion end points (Table 1) were compared with TADs reported in six recent studies (Fig. 3). In general, there were many differences between the TAD boundaries identified in these studies, which were performed over a span of 5 years and differed substantially in resolution. However, among the subclass of large bands showing UR, there was usually close correspondence to a TAD "consensus." For example, band 26C1, which comprises much or all of UR26C.1, matches approximately to a single large TAD in five of the six studies. Four studies agree that the left boundary of this TAD is near 6100 kb, 39 kb from the measured left UR boundary but still 17 kb to the right of the flanking l(2)10642 P element at 26B8-9. The right boundaries of four TAD measurements are all about 10–15 kb to the right of the UR, whereas two are past the flanking element l(2)k13720 at 26C2-3. Thus, a TAD similar to band 26C1-2 and UR26C.1 probably exists, but the UR may be slightly smaller and the TAD may comprise or be slightly larger than this band. It is easy to

Table 3. Gene ontology (GO) analysis of UR region genes

Category	Term	Count	P_{Value}	Benjamini
Annotation Cluster 1	Enrichment Score: 10.69			
INTERPRO	IPR007110:Immunoglobulin-like domain	50	5.90E−19	5.85E−16
INTERPRO	IPR013783:Immunoglobulin-like fold	52	8.30E−15	2.75E−12
Annotation Cluster 2	Enrichment Score: 6.69			
GOTERM_CC_DIRECT	GO:0005956~protein kinase CK2 complex	14	1.52E−10	4.79E−08
GOTERM_BP_DIRECT	GO:0080163~regulation of protein serine/threonine phosphatase activity	13	4.81E−10	6.36E−07
Annotation Cluster 3	Enrichment Score: 4.8			
GOTERM_BP_DIRECT	GO:0007606~sensory perception of chemical stimulus	26	1.52E−08	6.71E−06
GOTERM_MF_DIRECT	GO:0005549~odorant binding	27	4.39E−07	1.17E−04
Annotation Cluster 4	Enrichment Score: 4.1			
INTERPRO	IPR003591:Leucine-rich repeat, typical subtype	21	9.10E−06	7.51E−04
Annotation Cluster 5	Enrichment Score: 3.90			
GOTERM_CC_DIRECT	GO:0016021~integral component of membrane	305	4.66E−08	4.88E−06
Annotation Cluster 7	Enrichment Score: 2.59			
GOTERM_MF_DIRECT	GO:0005549~odorant binding	27	4.39E−07	1.17E−04
GOTERM_MF_DIRECT	GO:0004984~olfactory receptor activity	25	1.63E−05	2.16E−02
Annotation Cluster 8	Enrichment Score: 2.48			
UP_SEQ_FEATURE	DNA-binding region:Homeobox	16	1.24E−06	7.48E−04
Annotation Cluster 10	Enrichment Score: 2.30			
GOTERM_BP_DIRECT	GO:0050896~response to stimulus	8	3.45E−04	5.54E−02
INTERPRO	IPR006170:Pheromone/odorant binding protein	12	1.44E−03	5.35E−02

The 1404 protein-coding genes located with UR regions as defined by Yarosh and Spradling (2014) and updated here were subjected to GO analysis using the National Institutes of Health (NIH) DAVID website (Huang et al. 2009) and the release 6 annotation of the *Drosophila* genome. For brevity, some redundant or less relevant matches are not shown.

