© 2018. Published by The Company of Biologists Ltd | Development (2018) 145, dev164657. doi:10.1242/dev.164657

RESEARCH ARTICLE

Distinct *cis*-acting elements mediate targeting and clustering of *Drosophila* polar granule mRNAs

Whitby V. I. Eagle, Daniel K. Yeboah-Kordieh, Matthew G. Niepielko and Elizabeth R. Gavis*

ABSTRACT

Specification and development of *Drosophila* germ cells depend on molecular determinants within the germ plasm, a specialized cytoplasmic domain at the posterior of the embryo. Localization of numerous mRNAs to the germ plasm occurs by their incorporation, as single-transcript ribonucleoprotein (RNP) particles, into complex RNP granules called polar granules. Incorporation of mRNAs into polar granules is followed by recruitment of additional like transcripts to form discrete homotypic clusters. The *cis*-acting localization signals that target mRNAs to polar granules and promote homotypic clustering remain largely uncharacterized. Here, we show that the *polar granule component* (*pgc*) and *germ cell-less* (*gcl*) 3′ untranslated regions contain complex localization signals comprising multiple, independently weak and partially functionally redundant localization elements (LEs). We demonstrate that targeting of *pgc* to polar granules and self-assembly into homotypic clusters are functionally separable processes mediated by distinct classes of LEs. We identify a sequence motif shared by other polar granule mRNAs that contributes to homotypic clustering. Our results suggest that mRNA localization signal complexity may be a feature required by the targeting and self-recruitment mechanism that drives germ plasm mRNA localization.

KEY WORDS: Germ plasm, Polar granule, mRNA localization, *Drosophila*, *pgc*, *gcl*

INTRODUCTION

Subcellular mRNA localization is a prevalent mechanism for generating and maintaining the asymmetric distributions of proteins necessary for cellular and developmental polarity. In many organisms, including *Drosophila*, mRNA localization plays an essential and conserved role in germline specification (Houston, 2013; Schisa, 2012). In *Drosophila*, kinesin-mediated transport and localized translation of *oskar* (*osk*) mRNA initiates the formation of a highly specialized cytoplasm at the posterior of the oocyte – the germ plasm – that is both necessary and sufficient to induce formation of the primordial germ cells (pole cells) in the early embryo (Mahowald, 2001). Osk recruits proteins, including Vasa and Tudor, to assemble ribonucleoprotein (RNP) granules called polar granules (Lehmann, 2016). Numerous maternally synthesized mRNAs subsequently become localized to the germ plasm through their entrapment in polar granules. Several of these, including *nanos* (*nos*), *polar granule component* (*pgc*) and *germ cell-less* (*gcl*), have known roles in germline development (Hanyu-Nakamura et al., 2008; Jongens et al., 1994; Kobayashi et al., 1996).

These transcripts are produced by the ovarian nurse cells and are delivered to the oocyte *en masse* through a concerted contraction of the nurse cells. Initially, *nos*, *gcl* and *pgc* reside in separate, single-transcript RNP particles that disperse throughout the oocyte via diffusion facilitated by microtubule-mediated streaming of the oocyte cytoplasm (Forrest and Gavis, 2003; Little et al., 2015; Trcek et al., 2015). Localization begins as single-transcript RNPs become incorporated into developing polar granules. The first transcripts to be so incorporated serve as seeds that recruit additional like transcripts from the bulk cytoplasm, resulting in homotypic clusters. New seeding events and cluster growth continue into early embryogenesis (Niepielko et al., 2018). Because the different mRNAs are incorporated stochastically, polar granules are heterogeneous with regard to the assortment of mRNAs they contain (Little et al., 2015). Structured illumination microscopy (SIM) revealed that homotypic clusters of different mRNAs remain spatially distinct within the granule (Niepielko et al., 2018; Trcek et al., 2015). Ultimately, the polar granules become associated with embryonic nuclei that enter the germ plasm and accompany these nuclei as they bud from the posterior cortex to form pole cells (Lerit and Gavis, 2011). Co-packaging in polar granules facilitates the pole cell inheritance of mRNAs that encode proteins essential to germline development, viability and function.

mRNA localization is directed by *cis*-acting localization signals frequently found in the 3′ untranslated regions (3′UTRs) of transcripts. Sequences or structures within these signals are specifically recognized and bound by proteins that package the transcript into a localization competent, and often translationally repressed, RNP particle (Eliscovich and Singer, 2017). For those polar granule mRNAs characterized to date, the 3′UTR is both necessary and sufficient to direct localization (Gavis and Lehmann, 1992; Rangan et al., 2009). Despite their common destination, no similarity has been found among 3′UTRs of the various *Drosophila* germ plasm mRNAs (Gavis et al., 1996; Gavis and Lehmann, 1992; Jain and Gavis, 2008; Lécuyer et al., 2007). Deletion analysis of the *nos* 3′UTR showed that multiple localization elements (LEs) contribute to localization and wild-type germ plasm mRNA localization can be achieved by different combinations of these LEs (Gavis et al., 1996). *nos* LEs are thus distributed and display partial functional redundancy. Some *nos* LEs also exhibit additive or synergistic effects, with a specific 41 nucleotide sequence capable of generating substantial localization when present in three copies (Bergsten et al., 2001; Gavis et al., 1996; Gavis and Lehmann, 1992). The significance of such organizational complexity and functional redundancy, and whether it is common among *Drosophila* germ plasm transcript localization signals remains unknown. In *Xenopus*, germ plasm localization of *nanos1* (also called *Xcat2*) mRNA requires two distinct LEs: a general element that direct RNAs to the mitochondrial cloud in a diffusion

Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA.

*Author for correspondence (gavis@princeton.edu)

E.R.G., 0000-0003-0251-0760

Received 16 February 2018; Accepted 18 September 2018

Development (2018) 145, dev164657. doi:10.1242/dev.164657

and entrapment process (Chang et al., 2004; Zhou and King, 1996); and a specialized element that targets *nanos1* to germ granules within the cloud (Kloc et al., 2000). Whether different LEs can mediate unique aspects of germ plasm mRNA localization in *Drosophila* has yet to be elucidated.

To address these gaps, we used transgenic reporters, single molecule fluorescence *in situ* hybridization (smFISH), and quantitative image analysis to identify *cis*-acting elements that direct localization of two highly enriched and functionally important polar granule transcripts: *gcl* and *pgc*. These transcripts encode proteins required, respectively, for pole cell formation and specification of germline fate (Hanyu-Nakamura et al., 2008; Jongens et al., 1992; Lerit et al., 2017; Pae et al., 2017; Timinszky et al., 2008). Whereas previous studies of germ plasm localization signals were limited to qualitative assays (Gavis and Lehmann, 1992; Gavis et al., 1996; Rangan et al., 2009), we have now analyzed mRNA localization patterns quantitatively at the level of polar granules, with single-transcript resolution.

We find that the *gcl* 3′UTR contains multiple LEs that are widely dispersed, weak in isolation and partially functionally redundant. Although several regions in the *pgc* 3′UTR exhibit localization activity, we have delineated a 59-nucleotide segment of the *pgc* 3′ UTR that is responsible for approximately half of the average *pgc* content of polar granules. Within this region, we identified a 6-nucleotide motif that is also found in the *gcl* and *nos* 3′UTRs, and for each mRNA, mutation of this motif significantly reduces its accumulation within polar granules. We demonstrate that, in addition to being functionally separable processes, the targeting of *pgc* transcripts to polar granules and their self-assembly into homotypic clusters are mediated by distinct classes of LEs, with the 6-nucleotide motif dedicated to clustering. Together, our results reveal complexity in the organization of regulatory elements comprising germ plasm mRNA localization signals that serves the targeting and self-recruitment mechanism for enrichment in the germ plasm.

RESULTS

pgc and *gcl* 3′UTRs mediate localization to and accumulation within polar granules

To delineate LEs in the 3′UTRs of *gcl* and *pgc*, we used a transgenic reporter assay. 3′UTR sequences were inserted into a reporter construct containing the *nos* 5′UTR and green fluorescent protein (GFP) coding region under UAS control (Fig. 1A). All transgenes were inserted into the same chromosomal location and expression was induced with the maternal *α-tubulin-GAL4* driver. Transgenic reporter mRNAs were detected by smFISH using probes complementary to the GFP-coding sequence. Consistent with previous results (Rangan et al., 2009), the intact *gcl* and *pgc* 3′UTRs both directed robust localization to the germ plasm in late-stage oocytes (data not shown) and early embryos (Fig. 1B,C). In contrast, the unregulated *α-tubulin* 3′UTR (*tub*) did not confer detectable posterior enrichment (see below), and *gfp-tub3′UTR* RNA was evenly distributed throughout the cytoplasm (Fig. 1D).

