The Groucho co-repressor inhibits progression through the early transcription elongation checkpoint to repress gene expression

María Lorena Martínez Quiles and Barbara H Jennings

Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, United Kingdom.

ABSTRACT

The Groucho/Transducin-Like Enhancer of split (Gro/TLE) family proteins act as co-repressors for many transcription factors that act in several key pathways during development and cancer. Recent studies have shown that Gro is enriched at a subset of genes that exhibit RNA polymerase II (RNAP II) pausing in *Drosophila melanogaster* cells and that depletion of Gro can lead to pause release. Here we compare Gro requirement across three cell lines and find that although Gro target sites are largely cell-type specific, Gro is recruited across a similar profile of genomic features. Gro recruitment frequently overlaps that of general factors which promote RNAP II pausing but it does not preclude recruitment of factors that mediate release of paused RNAP II, indicating that Gro restricts the activity of pause release factors rather than their recruitment. Genetic interaction analyses reveal that the association of Gro with factors known to promote RNAP II pausing in *Drosophila* cell lines reflects a functional relationship in vivo. Knock-down of factors that promote RNAP II pausing enhance *gro* knock-down phenotypes during wing vein patterning. However, knocking down these factors in the eye concurrently with *gro* had different effects on patterning (enhancement) and growth (rescue), indicating that the molecular mechanisms underlying Gro function differ in these two processes. Our results identify a mechanism through which developmentally regulated transcription factors can establish RNAP II pausing to modulate transcription.

INTRODUCTION

The expression of key genes involved with cell fate determination during development is often controlled by spatially restricted localization or activity of transcriptional repressors. Many repressors do not have intrinsic repressive activity but recruit co-factors (co-repressors) that inhibit productive transcription.

The Groucho (Gro) family of co-repressors are recruited by many different families of transcription factors that act during animal development. This family includes a single ortholog in *Drosophila* (Gro), and four orthologs in humans (Transducin-Like Enhancer of split: TLE1-4) and mouse (Gro-related-gene: Grg1-4) which are recruited by a plethora of transcription factors that include Hes, Runx, Nkx, LEF1/Tcf, Pax, Six, and Fox proteins. They are key effectors of cell fate specification signalling pathways, including Notch, EGFR and Wnt, and participate in a diverse range of biological processes including neurogenesis, somitogenesis, osteogenesis haematopoiesis, stem cell maintenance and proliferation and cancer pathogenesis (Agarwal et al., 2015; Jennings and Ish-Horowicz, 2008).

Gro-mediated repression during development is frequently dynamic and rapidly reversible. For example, the production of neuroblasts in the *Drosophila* embryo depends on five short pulses of repression by E(spl)-bHLH/Gro complexes that occur within 4 hours (Jennings et al., 1994; Paroush et al., 1994) Although *Drosophila* Gro acts as a co-repressor, its binding to chromatin is not restricted to inactive chromatin. Profiling genome-wide recruitment of Gro in *Drosophila* cell culture cells and embryos revealed that it is enriched in chromatin that is permissive for transcription (Chambers et al., 2017; Kaul et al., 2014). Furthermore, when Gro is recruited to transcription start sites (TSSs), it does not block RNA polymerase II (RNAP II) recruitment or completely silence transcription. Thus, the properties of Gro-mediated repression are very different to those of the most intensively studied repressors; for example, the Polycomb Repressive Complexes (Blackledge and Klose, 2021).

Transcription is initiated by recruitment of the pre-initiation complex (PIC) that includes RNAP II and assorted general transcription factors. After clearing the promoter, RNAP II transcribes 20–120 nucleotides and must pass the early elongation checkpoint before transcribing the rest of the gene. This checkpoint can be regulated to repress or activate gene expression (reviewed in: Core and Adelman, 2019; Dollinger and Gilmour, 2021; Jonkers and Lis, 2015). Briefly, two protein complexes containing Spt4 and Spt5 (often referred to as 'DSIF') and the other known as Negative ELongation Factor (NELF), act together to inhibit transcript elongation beyond 50 nucleotides. For further elongation to occur, the Cdk 9 kinase subunit of Positive Transcription Elongation Factor b (P-TEFb) must

phosphorylate specific residues in NELF, Spt5 and RNAP II. This induces the dissociation of NELF from the polymerase complex, the switch in Spt5 from a negative to positive regulator of transcription and subsequent production of the full-length transcript by RNAP II. Where this checkpoint is rate-limiting for transcription, RNAP II accumulates at the 5' end of the gene and the RNAP II is described as "paused”.