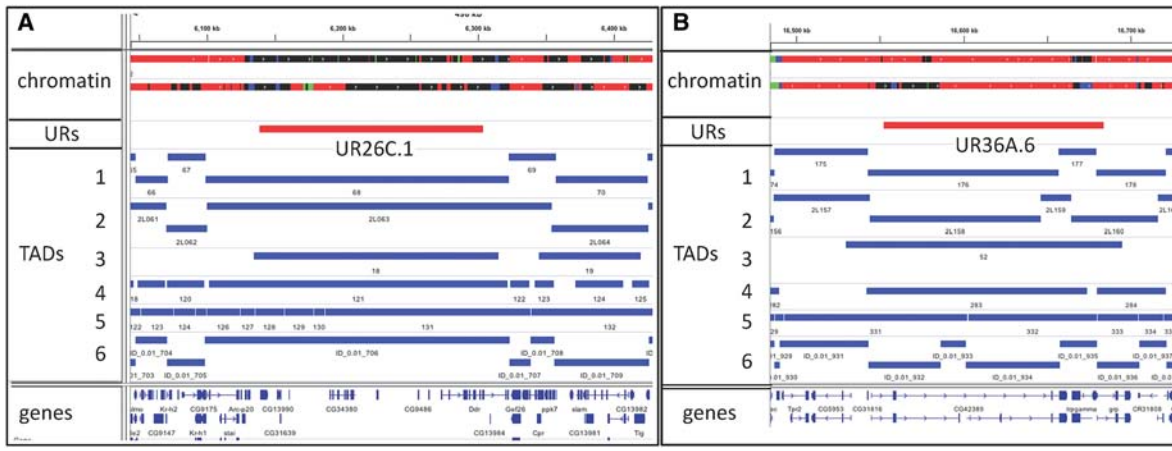

Figure 3. Correspondence between UR bands and topologically associated domains (TADs). (*A*) The 2L chromosome region surrounding UR26C.1, followed by two tracks showing chromatin types from S2 and BG3 tissue culture cells colored as in Filion et al. (2010). Below, the position of UR regions ("URs") mapped in Table 1 are shown in red. TADs mapped to this region by the indicated publications (1–6) are mapped below (blue). (1) Sexton et al. (2012); (2) Hou et al. (2012); (3) Eagen et al. (2015); (4) Eagen et al. (2017); (5) Stadler et al. (2017); (6) Ramirez et al. (2017). Genes in the region are also plotted. In five of six studies, UR26C.1 approximates one TAD. (*B*) The same tracks as in *A* are plotted for the genomic region surrounding UR36A.6. In this case, the UR is reported to contain from one to four TADs.

understand why a UR would be slightly smaller than a band, because replication forks might not stall instantly after encountering the band edge, whereas the slightly larger size of a TAD measurement might be due to insufficient Hi-C resolution. In general, the agreement between the UR, band, and TAD is quite striking.

Similar approximate agreement with a TAD was observed in at least 21 of the other 38 UR regions in Table 1. In many of these, the UR was consistently smaller than the TAD, especially in the case of URs that were judged to comprise only part of a large band. Nonetheless, it was also common to have URs split into two or more TADs in at least some of the reported data. For example, UR36A.6 (Fig. 3B) comprises band 36A6-7 and is unusual in being weak (six deletions) and in corresponding to a largely red ("active") chromatin domain. Although two studies identified TADs in this region with similar dimensions, the others split UR36A.6 into two to four TADs. Thus, it appears that UR regions associated with single strong bands frequently do correspond to TADs, but that high-resolution Hi-C may be required to precisely map their end points. However, Hi-C data sometimes go farther and break a UR and its corresponding band into subregions. The biological meaning of such subdivision is currently unclear.

TISSUE COMPARISON OF URs AND TADs

We compared the relationship of UR regions and bands between tissues by analyzing DNA from the *Drosophila* midgut, which contains 8C polytene enterocytes as a major cell type. After sequencing to a depth of 272 million reads, the midgut read depth profile showed clear underreplication (Fig. 4A). Moreover, the size and location of URs appeared to be about the same in salivary gland and midgut. However, midgut UR regions were less underreplicated in general than salivary gland URs, and individual URs varied substantially in their relative levels of UR between the two tissues. For example, UR34A.1, UR34F.1, UR35D.1, and UR35D.3 were much more fully replicated in midgut, suggesting that origin usage or timing differs in these regions between the tissues (Fig. 4A). We identified 1729 unique reads that define deletions between 10 and 500 kb from the midgut data, observed that they preferentially mapped to UR regions, and used the deletion end points to calculate the precise boundaries of 14 midgut UR regions on chromosome 2L (Fig. 4B). The values obtained in the case of all 14 midgut UR zones were the same within measurement error as the corresponding coordinates in salivary gland DNA. For example, the right boundary of UR35B.6 of 14970 ± 11.43 kb in salivary gland compared closely with 14920 ± 19.79 kb in midgut. Thus, consistent with the early appearance of a fixed pattern of TADs during development (Hug et al. 2017) and the invariant pattern of bands and TADs between tissues, the dimensions of URs are the same in midgut and salivary gland, reflecting a fundamental aspect of *Drosophila* genome structure. However, the level of underreplication does vary in particular UR zones, most likely because of differences in origin usage and timing between these tissues.