The efficacy of germ plasm localization was quantified using two approaches. First, we used a nearest-neighbor method to measure colocalization of reporter RNAs with a native germ plasm transcript, *nos*. This value allowed us to assess whether transcripts are targeted to and incorporated into polar granules (polar granule colocalization; Fig. 1E). Second, we quantified the average number of mRNA molecules per polar granule for a given mRNA species (polar granule mRNA content; Fig. 1F). This value primarily reflects the efficacy of mRNA recruitment into homotypic clusters (Niepielko et al., 2018). As expected, the *gfp-pgc3′UTR* and *gfp-gcl3′UTR* RNAs colocalized with *nos* and accumulated within polar granules similarly to their endogenous counterparts, indicating that they faithfully recapitulate the localization process (Fig. 1; Fig. S1). To account for random overlap due to the high density of particles within the imaging volume, we also measured the frequency with which *gfp-tub3′UTR* RNA appeared colocalized with *nos*. This value (18%) is consistent with previous measurements (Little et al., 2015) and provides a threshold above which transcripts detected as

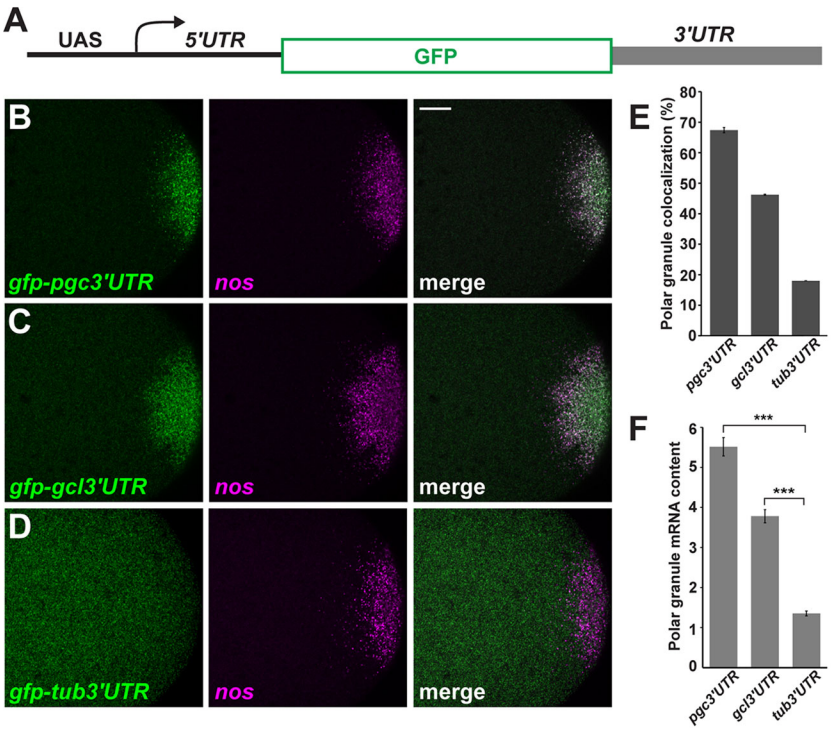

Fig. 1. *pgc* and *gcl* 3′UTRs mediate association with and accumulation in polar granules. (A) Schematic of the *UASp-gfp-3′UTR* reporter transgenes used to delineate localization elements. Transgenic reporter mRNA was detected by smFISH using probes complementary to the *gfp* sequences. (B-D) Confocal sections of the posterior region of 0-1 h old embryos (anterior towards the left, dorsal upwards). Embryos were probed simultaneously for *nos* mRNA (magenta) and reporter mRNA (green). Scale bar: 15 µm. (E) Colocalization quantification by nearest-neighbor method indicating the proportion of *nos* granules that also contain reporter mRNA (n=30 embryos each). (F) Quantification of polar granule reporter mRNA content (average number of transcripts). *gfp-pgc3′UTR* (5.5) and *gfp-pgc3′UTR* (3.8) are significantly enriched relative to *gfp-tub3′UTR* (1.3); n=10 embryos each. Values shown are mean±s.e.m.; ***P<0.001, as determined by a two-tailed t-test.

Development (2018) 145, dev164657. doi:10.1242/dev.164657

colocalized are considered to be co-packaged within the same granule (Fig. 1D,E).

Multiple individually weak localization elements are distributed throughout the *gcl* 3′UTR

Division of the *gcl* 3′UTR into two overlapping fragments, *gcl*(1-276) and *gcl*(247-525) (Fig. 2A), severely compromised its localization activity. Polar granule colocalization was observed for both the *gfp-gcl(1-276)* and *gfp-gcl(247-525)* reporter RNAs, but at greatly reduced frequency relative to *gfp-gcl3′UTR* (Fig. 2B,C). Polar granule mRNA content was also dramatically decreased, with only *gfp-gcl(1-276)* RNA accumulating to any significant extent (Fig. 2B,D; Fig. S2A-D). We therefore hypothesized that multiple sequences dispersed throughout the 3′UTR might be required for effective localization. To test this, we extended each fragment to include a larger region of the *gcl* 3′UTR [*gcl*(1-399) and *gcl*(150-525); Fig. 2A]. In each case, the extended fragments conferred significantly greater localization than their precursors, indicating that LEs reside in nucleotides 150-247 and nucleotides 276-399 (Fig. 2B-D; Fig. S2E,F). However, a reporter containing the central region (nucleotides 150-399) exhibited only weak accumulation (Fig. 2B-D; Fig. S2G). From these data, we conclude that nucleotides 150-399, although important, are insufficient to mediate robust germ plasm localization when isolated from the rest of the *gcl* 3′UTR. Consequently, we infer that additional LEs must reside in the terminal fragments encompassing nucleotides 1-150 and nucleotides 399-525. Notably, in nearly every case, colocalization of reporter mRNAs with *nos* correlated with their average polar granule mRNA content (Fig. 2C,D).

Because germ plasm transcript abundance can influence localization (Gavis and Lehmann, 1992; Jongens et al., 1994; Niepielko et al., 2018), we quantified reporter RNA levels in 0-1 h old embryos by RT-qPCR. Although we did find variation, there was no correlation between the level of a particular RNA and its localization efficacy (Fig. S2I). Indeed, *gfp-gcl3′UTR*, which had the highest polar granule mRNA content, was among the lowest in mRNA levels. Additionally, *gfp-gcl(150-525)* was more abundant than *gfp-gcl(1-399)*, but was localized to a lesser degree. Thus, the observed behaviors of different reporter RNAs reflect the activity of the 3′UTR segments they contain.

Functional redundancy among *gcl* localization elements

As the *gcl* 3′UTR appears to contain multiple independently weak LEs, we wondered whether they might act additively or synergistically in mediating localization. To test this, we generated transgenes containing two tandem repeats (×2) of the proximal (1-276), central (150-399) or distal (247-525) segments of the *gcl* 3′UTR (Fig. 3A). In each case, we observed substantial increases in both polar granule colocalization and average mRNA content for the ×2 transcripts when compared with their single counterparts. This was most dramatic for *gfp-gcl(150-399)×2*, which showed an 87% increase in polar granule mRNA content over *gfp-gcl(150-399)*, and colocalized with *nos* at a frequency comparable with *gfp-gcl3′UTR* (Fig. 3B-D; Fig. S3A-H). These results cannot be attributed to differences in RNA levels; for example, *gfp-gcl(150-399)×2* is present at a lower level than *gfp-gcl(150-399)* (Fig. S3I). The finding that multiple copies of one region can partly substitute for the loss of others in both polar granule targeting and homotypic cluster growth indicates that *gcl* LEs can function redundantly. Together, our data indicate that *gcl* germ plasm localization is mediated by a combination of multiple, widely dispersed, and partially functionally redundant LEs in its 3′UTR.

Identification of a 59-nucleotide region that promotes accumulation of *pgc* within polar granules

We took the same approach to delineate LEs in the *pgc* 3′UTR. A fragment containing the 5′ two-thirds of the *pgc* 3′UTR, *pgc*(1-273), conferred significant but not wild-type germ plasm

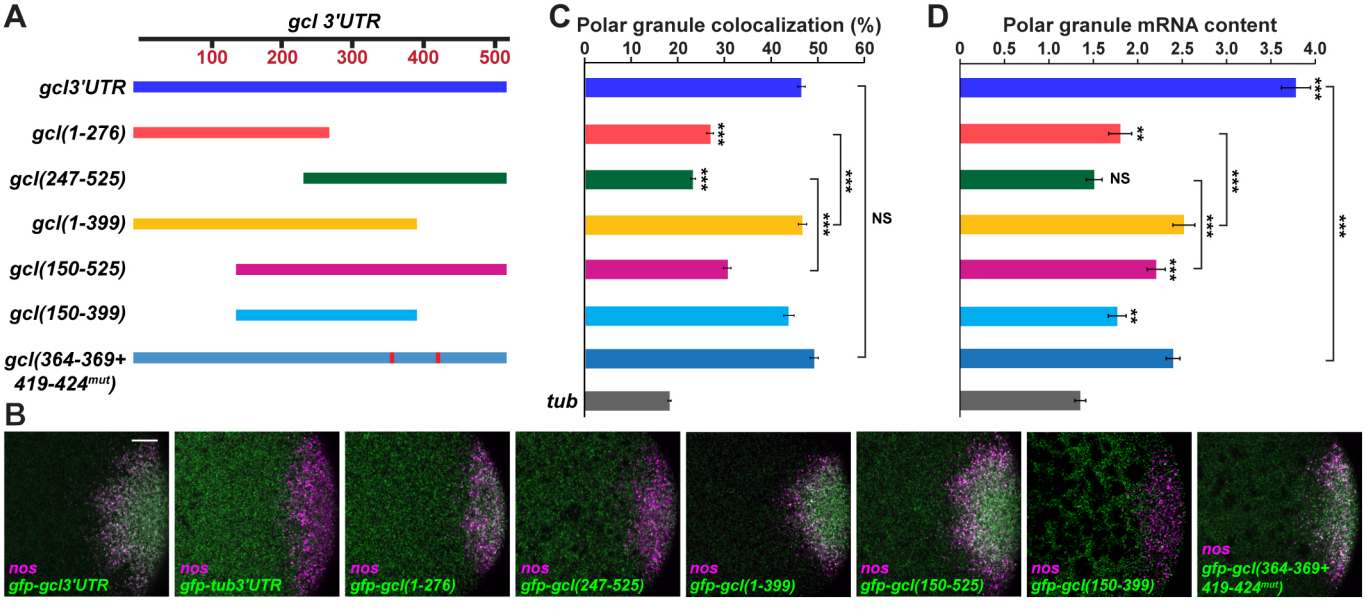

Fig. 2. Multiple individually weak localization elements are distributed throughout the *gcl* 3′UTR. (A) Schematic of *gcl* 3′UTR fragments tested. Red boxes indicate mutated sequences. (B) Confocal sections of the posterior region of 0-1 h old transgenic embryos (anterior towards the left, dorsal upwards). Embryos were probed simultaneously for *nos* mRNA (magenta) and reporter mRNA (green). Scale bar: 10 µm. (C) Nearest-neighbor quantification of colocalization between *nos* and reporter transcripts (*n*=20-30 embryos each). (D) Quantification of reporter mRNA content (average number of transcripts) in polar granules (*n*=10 embryos each). Values shown in C,D are mean±s.e.m., ***P*<0.01; ****P*<0.001, as determined by a two-tailed *t*-test. Unless specified, significance tests are in comparison with the *gfp-tub3′UTR* control.