CONCLUSION

Bands Are Fundamental Genomic Units That May Coordinate Replication and Transcription

Our results help clarify the nature of the genomic domains that are visualized in polytene cells as chromosomal bands. We showed that many of the strongest bands correspond almost exactly to UR regions, implicating bands in replication control, not just in gene regulation. These bands also largely match major TADs determined by Hi-C

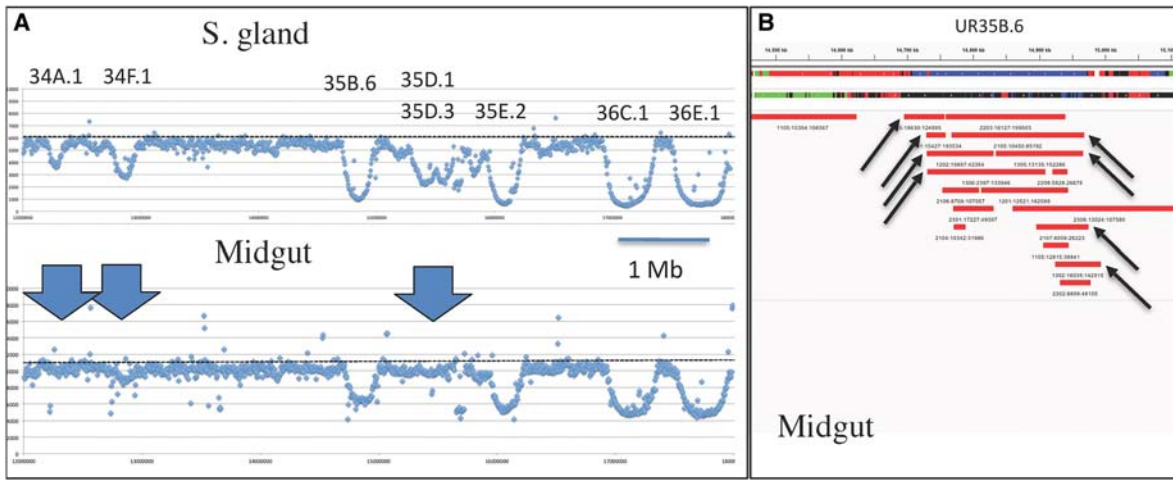

Figure 4. Precise correspondence in coordinates but not in depth of UR regions in larval salivary gland (S. gland) and adult midgut DNA. (*A*) Plot of read depth in 5-kb bins from region 34–36 of chromosome 2L from larval salivary gland (*upper*) and adult midgut (*lower*) DNA. Major UR regions show decreased read depth and are labeled. Blue arrows indicate four UR regions that are greatly diminished in midgut compared with salivary gland. (*B*) Plot of region surrounding UR35B.6 (see Fig. 1D), showing chromatin from BG3 and S2 cells (colored according to Filion et al. 2010, but H3K27me3-rich is green and H3K9me3-rich is blue), and deletions determined from midgut sequence reads. Arrows indicate the deletions used to calculate the *left* and *right* end coordinates of the midgut UR35B.6, which proved indistinguishable from salivary gland UR35B.6 end points (see text).

studies. Neither band structure nor TAD organization varies significantly between different tissues, and we showed that UR domains are also precisely the same between different cell types. Bands and TADs likely represent a sequence-encoded aspect of metazoan genome structure.

We propose that these fixed genomic domains serve to organize the temporal program of replication (Gilbert et al. 2010; Pope et al. 2014) in a manner that promotes appropriate gene expression. It is well-known from studies of unicellular organisms that DNA replication and transcription have the potential to significantly interfere with each other if not subject to regulation (see Merrikh et al. 2012). Highly transcribed genes including rRNA genes in bacteria are usually positioned so that replication forks will travel in the same direction as transcription to minimize interference, rather than in the opposite direction. In metazoans there is also a strong tendency for highly expressed genes (i.e., genes whose products are needed in large amounts) to be replicated early in S phase. Early replication is expected to increase product production by generating transcripts from a second template for as long as possible. The organization of the genome into domains, coupled with regulated origin activation, may help ensure these outcomes.

Bands May Promote Favorable Replication Timing for Genes Transcribed at All Levels

Whether replication timing is important for genes transcribed at a moderate or low rate has remained unclear. However, there are several ways in which replication timing might assist such genes independent of enhancing gene dosage. Interactions between enhancers and promoters within a TAD that involve the formation of looped contacts would likely be susceptible to disruption by replication fork passage. Afterward, gene expression would remain off for however long was required to restore these interactions and complete RNA polymerase passage. Genomic domains controlling replication timing might ensure that their component genes shut down at a time in the cell cycle that minimizes the impact on developmental and physiological events in which they are participating.

There may even be a currently unrecognized class of "slow-activating" genes that would benefit from a program of late replication. These might include genes with extremely complex regulation involving large transcription units and multiple enhancers, such that product production can only begin after a long delay while appropriate regulatory structures are established and the length of the gene is traversed by slowly moving polymerases. If slow-activating genes were to replicate early in S phase, this intricate regulatory organization along with nascent transcripts would be disrupted by fork passage quite soon after product production had finally begun. Only after a second lag of equal length would such genes be able to resume transcription, perhaps leaving insufficient time in the cell cycle for adequate production. We propose that such slow-activating genes would benefit from replicating late in the cell cycle, as this would ensure that they could produce product with only one major pause per cell cycle for regulatory assembly and initial transcription. These considerations might explain why the genome would evolve tissue-invariant late-replicating bands. Furthermore, these domains might contain some of the largest, most complex, and highly regulated genes in the genome.

Our studies suggest that UR domains house genes that could in this way benefit from late replication. URs are enriched for large genes and encode genes with apparently complex regulation in embryonic and neural development.

For example, the largest and most complex homeotic gene clusters, ANT-C and BX-C, are located in UR regions. Multiple very large genes in URs encode IgG superfamily proteins, at least some of which are required for neuronal pathfinding and other complex aspects of nervous system function. The fact that transcripts from most genes within URs have not been detected does not mean that these genes are unimportant. They may be expressed in understudied tissues such as polyploid neurons and glia or produce heterogeneous products that differ between cells.

Polyteny May Further Diversify Gene Regulation

The idea that bands are part of a system to time DNA replication and that late replicating bands house a subclass of genes with complex regulation provides a speculative but interesting rationale for the evolution of polyteny and underreplication after the origin of Diptera in the early Mesozoic Era. Extant cytological surveys indicate that polyploidy, but not polyteny, is widespread throughout the phylogenetic tree and in particular within many insect orders. Non-Dipteran polyploid cells, to the extent known, reset the endocycle near or within M phase and fully replicate their satellite DNA (Gage 1974). In contrast, polytene chromosomes are essentially confined among insects to the Diptera, one of the higher insect orders that arose in the Triassic (Grimaldi and Engel 2005), long after insect body plans and general physiology had been established. Polytene cells in most Dipterans probably underreplicate centromeric heterochromatin, based on their chromosome morphology (White 1973). Thus, the evolution of flies was associated with alterations in the endocycle that produced polyteny and centromeric UR. These changes were not needed to make a generic insect but are likely to have been part of the evolutionary innovation that put two-winged flies on the path to success across the globe.

Compartmentalized genome territories are widespread in metazoans and were presumably in place long before Dipteran evolution. We propose that the advent of polyteny and centromeric UR provided new flexibility in gene expression not available in other insect groups. By bringing multiple aligned copies of the genome close together in a polytene chromosome, opportunities for interstrand enhancer–promoter interactions via looping would be greatly increased. Novel regulatory interactions might be further multiplied by bringing homologs into proximity, which might also explain the origin of somatic pairing in both polytene and diploid cells, another characteristic of Diptera. Pairing-dependent interactions have been widely documented in *Drosophila* genetics and can frequently be explained by cross-strand enhancer action (Lewis 1954; Gelbart and Wu 1982; Bingham and Zachar 1985; Lee and Wu 2006; Mellert and Truman 2012).