Development (2018) 145, dev164657. doi:10.1242/dev.164657

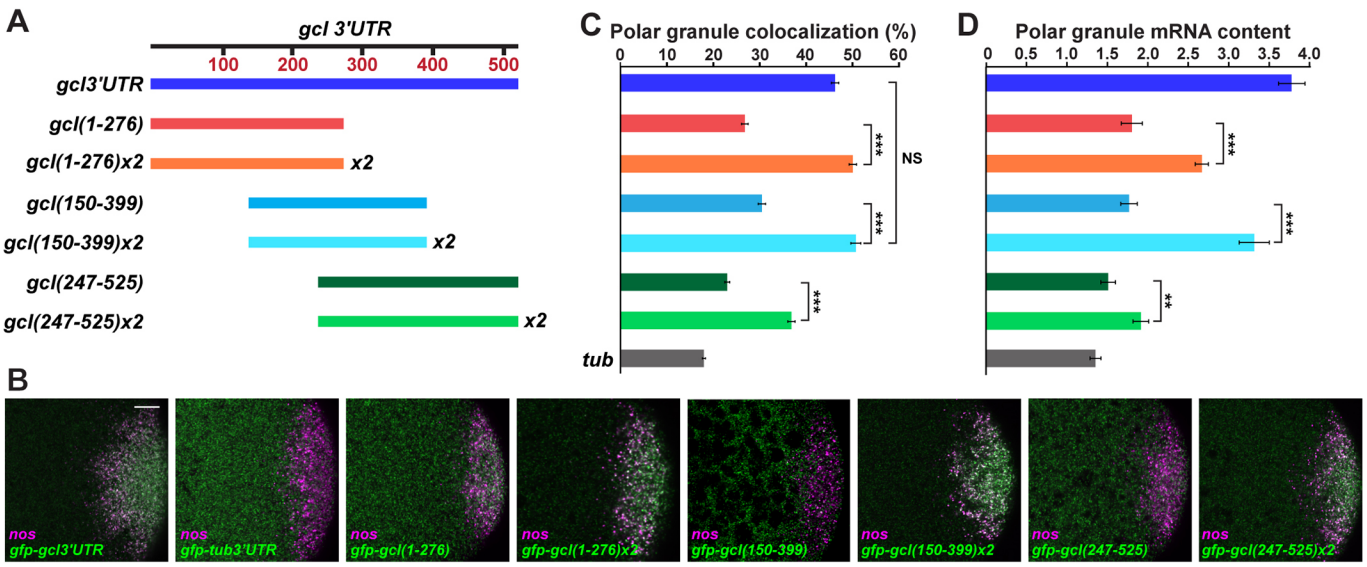

Fig. 3. Functional redundancy among *gcl* 3′UTR localization elements. (A) Schematic of *gcl* 3′UTR sequences tested for localization activity. Fragments that were tested as tandem dimers are indicated as ×2. (B) Confocal sections of the posterior region of 0-1 h old transgenic embryos (anterior towards the left, dorsal upwards). Embryos were probed simultaneously for *nos* mRNA (magenta) and reporter mRNA (green). Scale bar: 10 µm. (C) Nearest-neighbor quantification of colocalization between *nos* and reporter transcripts (*n*=20-30 embryos each). (D) Quantification of the average number of reporter mRNAs per polar granule (*n*=10 embryos each). Values shown in (C,D) are mean±s.e.m. ***P*<0.01; ****P*<0.001, as determined by a two-tailed *t*-test.

localization, whereas the remaining one-third, *pgc*(255-392), had no activity (Fig. 4; Fig. S4A-D). Although *pgc*(255-392) is insufficient in isolation, sequences in this region are important for wild-type localization efficacy, as shown by the 30% decrease in polar granule colocalization and 50% decrease in polar granule mRNA content when they are deleted from the *pgc* 3′UTR (Fig. 4C,D). When the 5′ fragment was shortened to the first 150 nucleotides, localization competence was lost. However, the central nucleotide 150-249 segment was also unable to direct localization on its own (Fig. 4; Fig. S4E,F). Therefore, the proximal nucleotide 1-150, the central nucleotide 150-249 and distal nucleotide 255-392 regions all contain LEs that work together to mediate localization.

Several mRNAs that are localized during early *Drosophila* development are known to rely on LEs characterized by secondary structures (Bullock et al., 2010; Cohen et al., 2005; Jambor et al., 2014). We therefore sought to identify potential secondary structure motifs in the *pgc* 3′UTR that are evolutionarily conserved among drosophilids, reasoning that these might identify LEs (see Materials and Methods). This analysis revealed a predicted, conserved stem-loop encompassing nucleotides 184-243 (Fig. 5B). Deletion of this sequence, *pgc*(*Δ184-243*), had no effect on colocalization of the

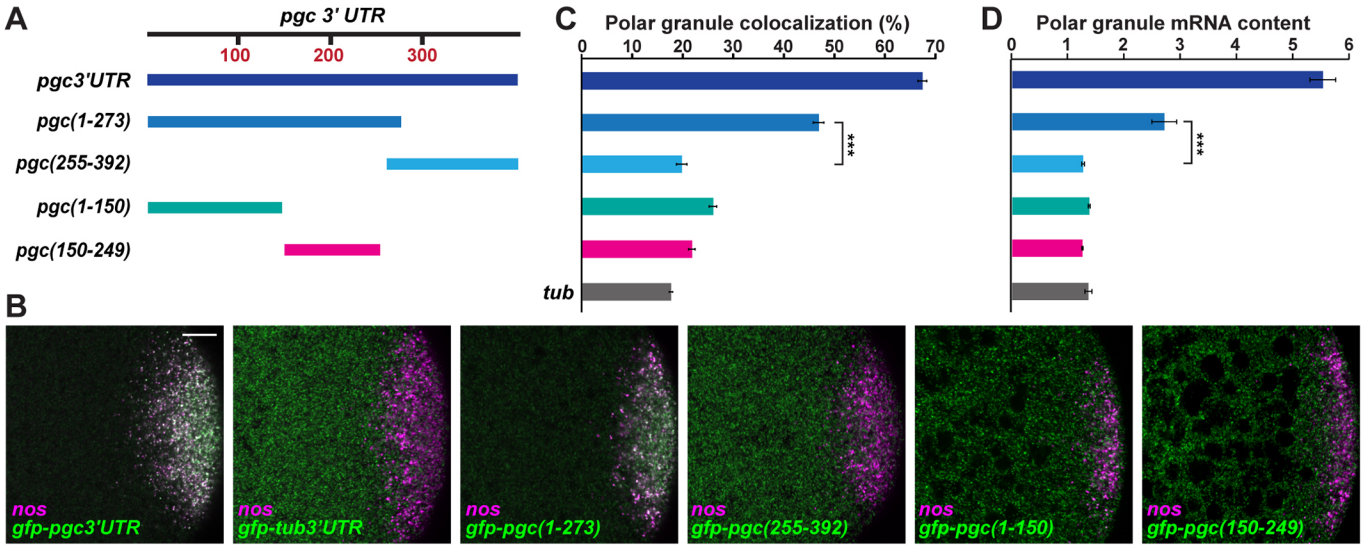

Fig. 4. The proximal, central and distal regions of the *pgc* 3′UTR all collaborate to mediate localization. Schematic of *pgc* 3′UTR sequences tested for localization activity. (B) Confocal sections of the posterior region of 0-1 h old transgenic embryos (anterior towards the left, dorsal upwards). Embryos were probed simultaneously for *nos* mRNA (magenta) and reporter mRNA (green). Scale bar: 10 µm. (C) Nearest-neighbor quantification of colocalization between *nos* and reporter transcripts (*n*=20-30 embryos each). (D) Quantification of the average number of reporter mRNAs per polar granule (*n*=10 embryos). Values shown in C,D are mean±s.e.m., ****P*<0.001, as determined by a two-tailed *t*-test.