Somatic Genome Instability May Have Contributed to Dipteran Evolution

Whether somatic genome instability within euchromatic UR regions evolved at the same time as polyteny within early Dipteran groups or had a more recent origin during the expansion of the higher Diptera after the Cretaceous (Wiegmann et al. 2011) is currently unknown. Very few species have been tested for euchromatic UR, and no reliable correlate of this process in polytene chromosome morphology is currently available. A late origin is suggested by the limited phylogenetic distribution of clear SUUR homologs only within the genus *Drosophila* and a few other higher flies.

Despite specialized mouthparts limiting them to liquid food, Diptera have adapted with unprecedented success (Grimaldi and Engel 2005). Euchromatic UR might have contributed to their adaptability in concert with polyteny by further enhancing regulatory flexibility. UR-generated chromosomal deletions may alter gene regulation and protein structure by juxtaposing novel genomic regions. In addition, UR deletions are expected to perturb the local alignment of sister strands in a manner that would further expand the range of possible enhancer–promoter interactions. Somatic instability within euchromatic UR regions may increase the ability of flies to adapt to novel environments because these regions preferentially contain large genes with complex regulation that affect neural function and behavior, as indicated by GO analysis (Table 3). The diversification of UR region genes involved in "sensory perception of chemical stimulus," "odorant binding," and "olfactory receptor activity" may have helped flies to identify and exploit rare and transient food sources. It may have allowed them to adapt their sensory perception and behavior rapidly, contributing to their success in dominating diverse ecological niches. SUUR mutants might survive in laboratory conditions, but have difficulty finding dispersed food resources in wild environments. Thus, a deeper understanding of bands, polyteny, and UR may help us learn how flies have been able to make an outsized impact on Earth and on human health.

Polytene Chromosomes Remain Valuable for Understanding Chromosome and Nuclear Biology

In conclusion, polytene chromosomes have contributed significantly to our understanding of chromosome organization and genome function for the last 80 years. Today, however, the opinion has grown that polytene cells differ greatly from mammalian cells, and that giant chromosomes have been surpassed as tools by chromosome conformation capture and other genomic techniques. This study has emphasized that polytene chromosomes arose primarily as a cell cycle change, not a developmental change, and that strong experimental evidence shows they differ little in structure and function from diploid chromosomes. These natural quasicrystals of unit chromatin strands provide a high resolution view of structural and functional processes that function in nuclei throughout the animal kingdom. The ability to directly visualize the fundamental mechanisms of cell nuclei, in conjunction with molecular and genomic techniques, will continue to reward those who choose to use polytene cells to investigate the many outstanding questions in chromosome biology.

ACKNOWLEDGMENTS

The author is especially grateful to Will Yarosh and Don Fox, former members of the Spradling laboratory, for DNA sequencing of *Drosophila* tissues and to Shelly Paterno and Mike Buszczak for the protein trap photo in Figure 1A. He thanks current laboratory members Chenhui Wang, Steve Deluca, and Bob Levis for comments on the manuscript. He is particularly grateful to Todd Laverty and Gerry Rubin for sharing data from the Laverty P element in situ hybridization collection. Allison Pinder provided expert assistance with DNA sequencing and Fred Tan gave valuable advice on sequence analysis.

REFERENCES

Ashburner M. 1970. Function and structure of polytene chromosomes during insect development. *Adv Insect Physiol* **7:** 1–95.

Belyaeva ES, Zhimulev IF, Volkova EI, Alekseyenko AA, Moshkin YM, Koryakov DE. 1998. *Su(UR)ES*: A gene suppressing DNA underreplication in intercalary and pericentric heterochromatin of *Drosophila melanogaster* polytene chromosomes. *Proc Natl Acad Sci* **95:** 7532–7537.

Belyakin SN, Christophides GK, Alekseyenko AA, Kriventseva EV, Belyaeva ES, Nanayev RA, Makunin IV, Kafatos FC, Zhimulev IF. 2005. Genomic analysis of *Drosophila* chromosome underreplication reveals a link between replication control and transcriptional territories. *Proc Natl Acad Sci* **102:** 8269–8274.

Bingham PM, Zachar Z. 1985. Evidence that two mutations, wDZL and z1, affecting synapsis-dependent genetic behavior of white are transcriptional regulatory mutations. *Cell* **40:** 819–825.

Bridges CB. 1935. Salivary chromosome maps: With a key to the banding of the chromosomes of *Drosophila melanogaster*. *J Hered* **26:** 60–64.

Buszczak M, Paterno S, Lighthouse D, Bachman J, Plank J, Owen S, Skora A, Nystul T, Ohlstein B, Allen A, et al. 2007. The Carnegie protein trap library: A versatile tool for *Drosophila* developmental studies. *Genetics* **175:** 1505–1531.

Duncan AW, Taylor MH, Hickey RD, Hanlon-Newell AE, Lenzi ML, Olson SB, Finegold MJ, Grompe M. 2010. The ploidy conveyor of mature hepatocytes as a source of genetic variation. *Nature* **467:** 707–710.

Eagen KP, Hartl TA, Kornberg RD. 2015. Stable chromosome condensation revealed by chromosome conformation capture. *Cell* **163:** 934–946.

Eagen KP, Aiden EL, Kornberg RD. 2017. Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. *Proc Natl Acad Sci* **114:** 8764–8769.

Edgar BA, Zielke N, Gutierrez C. 2014. Endocycles: A recurrent evolutionary innovation for post-mitotic cell growth. *Nat Rev Mol Cell Biol* **15:** 197–210.

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **143:** 212–224.

Fox D, Gall JG, Spradling AC. 2010. Error-prone polyploid mitosis during normal *Drosophila* development. *Genes Dev* **24:** 2294–2302.

Fujiwara T, Bandi M, Nitta M, Ivanova EV, Bronson RT, Pellman D. 2005. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437:** 1043–1047.

Gage LP. 1974. Polyploidization of the silk gland of *Bombyx mori*. *J Mol Biol* **86:** 97–108.

Gall JG, Cohen EH, Polan ML. 1971. Repetitive DNA sequences in *Drosophila*. *Chromosoma* **33:** 319–344.

Gelbart WM, Wu CT. 1982. Interactions of zeste mutations with loci exhibiting transvection effects in *Drosophila melanogaster*. *Genetics* **102:** 179–189.

Gerasimova TI, Lei EP, Bushey AM, Corces VG. 2007. Coordinated control of dCTCF and gypsy chromatin insulators in *Drosophila*. *Mol Cell* **28:** 761–772.

Ghirlando R, Felsenfeld G. 2016. CTCF: Making the right connections. *Genes Dev* **30:** 881–891.

Gilbert DM, Takebayashi SI, Ryba T, Lu J, Pope BD, Wilson KA, Hiratani I. 2010. Space and time in the nucleus: Developmental control of replication timing and chromosome architecture. *Cold Spring Harb Symp Quant Biol* **75:** 143–153.

Glaser RL, Karpen GH, Spradling AC. 1992. Replication forks are not found in a *Drosophila* mini-chromosome demonstrating a gradient of polytenization. *Chromosoma* **102:** 15–19.

Grimaldi D, Engel MS. 2005. *Evolution of the insects*. Cambridge University Press, Cambridge.

Hammond MP, Laird CD. 1985. Chromosome structure and DNA replication in nurse and follicle cells of *Drosophila melanogaster*. *Chromosoma* **91:** 267–278.

Hannibal RL, Chuong EB, Rivera-Mulia JC, Gilbert DM, Valouev A, Baker JC. 2014. Copy number variation is a fundamental aspect of the placental genome. *PLoS Genet* **10:** e1004290.

Heino TI. 1989. Polytene chromosomes from ovarian pseudonurse cells of the *Drosophila melanogaster otu* mutant. *Chromosoma* **97:** 363–373.

Hochstrasser M. 1987. Chromosome structure in four wild-type polytene tissues of *Drosophila melanogaster*. The 87A and 87C heat shock loci are induced unequally in the midgut in a manner dependent on growth temperature. *Chromosoma* **95:** 197–208.

Hou C, Li L, Qin ZI, Corces VG. 2012. Gene density, transcription and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol Cell* **48:** 471–484.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57.

Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. 2017. Chromatin architecture emerges during zygotic genome activation independent of transcription. *Cell* **169:** 216–228.e19.