Development (2018) 145, dev164657. doi:10.1242/dev.164657

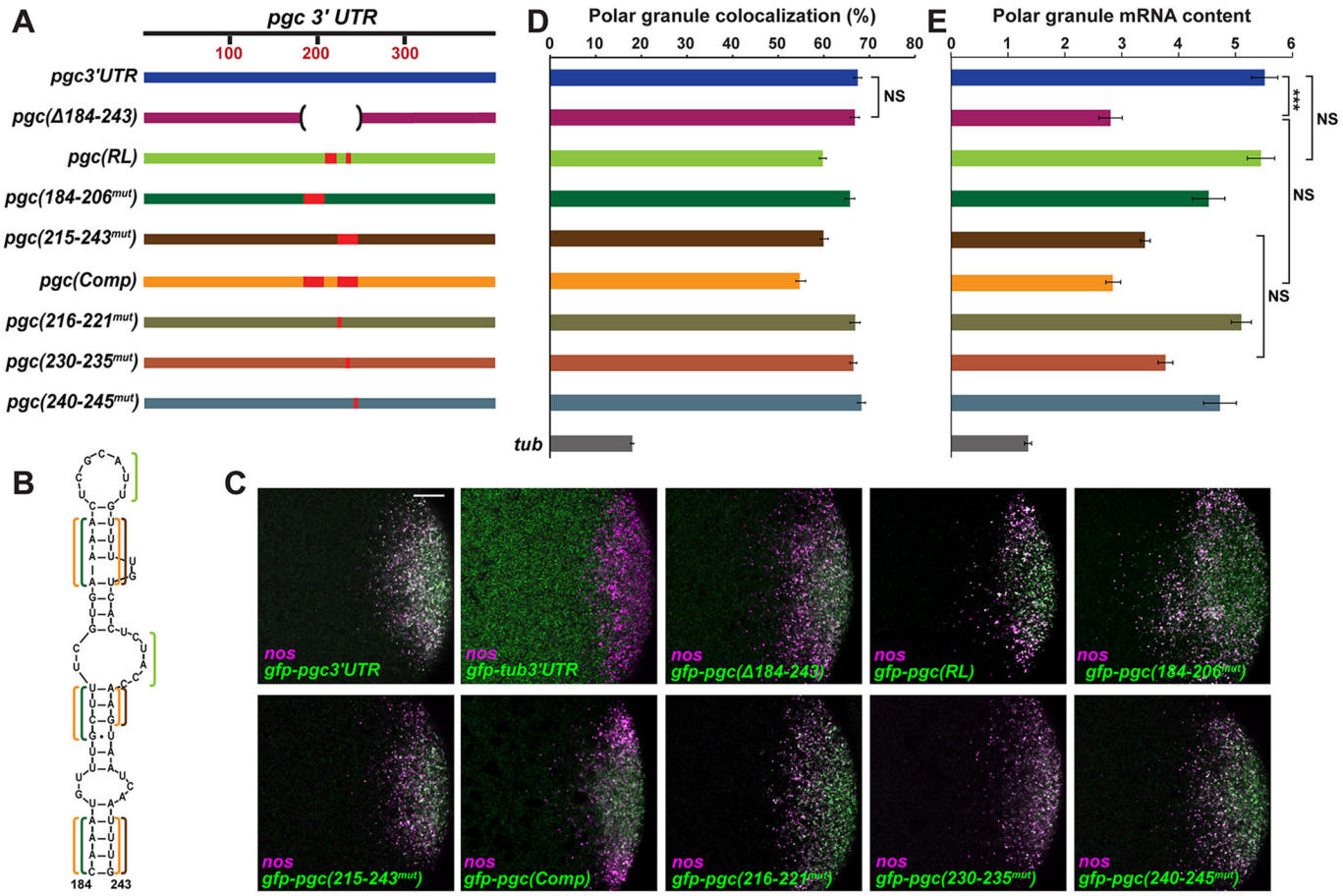

Fig. 5. A 59-nucleotide region of the *pgc* 3′UTR contains primary sequence elements responsible for polar granule *pgc* mRNA content. (A) Schematic of *pgc* 3′UTR sequences tested for localization activity. Brackets indicate deletions. Red boxes indicate mutated sequences. (B) Schematic of the predicted, conserved stem loop encompassing nucleotides 184-243 of the *pgc* 3′UTR. Colored brackets highlight mutated sequences: *pgc(RL)*, light green; *pgc(184-206^{mut})*, dark green; *pgc(215-243^{mut})*, brown; *pgc(Comp)*, orange. (C) Confocal sections of the posterior region of 0-1 h old transgenic embryos (anterior towards the left, dorsal upwards). Embryos were probed simultaneously for *nos* mRNA (magenta) and reporter mRNA (green). Scale bar: 10 µm. (D) Nearest-neighbor quantification of colocalization between *nos* and reporter transcripts (*n*=20-30 embryos each). (E) Quantification of the average number of reporter mRNAs per polar granule (*n*=10-13 embryos). Values shown in D,E are mean±s.e.m.; NS, not significant; ***P<0.001, as determined by a two-tailed *t*-test.

reporter RNA with *nos* (Fig. 5A-D). By contrast, the polar granule mRNA content for *gfp-pgc(Δ184-243)* was decreased by 49% when compared with *gfp-pgc3′UTR* (Fig. 5E; Fig. S4G).

To directly test whether the function of this 59-nucleotide region is mediated by the putative stem loop, we analyzed three additional sets of mutations (Fig. 5A,B). Substitution of nucleotides in the loop and bulge regions with random sequences that maintain the predicted structure in *gfp-pgc(RL)* had little or no effect on either polar granule colocalization or mRNA content (Fig. 5C-E; Fig. S4H), indicating that these sequences are dispensable. Two sets of mutations, within nucleotides 184-206 and nucleotides 215-243, which are each predicted to disrupt folding of the stem loop, reduced polar granule mRNA content by 15% and 35%, respectively, but had little or no effect on polar granule targeting (Fig. 5C-E; Fig. S4I,J). Surprisingly, although these mutations are predicted to restore the stem-loop structure when combined in *pgc(Comp)* (Fig. 5A,B), polar granule mRNA content was reduced by 49%, with little effect on polar granule colocalization (Fig. 5C-E; Fig. S4K). The failure of the compensatory mutations to restore mRNA content leads us to conclude that rather than secondary structure, primary sequences within nucleotides 184-206 and nucleotides 215-243 of the *pgc* 3′UTR together are responsible for nearly half of polar granule *pgc* content. The ability of these sequences to selectively affect polar granule mRNA content independent of polar granule colocalization suggests that distinct elements in the *pgc* 3′UTR mediate targeting to polar granules and the subsequent growth of homotypic clusters.

Identification of a *pgc* homotypic clustering element

Using SIM, we recently showed that posteriorly localized RNAs can form two or more homotypic clusters within a polar granule. These likely occur through multiple, independent targeting events followed by recruitment of like transcripts (Niepielko et al., 2018). Thus, although polar granule mRNA content is determined primarily by homotypic cluster size (Niepielko et al., 2018), sequences such as those within *pgc* nucleotides 184-243 could affect mRNA content by contributing to these additional targeting events, as well as to homotypic cluster growth. To address this possibility, we performed SIM imaging on *gfp-pgc3′UTR* and *gfp-pgc(215-243^{mut})* RNAs.

As detected by SIM, both *gfp-pgc3′UTR* RNA and *gfp-pgc(215-243^{mut})* RNAs form multiple homotypic clusters within polar granules (marked with Osk-GFP), similarly to endogenous *pgc* (Fig. 6A). To confirm that detection of multiple clusters was not due to random overlap of particles in the imaging volume, we analyzed a non-localizing transcript, *lost*, which shows minimal colocalization

Development (2018) 145, dev164657. doi:10.1242/dev.164657

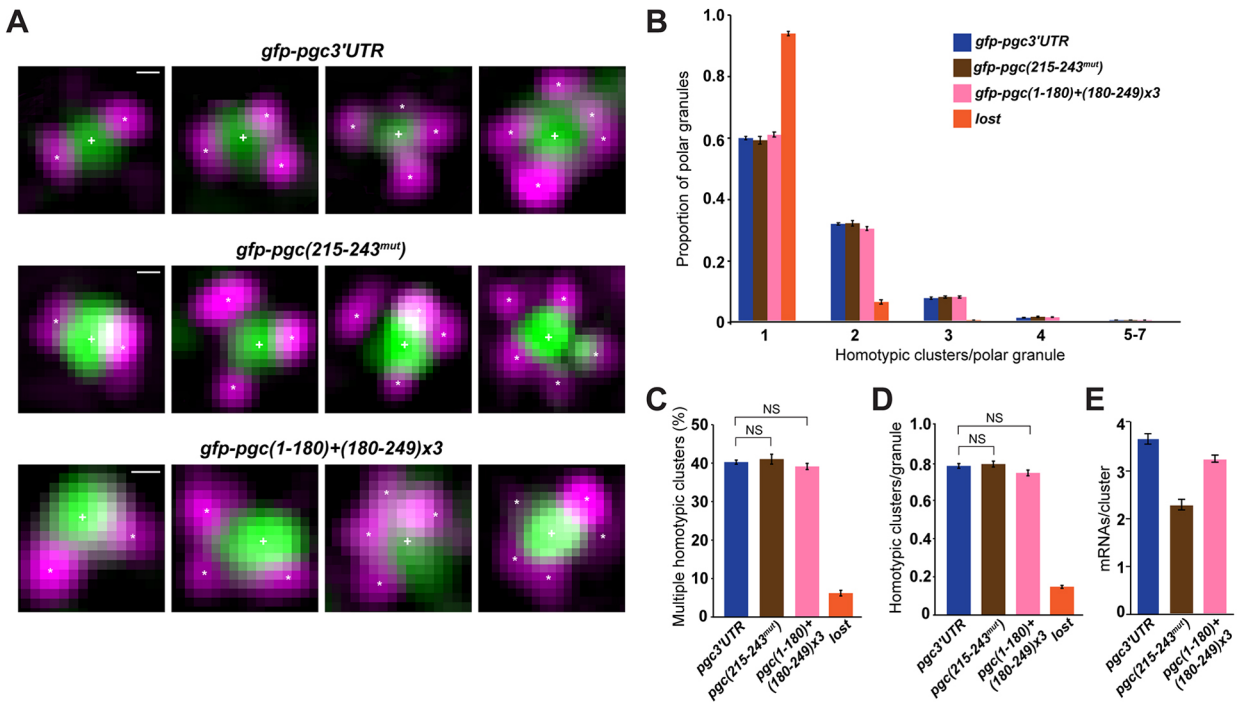

Fig. 6. Occurrence of multiple homotypic clusters for *gfp-pgc3′UTR* reporter RNAs. (A) Examples of polar granules with multiple (2-5) homotypic clusters of *gfp-pgc3′UTR*, *gfp-pgc(215-243^{mut})* and *gfp-pgc(1-180)+(180-249)×3* RNAs in 0-1 h old embryos. Individual polar granules are identified by Osk-GFP, detected by direct GFP fluorescence (green). mRNA was detected by smFISH (magenta). Scale bars: 100 nm. (B) Quantification of the proportion of polar granules (with at least one homotypic cluster) that contain a given number of homotypic clusters; *n*=6 (*lost*), 15 (*gfp-pgc3′UTR*), 14 (*gfp-pgc215-243^{mut}*) and 11 (*gfp-pgc1-180 +180-249×3*) embryos. (C) For polar granules (with at least one homotypic cluster), the proportion that contain multiple clusters was quantified. (D) Quantification of the total number of colocalized homotypic clusters divided by the total number of polar granules detected. (E) Quantification of the average number of mRNAs relative to the average number of homotypic clusters of that mRNA per granule. Values shown in B-D are mean±s.e.m.; values in E are mean±s.d.

with polar granules (Niepielko et al., 2018). Among the rare polar granules where overlap with *lost* was detected, the majority (94%) had only one *lost* particle (Fig. 6B,C). To determine how disruption of sequences in *gfp-pgc(215-243^{mut})* affects polar granule mRNA content, we measured the frequency distribution of multiple homotypic clusters of *gfp-pgc3′UTR* and *gfp-pgc(215-243^{mut})* RNA, and quantified the average number of mRNAs per cluster. Both the distribution and the average overall frequency of multiple homotypic clusters were indistinguishable for the two transcripts (Fig. 6B-D). Moreover, the average number of mRNA/cluster decreased by 38% for *gfp-pgc(215-243^{mut})* when compared with *gfp-pgc3′UTR*, a value identical to the decrease in granule mRNA content (Fig. 6E, compare with Fig. 5E). Thus, the decrease in mRNA content resulting from perturbation of nucleotides 215-243 can be attributed to a reduction in average homotypic cluster size, defining the sequences within this region as a clustering element.

Discovery of a conserved motif that regulates homotypic clustering

To further delimit the responsible sequences within nucleotides 215-243, we introduced sets of 6-nucleotide mutations across this region: *pgc(216-221^{mut})*, *pgc(230-235^{mut})* and *pgc(240-245^{mut})* (Fig. 5A). Whereas none of these mutations impacted polar granule colocalization (Fig. 5D), *gfp-pgc(230-235^{mut})* behaved comparably to *gfp-pgc(215-243^{mut})*, decreasing polar granule content for the reporter RNA by 32% (Fig. 5E). Although flanking sequences make minor contributions, nucleotides 230-235 likely constitute the bulk of the homotypic clustering element.

A priori, sequences that function in homotypic clustering might be expected to be RNA specific. Therefore, we were surprised to find two matches or close matches to the *pgc*(230-235) sequence, CAAGUA, in both the *gcl* 3′UTR (UAAGUA and CAAGUU) and the *nos* 3′UTR (CAAGUC and CAAGUA) (Fig. S2J, Fig. S5B). Remarkably, mutation of the *gcl* sequences to generate *gfp-gcl(364-369+419-424^{mut})* (Fig. 2A; Fig. S2J) resulted in a 37% decrease in polar granule mRNA content when compared with *gfp-gcl3′UTR* without affecting polar granule colocalization (Fig. 2B-D; Fig. S2H,I). Similarly, mutation of the *nos* sequences to generate *gfp-nos(217-222+645-650^{mut})* (Fig. S5A,B) resulted in a 42% decrease in polar granule mRNA content when compared with *gfp-nos3′UTR*, with minimal effect on colocalization (Fig. S5C-E). These data suggest that the conserved motifs have a generalized function in mRNA clustering.

Distinct elements mediate targeting of *pgc* to polar granules and homotypic clustering

The ability of mutations that disrupt sequences within nucleotides 184-243 of the *pgc* 3′UTR to affect polar granule mRNA content without affecting polar granule colocalization suggests that distinct elements within the *pgc* 3′UTR mediate *pgc* targeting to polar granules and homotypic clustering. Indeed, the nucleotide 184-243 region is ineffective on its own, with even three tandem copies [*gfp-pgc(180-249)×3*] failing to confer localization (Fig. 7; Fig. S6D,E). Thus, whereas the *gfp-pgc(180-249)* and *gfp-pgc(180-249)×3* mRNAs contain clustering elements, they may lack a requisite targeting element. The partial localization of *gfp-pgc(1-249)* suggests that such a targeting element resides within the first 180 nucleotides

Development (2018) 145, dev164657. doi:10.1242/dev.164657

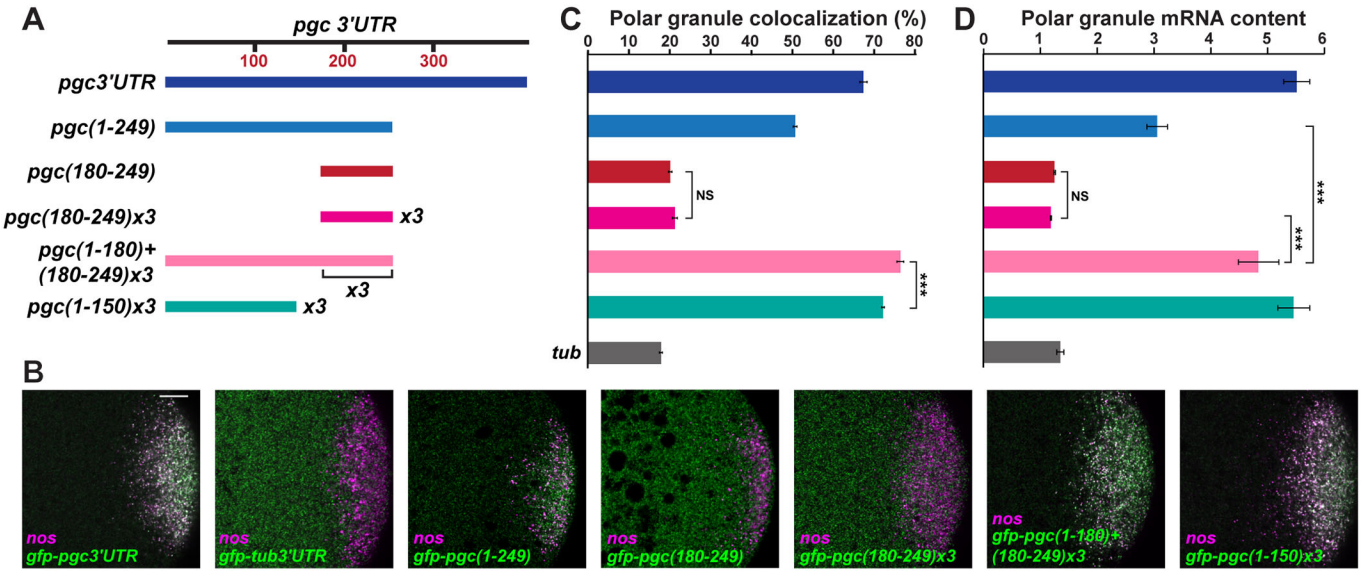

Fig. 7. *pgc* polar granule targeting and homotypic cluster growth are mediated by distinct 3′UTR elements. (A) Schematic of *pgc* 3′UTR sequences and multimers tested for localization activity. (B) Confocal sections of the posterior region of 0-1 h old transgenic embryos (anterior towards the left, dorsal upwards). Embryos were probed simultaneously for *nos* mRNA (magenta) and reporter mRNA (green). Scale bar: 10 µm. (C) Nearest-neighbor quantification of colocalization between *nos* and reporter transcripts (n=30 embryos each). (D) Quantification of the average number of reporter mRNAs per polar granule (n=10 embryos each). Values shown in (B,C) are mean±s.e.m.; NS, not significant; ***P<0.001, as determined by a two-tailed t-test.

of the *pgc* 3′UTR (Fig. 7, Fig. S6C). To test this, we sought to rescue localization of *gfp-pgc(180-249)×3* by adding back this 180-nucleotide region, generating *gfp-pgc(1-180)+(180-249)×3* (Fig. 7A). Indeed, polar granule colocalization and mRNA content of *gfp-pgc(1-180)+(180-249)×3* RNA were nearly comparable with that of *gfp-pgc3′UTR* (Fig. 7B-D; Fig. S6F). Using SIM, we confirmed that the number of homotypic clusters per granule and the number of mRNAs per cluster were minimally affected (Fig. 6). These data provide evidence for distinct targeting and clustering elements that together confer wild-type germ plasm localization. The fact that additional copies of the nucleotide 180-249 region can substitute for the loss of nucleotides 250-392 indicates that there is functional redundancy among *pgc* clustering elements.

Our initial dissection of the *pgc* 3′UTR (Fig. 4) provided evidence that LEs within nucleotides 1-150 and nucleotides 150-249 work together to mediate localization. Given the evidence that one or more LEs within nucleotides 1-150 functions in mRNA targeting, we wondered whether multimerization of this sequence could promote stable polar granule interaction in the absence of a capacity for homotypic clustering. To achieve this, we made *gfp-pgc(1-150)×3* (Fig. 7A). Strikingly, not only did this RNA very effectively colocalize with polar granules, it also showed dramatic accumulation within the polar granules (Fig. 7B-D; Fig. S6H). Consistent with the previous analyses, localization behavior could not be explained by transcript levels (Fig. S6I).

From these data, we derive several conclusions. First, mRNA localization is directed by both targeting and clustering elements. Without a targeting element, the number of clustering elements has little bearing, as exemplified by *gfp-pgc(180-249)×3*. Second, the 58% increase in polar granule mRNA content for *gfp-pgc(1-180)+(180-249)×3* relative to *gfp-pgc(1-249)* indicates that the centrally located clustering elements can indeed act additively, but only in the context of a sequence that confers competency to interact with polar granules. Finally, either some *pgc* targeting elements can promote homotypic clustering when multimerized, or an additional weak clustering element is present in nucleotides 1-150.