Karpen GH, Spradling AC. 1990. Reduced DNA polytenization of a minichromosome region undergoing position-effect variegation in *Drosophila*. *Cell* **63:** 97–107.

Kaufmann BP. 1939. Distribution of induced breaks along the X-chromosome of *Drosophila melanogaster*. *Proc Natl Acad Sci* **25:** 571–577.

Laird CD. 1980. Structural paradox of polytene chromosomes. *Cell* **22:** 869–874.

Laverty T, Rubin GM. 2000. Encyclopedia of *Drosophila*, vol. 2. CD distributed by the Berkeley *Drosophila* Genome Project, http://www.fruitfly.org/index.html.

Lee AM, Wu CT. 2006. Enhancer-promoter communication at the *yellow* gene of *Drosophila melanogaster*: Diverse promoters participate in and regulate trans interactions. *Genetics* **174:** 1867–1880.

Lefevre G Jr. 1976. A photographic representation and interpretation of the polytene chromosomes of *Drosophila melanogaster* salivary glands. In *The genetics and biology of* Drosophila, Vol. 1a (ed. Ashburner M, Novitski E), pp. 32–64. Academic, New York.

Lewis EB. 1954. The theory and application of a new method of detecting chromosomal rearrangements in *Drosophila melanogaster*. *Am Nat* **88:** 225–239.

Losick VP, Fox DT, Spradling AC. 2013. Polyploidization and cell fusion contribute to wound healing in the adult *Drosophila* epithelium. *Curr Biol* **23:** 2224–2232.

Losick VP, Jun AS, Spradling AC. 2016. Wound-induced polyploidization: Regulation by hippo and JNK signaling and conservation in mammals. *PLoS ONE* **11:** e0151251.

Marygold SJ, Crosby MA, Goodman JL; FlyBase Consortium. 2016. Using FlyBase, a database of *Drosophila* genes and genomes. *Methods Mol Biol* **1478:** 1–31.

Mathog E, Hochstrasser M, Gruenbaum Y, Saumweber H, Sedat J. 1984. Characteristic folding pattern of polytene chromosomes in *Drosophila* salivary gland nuclei. *Nature* **308:** 414–421.