DISCUSSION

All characterized mRNA localization mechanisms function through *cis*-acting localization signals. For transcripts such as *nos*, *gcl* and *pgc*, which localize by a diffusion and entrapment process, 3′UTRs are sufficient to direct localization (Gavis and Lehmann, 1992; Rangan et al., 2009) but the specific features responsible for their function have remained largely unknown. Here, we have used high-resolution imaging and quantitative analysis of mRNA localization at the level of polar granules to map and characterize the functions of LEs in the *gcl* and *pgc* 3′UTRs. Our results reveal parallels to *nos*, with robust localization of reporter RNAs requiring the function of multiple regulatory elements spread widely across the 3′UTR. Although weak in isolation, these LEs can function redundantly such that multiple copies of one can at least partly substitute for lack of another. Thus, complexity in localization signal organization appears to be a feature of *Drosophila* germ plasm transcripts. Our results also provide evidence for another layer of complexity, by showing that targeting to polar granules and formation of homotypic clusters are separable processes regulated by distinct classes of *cis*-acting elements.

Compound mRNA localization signals with elements that direct discrete intermediate steps of a localization pathway have been found in some actively transported mRNAs such as *Drosophila bicoid* and *osk* (Jambor et al., 2014; Kim-Ha et al., 1993; Macdonald and Kerr, 1998; Macdonald et al., 1993). These elements allow the sequential association of an mRNA with different transport machineries. The compound nature of *Drosophila* germ plasm mRNA localization signals also allows for a process involving multiple events, in this case the targeting of mRNAs to polar granules and the growth of homotypic clusters within the granules. These roles are potentially similar to roles played by LEs comprising the compound *Xenopus nanos1* mRNA localization signal: the mitochondrial cloud localization element (MCLE) that directs *nanos1* along with various RNAs to the cloud and the germ granule localization element (GGLE), which targets *nanos1* to the germ granules. Like the *pgc* clustering elements,

Development (2018) 145, dev164657. doi:10.1242/dev.164657

which function only when paired with a targeting element, the GGLE functions only in the presence of a mitochondrial cloud targeting element – i.e. the RNA must enter the cloud in order to associate with germ granules (Chang et al., 2004; Kloc et al., 2000). Intriguingly, different *Xenopus* germline RNAs exhibit different distributions within the mitochondrial cloud, with *nanos1* sequestered within granules and other RNAs on the outside or associated with the matrix between granules (Kloc et al., 2002). This pattern hints that *Xenopus* germ granule assembly may also involve mRNA clustering and that the *nanos1* GGLE might function similarly to the clustering elements found in *Drosophila* germ granule mRNAs.

The presence of multiple elements that act additively, and often redundantly, to confer localization is a feature of many RNA localization signals, regardless of the method of localization (Kloc and Etkin, 2005; Shahbabian and Chartrand, 2012). A multiplicity of partially redundant LEs within *Drosophila* germ plasm mRNA localization signals might be advantageous for several reasons. The presence of multiple clustering elements may counterbalance an inherently inefficient localization process as 4% or less of the various transcripts ultimately become localized to the germ plasm (Bergsten and Gavis, 1999; Little et al., 2015; Trcek et al., 2015). Given the large volume of the oocyte and the narrow posterior cortical domain in which germ granules form (Little et al., 2015), contact of the diffusing single-transcript RNPs with nascent polar granules is likely to be infrequent. The presence of multiple elements and cognate binding proteins would thus increase the likelihood of interaction.

Additionally, the inclusion of multiple clustering elements within an mRNA molecule may create multivalency that can control clustering either through RNA-RNA interactions or through interactions with RNA-binding proteins, which may themselves be multivalent. Each newly recruited mRNA would in turn provide additional interaction surfaces, allowing for continuous cluster growth (Little et al., 2015). This is consistent with our observation that every reporter RNA that produces large homotypic clusters is predicted to include at least three clustering elements. Furthermore, RNAs with only one predicted clustering element form homotypic clusters that contain, on average, only slightly more than two transcripts. As pole cell survival correlates with germ plasm inheritance (Slaidina and Lehmann, 2017), the additive activity of multiple targeting and clustering elements together would ensure that a sufficient quantity of each RNA makes it into the germline to facilitate proper development.

Self-association of germ plasm transcripts in homotypic clusters occurs exclusively within the polar granules, never in the bulk cytoplasm (Little et al., 2015). Thus, there must be a permissive switch upon interaction of a single-transcript RNP with a polar granule that allows for the recruitment of additional like-transcripts, and the subsequent growth of homotypic clusters. The nature of this switch remains unknown, but it is notable that multiple RNA helicases are present in the polar granules, including Vasa, Belle and Me31B (Hay et al., 1988; Johnstone et al., 2005; Thomson et al., 2008). These factors might direct remodeling of single-transcript RNPs upon their interaction with a polar granule, thereby unmasking previously unexposed clustering elements.

Evidence for localization signal sequences or structures shared among *Drosophila* germ plasm mRNAs has been lacking, despite their common destination. Whether many different targeting elements are recognized degenerately by one or more *trans*-acting factors or whether different sets of proteins can all mediate polar granule targeting remains an unresolved issue. In contrast to sharing the ability to incorporate into polar granules, the capacity for *nos*, *pgc* and *gcl* transcripts to form spatially segregated homotypic clusters within polar granules (Little et al., 2015) predicts that each harbors unique clustering elements. Unexpectedly, at least one clustering element is shared by the *pgc*, *gcl* and *nos* 3′UTRs. This suggests that in addition to providing multivalency, clustering elements may act combinatorially to generate an RNA-specific signature. Analyzing the clustering behavior of reporters with chimeric 3′UTRs containing combinations of *pgc*, *gcl* and/or *nos* LEs may provide insight into the contributions of these elements to RNA self-recruitment.

MATERIALS AND METHODS

Fly stocks

The following *Drosophila melanogaster* alleles and transgenes were used: y^1, w^{67c23} (Bloomington Stock Center 6599) as the wild-type strain; and *osk-GFP* (Sarov et al., 2016). The *matα4-GAL-VP16* driver (Bloomington 7063) was used to express all *UASp-gfp-3′UTR* transgenes.

Transgene construction

The *gfp*-coding sequence was PCR amplified from pEGFP-N1 (Clontech) and inserted into a pBS-SK plasmid containing the *nos* 5′UTR at the position of the *nos* ATG. A fragment containing the *nos* polyadenylation signal and 460 bp of 3′ genomic DNA was excised from a genomic *nos* plasmid (Gavis and Lehmann, 1992) and inserted downstream of the *gfp*-coding sequence. *pgc* 3′UTR sequences were PCR amplified from y^1, w^{67c23} genomic DNA. The *pgc 3′UTR* we isolated was missing 12 bp (nucleotides 297-308) when compared with the annotated sequence on FlyBase. *gcl 3′ UTR* sequences were amplified from cDNA clone LD23660 (Drosophila Genome Resource Center). The *tubulin 3′UTR* sequence was PCR amplified from a Bluescript plasmid containing nucleotides 1-198 of the *alphaTub84B* 3′UTR (Theurkauf et al., 1986). Each of these sequences was inserted 3′ to the *gfp*-coding sequence. For germline integration, we modified the pattB insertion vector by inserting the GAL4-binding site cassette from pUASp (nucleotides 114-734) into the *pattB* insertion vector (www.flyc31.org/sequences_and_vectors.php) between *Eco*RI and *Xba*I to generate pattB-UASp. The nos-gfp-3′UTR sequences were then cloned into pattB-UASp and integrated into the attp40 site via injection and phiC31-mediated recombination.

Mutations generated within *pgc* 3′UTR nucleotides 184-243 with altered nucleotides underlined:

wild type, CAAATGTTTGCTTTCGTGAAAACTCGCATTGTTTTGTCACTCTACCAAGTAATCAATTTG; *pgc*(RL), CAAATGTTTGCTTTCGTGAAAACTTCGGTTTTGTCACAACCTTAAGTAATCAATTTG; *pgc*(184-206^{mut}), ACCCTGTTTAAGGTCGTGCGCGCTCGCATTGTTTTGTCACTCTACCAAGTAATCAATTTG; *pgc*(215-243^{mut}), CAAATGTTTGCTTTCGTGAAAACTCGCATTGCGCTGGCACTCTACCCCTTAATCAAGGGT; *pgc*(216-221^{mut}), CAAATGTTTGCTTTCGTGAAAACTCGCATTGCGCCAGCACTCTACCAAGTAATCAATTTG; *pgc*(230-235^{mut}), CAAATGTTTGCTTTCGTGAAAACTCGCATTGTTTTGTCACTCTACACCTGCATCAATTTG; and *pgc*(240-245^{mut}), CAAATGTTTGCTTTCGTGAAAACTCGCATTGTTTTGTCACTCTACCAAGTAATCACGGGTC.

Single molecule fluorescent *in situ* hybridization

smFISH probe sets of 20-nucleotide oligonucleotides with 2-nucleotide spacing, complementary to *nos* (CG5637; 63 oligos), to the *nos*-coding region (34 oligos), to *pgc* (CG32885; 52 oligos), to *gcl* (48 oligos), to *lost* (48 oligos) or to *gfp* (32 oligos) were designed using the Stellaris Probe Designer. Oligonucleotides were obtained from Biosearch Technologies with a 3′ NHS ester modification and conjugated to Atto 565 (Sigma 72464) or Atto 647N (Sigma 18373) dyes, then purified via HPLC as described previously (Raj et al., 2008). smFISH was performed as previously described (Little et al., 2015) on 0-1 h old embryos. Embryos were mounted in Prolong Diamond medium (Fischer Scientific P36965) and cured for 3-4 days at room temperature in the dark prior to imaging.

Development (2018) 145, dev164657. doi:10.1242/dev.164657

Microscopy

Confocal imaging was performed on a Nikon A1-RS laser-scanning confocal microscope using a 60×1.4 NA oil immersion objective at either 2× or 3× optical zoom, with pixels of 102×102 nm or 72×72 nm, respectively. Confocal sections were acquired with 16× line averaging. Super-resolution images were taken on a Nikon Structured Illumination Microscope using a 100×1.49 NA oil immersion objective, with pixels of 33×33 nm. For each embryo, a *z* series of 21 slices were taken, with a step size of 150 nm. In each experiment, laser power and gain were adjusted to avoid signal saturation while maximizing separation of signal and noise. Embryos selected for imaging were oriented such that the germ plasm was facing the cover slip. To minimize fluorescent signal distortion, all images were taken within 5 µm of the embryo cortex. When the same transgene is shown in more than one figure, the same accompanying confocal image is used each time to allow comparison across figures.

Image quantification

Confocal colocalization analysis

Colocalization of *nos* with mRNAs in the germ plasm reflects their incorporation into polar granules (Little et al., 2015; Niepielko et al., 2018). This metric is therefore a proxy for colocalization of reporter mRNAs with polar granules. All analysis of confocal images was carried out on single sections. In MATLAB (MathWorks), a region minimally outlining the germ plasm of each embryo was specified by a manually drawn polygon (Fig. S1E). The spotDetector algorithm (Aguet et al., 2013) was then used to detect and assign the *x*, *y* coordinates of every mRNA signal in each channel (~500-1000 particles per embryo) (Fig. S1E,F). Distances were calculated from each point in the *nos* channel to the nearest point in the reporter mRNA channel. Functional colocalization is defined as a distance less than 300 nm, on the basis of average polar granule radius (Amikura et al., 2001; Houston, 2013; Illmensee and Mahowald, 1974). The colocalization frequency in an embryo is the percentage of total RNA molecules within the germ plasm that fall below this threshold (Fig. S1G). Colocalization frequencies are displayed as the mean value measured from 20-30 embryos per genotype±s.e.m.

Polar granule RNA content

The absolute mRNA content per granule was determined by normalization of granule fluorescence intensity to fluorescence per single-copy particle (Little et al., 2015). For each embryo, sequential images were collected under two conditions: one (high gain) optimized for the unlocalized, single-copy RNPs, which results in saturation of the germ plasm fluorescent signal; and one (low gain) optimized for the germ plasm signal.

The average fluorescence intensity of single-copy RNPs was determined by manually demarcating an ROI in the bulk cytoplasm of the high gain image containing >500 particles. A spot detection algorithm (Little et al., 2015, 2011) was then used to assign the *x*, *y* coordinates and measure the fluorescence intensity of every RNP in that region, using a threshold set manually to eliminate false positives. A 13×13 pixel mask was fitted to every detected RNP and the intensity of every pixel of that grid for all particles was averaged to obtain an intensity value for an idealized single-copy RNP. To obtain a conversion factor for normalization of germ granule fluorescence intensity, the same ROI was then analyzed similarly in the low-gain image. The intensity of every RNP in the low-gain image was plotted versus its intensity in the high-gain image and a linear regression analysis was used to fit a line to the data. The slope of this line yields the conversion factor.

Average fluorescence intensity of the germ granules was obtained from the low-intensity image. A polygon was drawn manually to bound the germ plasm, which was demarcated using endogenous *nos* RNA, and the spot detection algorithm was used to detect all RNPs within the germ plasm. The data were fit to a Gaussian for measurement of fluorescence intensity. Intensity values were multiplied by the conversion factor, and divided by the intensity of the idealized unlocalized RNP to generate an estimation for the number of transcripts within each polar granule. These values were then averaged for the germ plasm of a given embryo. The germ plasm of 10 embryos was quantified per genotype to generate the average polar granule mRNA content.

Homotypic cluster quantification

At the start of every experiment, an image of TetraSpeck beads (ThermoFisher, T7279) was taken and used in conjunction with Nikon software to generate alignment correction parameters that were applied to germ plasm images. RNA particle detection and image quantification was carried out using a custom Matlab program as previously described in detail (Niepielko et al., 2018). In summary, N-SIM images were filtered by a balanced circular difference-of-Gaussian with a center radius size of 1.2 pixels and surround size of 2.2 pixels. We clipped a 13×13 pixel mask centered around each image punctae and fit a Gaussian distribution in three consecutive *z* slices and as well as in *x*-*y* to identify the *x*-*y*-*z* locations of candidate Osk protein and mRNA particles. True particles were determined by intensity thresholds set manually to eliminate false positives. To determine whether Osk and RNA particles are colocalized, we first calculated distances in *x*-*y* for all particle pairs. Next, we selected colocalized pairs based on the following criteria: (1) two particles must be within a *z* distance four slices, which accounts for chromatic aberration while eliminating pairs that may colocalize in *x*-*y* but cannot be in the same granule due to the size limitation of germ granules; and (2) a colocalized particle pair must also be within a distance limit of 200 nm in *x*-*y*. A conservative distance was chosen based on the average size of a germ granule that was previously used to calculate colocalization frequency among germ plasm mRNAs (Little et al., 2015). The number of discrete RNA particles that satisfy the colocalization criteria for each Osk particle was recorded. For our analysis, any colocalized RNA particle was considered to be a homotypic cluster, even if it contained a single transcript.

RT-qPCR

RNA was isolated using the RNeasy mini kit (Qiagen) from 0-1 h old embryos. RNA (~1 µg) was reverse transcribed to cDNA using the Quantitect Reverse Transcription kit (Qiagen). qPCR was run using an Applied Biosystems 7900HT standard 96-well qPCR instrument. The following TaqMan Gene Expression probes from ThermoFisher were used: *gcl* (Dm01812234, 4351372); *pgc* (custom - DmAJHSOQG, 4441114); *egfp* (Mr04097229, 4331182); and *rpl7* (Dm01817653, 4351372). All experiments used a CT threshold of 0.6613619. Three technical replicates were quantified from each cDNA pool. DNA standard curves were generated for each probe set using serial dilutions of plasmids containing target sequences. CT values from each assay were fit to the standard curves, corrected for the experimentally determined doubling efficiency of each probe, and normalized to the housekeeping gene *rpl7*. Doubling efficiencies were as follows: *pgc* (87.5%), *gcl* (92.7%), *gfp* (97.2%) and *rpl7* (88.5%). Finally, reporter mRNA expression levels were normalized to endogenous *pgc* or *gcl* expression in 0-1 h old wild-type embryos. Expression levels are displayed as mean±s.d.

Secondary structure prediction

The *D. melanogaster pgc* 3′UTR sequence was obtained from FlyBase (flybase.org/). To obtain *pgc* 3′UTRs from representative species across the *Drosophila* phylogenetic tree, the extended *pgc* gene regions and coding sequences from *D. melanogaster*, *D. sechellia*, *D. erecta*, *D. ananassae*, *D. willistoni* and *D. virilis* were downloaded from FlyBase. A custom Python script 'three_prime_detector.py' based on the SeqIO and AlignIO modules from the Biopython library (Cock et al., 2009; available at biopython.org/) was used to extract each 3′UTR sequence by first generating a Clustal W 2.0 (Larkin et al., 2007) alignment of the extended gene region and coding region, and then selecting the nucleotide sequence in the extended gene region beginning downstream of the stop codon and extending the length of the *D. melanogaster pgc* 3′UTR+50 bp. Local secondary structures conserved among the different *pgc* 3′UTR sequences were predicted using RNApromo (Rabani et al., 2008), a tool that predicts structural motifs common to a set of RNA sequences, in conjunction with the ViennaRNA package (Lorenz et al., 2011). The output of RNApromo is a log-likelihood score for each instance of secondary structure motif discovered in a set of sequences. For each motif, the sum of the individual log-likelihood scores is a measure of its relative conservation. To identify 'true positive' structures, we used a custom Python script 'random_seq_generator.py'. to generate 100 randomly shuffled *pgc* 3′UTR sequences for each of the six species. RNApromo was used to produce a dummy dataset of conserved local secondary structures from these shuffled sequences. We then used a custom R script 'cutoff_secondary_score.R' to calculate a 'secondary score' for each motif by first summing the log-

likelihood scores of each secondary structure motif found in the random dummy dataset and then dividing by the square root of the length of that motif to control for increased likelihood due to length. Secondary scores were sorted in descending order and a threshold was set at the 95% percentile score. Secondary structures in the true dataset with log-likelihood scores above this threshold (<5% chance of being a false positive) were considered further. Two secondary structures in different regions of the *pgc* 3′UTR were identified whose log-likelihood scores exceeded the cutoff score. The functional significance of the top scoring motif was experimentally tested (Fig. 5). Custom scripts provided upon request.

Acknowledgements
We thank S. Little and J. Lee for assistance with the MatLab scripts used to quantify transgene colocalization and intensity phenotypes. We thank G. Laevsky for assistance with microscopy. We thank S. Little, C. Ruesch, P. Schedl and J. Tamayo for critical comments on the manuscript.

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: W.V.I.E., E.R.G.; Methodology: W.V.I.E., E.R.G.; Software: D.K.Y.-K., M.G.N.; Formal analysis: W.V.I.E.; Investigation: W.V.I.E.; Writing - original draft: W.V.I.E., E.R.G.; Writing - review & editing: W.V.I.E., M.G.N., E.R.G.; Supervision: E.R.G.; Project administration: E.R.G.; Funding acquisition: E.R.G.

Funding
This work was supported by National Institute of Health grants R01 GM067758 to E.R.G. and F32 GM119200 to M.G.N. W.V.I.E. was supported by the National Institute of Health training grant T32 GM007388. Deposited in PMC for release after 12 months.

Supplementary information
Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.164657.supplemental

References
Aguet, F., Antonescu, C. N., Mettlen, M., Schmid, S. L. and Danuser, G. (2013). Advances in analysis of low signal-to-noise images link dynamin and AP2 to the functions of an endocytic checkpoint. *Dev. Cell* **26**, 279-291.

Amikura, R., Hanyu, K., Kashikawa, M. and Kobayashi, S. (2001). Tudor protein is essential for the localization of mitochondrial RNAs in polar granules of *Drosophila* embryos. *Mech. Dev.* **107**, 97-104.