Mellert DJ, Truman JW. 2012. Transvection is common throughout the *Drosophila* genome. *Genetics* **191:** 1129–1141.
Merrikh H, Zhang Y, Grossman AD, Wang J. 2012. Replication-transcription conflicts in bacteria. *Nat Rev Microbiol* **10:** 449–458.
Moshkin YM, Alekseyenko AA, Semeshin VF, Spierer A, Spierer P, Makarevich GF, Balyaeva ES, Zhimulev IF. 2001. The bithorax complex of *Drosophila melanogaster*: Underreplication and morphology in polytene chromosomes. *Proc Natl Acad Sci* **98:** 570–574.
Nagl W. 1978. *Endopolyploidy and polyteny in differentiation and evolution*. Elsevier Science, Amsterdam.
Neiman M, Beaton MJ, Hessen DO, Jeyasingh PD, Weider LJ. 2017. Endopolyploidy as a potential driver of animal ecology and evolution. *Biol Rev* **92:** 234–247.
Nordman J, Orr-Weaver TL. 2012. Regulation of DNA replication during development. *Development* **139:** 455–464.
Nordman J, Li S, Eng T, Macalpine D, Orr-Weaver TL. 2011. Developmental control of the DNA replication and transcription programs. *Genome Res* **21:** 175–181.
Nordman JT, Kozhevnikova EN, Verrijzer CP, Pindyurin AV, Andreyeva EN, Shloma VV, Zhimulev IF, Orr-Weaver TL. 2014. DNA copy-number control through inhibition of replication fork progression. *Cell Rep* **9:** 841–849.
Orr-Weaver TL. 2015. When bigger is better: The role of polyploidy in organogenesis. *Trends Genet* **31:** 307–315.
Pai CY, Lei EP, Ghosh D, Corces VG. 2004. The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol Cell* **16:** 737–748.
Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen R, Canfield TK, et al. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature* **515:** 402–405.
Ramirez F, Bhardwaj V, Villaveces J, Arrigoni L, Grüning BA, Lam KC, Habermann B, Akhtar A, Manke T. 2017. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *bioRxiv* doi: http://dx.doi.org/10.1101/115063; see also http://chorogeome.ie-freiburg.mpg.de.
Richards G. 1980. The polytene chromosomes in the fat body nuclei of *Drosophila melanogaster*. *Chromosoma* **79:** 241–250.
Rowley MJ, Corces VG. 2016. The three-dimensional genome: Principles and roles of long distance interactions. *Cell* **162:** 703–705.
Schoenfelder KP, Fox DT. 2015. The expanding implications of polyploidy. *J Cell Biol* **209:** 485–491.
Schoenfelder KP, Montague RA, Paramore SV, Lennox AL, Mahowald AP, Fox DT. 2014. Indispensable pre-mitotic endocycles promote aneuploidy in the *Drosophila* rectum. *Development* **141:** 3551–3560.
Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148:** 458–472.
Sher N, Bell GW, Li S, Nordman J, Eng T, Eaton ML, MacAlpine DM, Orr-Weaver TL. 2012. Developmental control of gene copy number by repression of replication initiation and fork progression. *Genome Res* **22:** 64–75.
Spierer A, Spierer P. 1984. Similar level of polyteny in bands and interbands of *Drosophila* giant chromosomes. *Nature* **307:** 176–178.
Spradling AC. 1993. Position effect variegation and genomic instability. *Cold Spring Harbor Symp Quant Biol* **58:** 585–596.
Stadler MR, Jaines JE, Eisen MB. 2017. Convergence of topological domain boundaries, insulators and polytene interbands revealed by high-resolution mapping of chromatin contacts in the early *Drosophila melanogaster* embryo. *bioRxiv* doi.org/10.1101/149344.
Stormo BM, Fox DT. 2017. Polyteny: Still a giant player in chromosome research. *Chromosome Res* doi: 10.1007/s10577-017-9562-z.
Ullanov SV, Khrameeva EE, Gavirlov AA, Flyamer IM, Kos P, Mikhaleva EA, Penin AA, Logacheva MD, Imakaev MV, Chertovich A, et al. 2016. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res* **26:** 70–84.
Vatolina TY, Boldyreva LV, Demakova OV, Demakov SA, Kokoza EB, Semeshin VF, Babenko VN, Goncharov FP, Balyaeva ES, Zhimulev IF. 2011. Identical functional organization of nonpolytene and polytene chromosomes in *Drosophila melanogaster*. *PLoS ONE* **6:** e25960.
White MJD. 1973. *Animal cytology and evolution*. Cambridge University Press, Cambridge.
Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim JW, Lambkin C, Bertone MA, Cassel BK, Bayless KM, Heimberg AM, et al. 2011. Episodic radiations in the fly tree of life. *Proc Natl Acad Sci* **108:** 5690–5695.
Yarosh W, Spradling A. 2014. Incomplete replication generates somatic DNA alterations within *Drosophila* polytene salivary gland cells. *Genes Dev* **28:** 1840–1855.
Zhimulev IF. 1996. Morphology and structure of polytene chromosomes. *Adv Genet* **34:** 1–490.
Zhimulev IF, Belyaeva ES. 2003. Intercalary heterochromatin and genetic silencing. *BioEssays* **25:** 1040–1051.
Zhimulev IF, Zykova TY, Goncharov FP, Khoroshkoa VA, Demakov OV, Semeshin VF, Pokholkova GV, Boldyreva LV, Demidova DS, Babenko VN, et al. 2014. Genetic organization of interphase chromosome bands and interbands in *Drosophila melanogaster*. *PLoS ONE* **9:** e101631.
Zielke T, Glotov A, Saumweber H. 2016. High-resolution in situ hybridization analysis on the chromosomal interval 61C7-61C8 of *Drosophila melanogaster* reveals interbands as open chromatin domains. *Chromosoma* **125:** 423–435.

Cold Spring Harbor Symposia on Quantitative Biology

Polytene Chromosome Structure and Somatic Genome Instability

Allan C. Spradling

Cold Spring Harb Symp Quant Biol 2017 82: 293-304 originally published online November 22, 2017
Access the most recent version at doi:10.1101/sqb.2017.82.033670