Bergsten, S. E. and Gavis, E. R. (1999). Role for mRNA localization in translational activation but not spatial restriction of *nanos* RNA. *Development* **126**, 659-669.

Bergsten, S. E., Huang, T., Chatterjee, S. and Gavis, E. R. (2001). Recognition and long-range interactions of a minimal *nanos* RNA localization signal element. *Development* **128**, 427-435.

Bullock, S. L., Ringel, I., Ish-Horowicz, D. and Lukavsky, P. J. (2010). A′-form RNA helices are required for cytoplasmic mRNA transport in *Drosophila*. *Nat. Struct. Mol. Biol.* **17**, 703-709.

Chang, P., Torres, J., Lewis, R. A., Mowry, K. L., Houliston, E. and King, M. L. (2004). Localization of RNAs to the mitochondrial cloud in *Xenopus* oocytes through entrapment and association with endoplasmic reticulum. *Mol. Biol. Cell* **15**, 4669-4681.

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B. et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422-1423.

Cohen, R. S., Zhang, S., Dollar, G. L. (2005). The positional, structural, and sequence requirements of the Drosophila TLS RNA localization element. *RNA* **11**, 1017-1029.

Eliscovich, C. and Singer, R. H. (2017). RNP transport in cell biology: the long and winding road. *Curr. Opin. Cell Biol.* **45**, 38-46.

Forrest, K. M. and Gavis, E. R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for *nanos* mRNA localization in *Drosophila*. *Curr. Biol.* **13**, 1159-1168.

Gavis, E. R. and Lehmann, R. (1992). Localization of *nano*s RNA controls embryonic polarity. *Cell* **71**, 301-313.

Gavis, E. R., Curtis, D. and Lehmann, R. (1996). Identification of cis-acting sequences that control *nanos* RNA localization. *Dev. Biol.* **176**, 36-50.

Hanyu-Nakamura, K., Sonobe-Nojima, H., Tanigawa, A., Lasko, P. and Nakamura, A. (2008). Drosophila Pgc protein inhibits P-TEFb recruitment to chromatin in primordial germ cells. *Nature* **451**, 730-733.

Hay, B., Ackerman, L., Barbel, S., Jan, L. Y. and Jan, Y. N. (1988). Identification of a component of *Drosophila* polar granules. *Development* **103**, 625-640.

Houston, D. W. (2013). Regulation of cell polarity and RNA localization in vertebrate oocytes. *Int. Rev. Cell. Mol. Biol.* **306**, 127-185.

Illmensee, K. and Mahowald, A. P. (1974). Transplantation of posterior polar plasm in *Drosophila*. Induction of germ cells at the anterior pole of the egg. *Proc. Nat. Acad.Sci. USA* **71**, 1016-1020.

Jain, R. A. and Gavis, E. R. (2008). The *Drosophila* hnRNP M homolog, Rumpelstiltskin, regulates *nanos* mRNA localization. *Development* **135**, 973-982.

Jambor, H., Mueller, S., Bullock, S. L. and Ephrussi, A. (2014). A stem-loop structure directs *oskar* mRNA to microtubule minus ends. *RNA* **20**, 429-439.

Johnstone, O., Deuring, R., Bock, R., Linder, P., Fuller, M. T. and Lasko, P. (2005). Belle is a *Drosophila* DEAD-box protein required for viability and in the germ line. *Dev. Biol.* **277**, 92-101.

Jongens, T. A., Hay, B., Jan, L. Y. and Jan, Y. N. (1992). The *germ cell-less* gene product: a posteriorly localized component necessary for germ cell development in *Drosophila*. *Cell* **70**, 569-584.

Jongens, T. A., Ackerman, L. D., Swedlow, J. R., Jan, L. Y. and Jan, Y. N. (1994). *Germ cell-less* encodes a cell type-specific nuclear pore-associated protein and functions early in the germ-cell specification pathway of *Drosophila*. *Genes Dev.* **8**, 2123-2136.

Kim-Ha, J., Webster, P. J., Smith, J. L. and Macdonald, P. M. (1993). Multiple RNA regulatory elements mediate distinct steps in localization of *oskar* mRNA. *Development* **119**, 169-178.

Kloc, M. and Etkin, L. D. (2005). RNA localization mechanisms in oocytes. *J. Cell Sci.* **118**, 269-282.

Kloc, M., Bilinski, S., Pui-Yee Chan, A. and Etkin, L. D. (2000). The targeting of Xcat2 mRNA to the germinal granules depends on a cis-acting germinal granule localization element within the 3′UTR. *Dev. Biol.* **217**, 221-229.

Kloc, M., Dougherty, M. T., Bilinski, S., Chan, A. P., Brey, E., King, M. L., Patrick, C. W., Jr and Etkin, L. D. (2002). Three-dimensional ultrastructural analysis of RNA distribution within germinal granules of Xenopus. *Dev. Biol.* **241**, 79-93.

Kobayashi, S., Yamada, M., Asaoka, M. and Kitamura, T. (1996). Essential role of the posterior morphogen *nanos* for germline development in *Drosophila*. *Nature* **380**, 708-711.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T. R., Tomancak, P. and Krause, H. M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174-187.

Lehmann, R. (2016). Germ plasm biogenesis–an Oskar-centric perspective. *Curr. Top. Dev. Biol.* **116**, 679-707.

Lerit, D. A. and Gavis, E. R. (2011). Transport of germ plasm on astral microtubules directs germ cell development in *Drosophila*. *Curr. Biol.* **21**, 439-448.

Lerit, D. A., Shebelut, C. W., Lawlor, K. J., Rusan, N. M., Gavis, E. R., Schedl, P. and Deshpande, G. (2017). Germ cell-less promotes centrosome segregation to induce germ cell formation. *Cell. Rep.* **18**, 831-839.

Little, S. C., Tkačik, G., Kneeland, T. B., Wieschaus, E. F. and Gregor, T. (2011). The formation of the Bicoid morphogen gradient requires protein movement from anteriorly localized mRNA. *PLoS Biol.* **9**, e1000596.

Little, S. C., Sinsimer, K. S., Lee, J. J., Wieschaus, E. F. and Gavis, E. R. (2015). Independent and coordinate trafficking of single *Drosophila* germ plasm mRNAs. *Nat. Cell Biol.* **17**, 558-568.

Lorenz, R., Bernhart, S. H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F. and Hofacker, I. L. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26.

Macdonald, P. M. and Kerr, K. (1998). Mutational analysis of an RNA recognition element that mediates localization of *bicoid* mRNA. *Mol. Cell. Biol.* **18**, 3788-3795.

Macdonald, P. M., Kerr, K., Smith, J. L. and Leask, A. (1993). RNA regulatory element BLE1 directs the early steps of *bicoid* mRNA localization. *Development* **118**, 1233-1243.

Mahowald, A. P. (2001). Assembly of the *Drosophila* germ plasm. *Int. Rev. Cytol.* **203**, 187-213.

Niepielko, M. G., Eagle, W. V. I. and Gavis, E. R. (2018). Stochastic seeding coupled with mRNA self-recruitment generates heterogeneous *Drosophila* germ granules. *Curr. Biol.* **28**, 1872-1881.

Pae, J., Cinalli, R. M., Marzio, A., Pagano, M. and Lehmann, R. (2017). GCL and CUL3 control the switch between cell ineages by mediating localized degradation of an RTK. *Dev. Cell* **42**, 130-142.

Rabani, M., Kertesz, M. and Segal, E. (2008). Computational prediction of RNA structural motifs involved in posttranscriptional regulatory processes. *Proc. Natl. Acad. Sci. USA* **105**, 14885-14890.

Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877-879.

Rangan, P., DeGennaro, M., Jaime-Bustamante, K., Coux, R.-X., Martinho, R. G. and Lehmann, R. (2009). Temporal and spatial control of germ-plasm RNAs. *Curr. Biol.* **19**, 72-77.

Sarov, M., Barz, C., Jambor, H., Hein, M. Y., Schmied, C., Suchold, D., Stender, B., Janosch, S., Vikas KJ, V., Krishnan, R. et al. (2016). A genome-wide resource for the analysis of protein localisation in *Drosophila*. *eLife* **5**, e12068.

Schisa, J. A. (2012). New insights into the regulation of RNP granule assembly in oocytes. *Int. Rev. Cell. Mol. Biol.* **295**, 233-289.

Shahbabian, K. and Chartrand, P. (2012). Control of cytoplasmic mRNA localization. *Cell. Mol. Life Sci.* **69**, 535-552.

Slaidina, M. and Lehmann, R. (2017). Quantitative differences in a single maternal factor determine survival probabilities among *Drosophila* germ cells. *Curr. Biol.* **27**, 291-297.

Theurkauf, W. E., Baum, H., Bo, J. and Wensink, P. C. (1986). Tissue-specific and constitutive alpha-tubulin genes of *Drosophila melanogaster* code for structurally distinct proteins. *Proc. Natl. Acad. Sci. USA* **83**, 8477-8481.

Thomson, T., Liu, N., Arkov, A., Lehmann, R. and Lasko, P. (2008). Isolation of new polar granule components in *Drosophila* reveals P body and ER associated proteins. *Mech. Dev.* **125**, 865-873.

Timinszky, G., Bortfeld, M. and Ladurner, A. G. (2008). Repression of RNA polymerase II transcription by a *Drosophila* oligopeptide. *PLoS ONE* **3**, e2506.

Trcek, T., Grosch, M., York, A., Shroff, H., Lionnet, T. and Lehmann, R. (2015). *Drosophila* germ granules are structured and contain homotypic mRNA clusters. *Nat. Commun.* **6**, 7962.

Zhou, Y. and King, M. L. (1996). Localization of *Xcat-2* RNA, a putative germ plasm component, to the mitochondrial cloud in Xenopus stage I oocytes. *Development* **122**, 2947-2953.