The phenomenon of RNAP II pausing is widespread in *Drosophila* and mammalian cells, especially at genes involved in developmental and stimulus-responsive pathways yet very little is known about how developmental cues establish RNAP II pausing in a gene-specific manner (Core and Adelman, 2019; Gaertner and Zeitlinger, 2014; Jennings, 2013). While promoter swapping experiments in *Drosophila* indicated that sequences around the promoter are sufficient to invoke RNAP II pausing (Lagha et al., 2013), a model for how promoter elements regulate RNAP II pausing remains elusive (Dollinger and Gilmour, 2021).

Several observations support a model where developmentally regulated transcription factors that recruit Gro can establish RNAP II pausing in a gene-specific manner. Previously, we and others have shown that Gro is frequently recruited to genes that exhibit RNAP II pausing in *Drosophila* cell lines and embryos (Chambers et al., 2017; Kaul et al., 2014) However, depletion of Gro does not lead to global effects on RNAP II pausing indicating that Gro is not part of the core machinery regulating pausing, but rather a factor that can regulate pausing at specific targets (Kaul et al., 2014). For example, depletion of Gro from *Drosophila* Kc167 cells does not change the overall amount of RNAP II recruited at the TSS at a known target gene *E(spl)HLH-mβ*, but it does lead to an increase in RNAP II Ser2-P (the elongating form of RNAP II) and corresponding release of RNA polymerase II pausing and increased full-length transcript production (Kaul et al., 2014).

Here, we extend our previous analysis of the genome-wide recruitment of Gro to show that recruitment to open chromatin is a general feature and while there are common Gro target sites across cell types, most Gro recruitment is cell-type specific. In addition, Gro recruitment frequently overlaps that of factors known to regulate RNAP II pausing and *gro* genetically interacts with genes encoding these factors during both wing and eye development. Taken together, our results indicate that Gro inhibits activation of P-TEFb to restrict the production of full-length transcripts at target genes.

RESULTS

Genome-wide profile of Gro recruitment in ML-DmBG3-c2 cells

We have previously profiled the recruitment of Gro in Kc167 and S2R+ cell lines (Kaul et al., 2014) Both of these cell lines were originally derived from *Drosophila* embryos and have properties related to plasmatocytes (phagocytic blood cells) (Cherbas et al., 2011). To determine which features of Gro recruitment are shared with cells derived from a different tissue, we have extended our analysis of Gro recruitment to the ML-DmBG3-c2 (BG3) cell line. BG3 cells were derived from the central nervous system (CNS) of third instar larvae (Cherbas et al., 2011) and were characterized extensively for genome-wide chromatin modifications as part of the modENCODE project (Kharchenko et al., 2011).

To profile genome-wide Gro binding in BG3 cells, we performed ChIP-seq using an anti-Gro antibody as described previously for Kc167 and S2R+ cells (Kaul et al., 2014). Gro binding sites were determined by combining peaks with a q value of < 0.1 that were present in two biological replicates (see materials and methods for further details). Comparison of Gro peaks in BG3, Kc167 and S2R+ cells revealed that Gro target sites are largely cell-type specific (Figure 1A). There were 831 (from total 1248) peaks unique to BG3 cells and only 175 peaks that were found in all three cell types. 133 of the Gro peaks in BG3 cells overlap Gro peaks in Kc167 cells but not peaks in S2R+ cells. Conversely, 109 peaks found in BG3 cells overlap with peaks in S2R+ cells but not Kc167 cells.

Although Gro is predominantly recruited to targets in a cell-specific manner, Gro is recruited across a similar profile of genomic features in BG3, Kc167 and S2R+ cells (Figure 1B). Gro peaks are predominantly located in the promoter and intronic regions of genes in all three cell types. This profile of binding is consistent with that previously seen in embryos (Chambers et al., 2017). In addition, Gro binds in discrete peaks rather than over broad domains in BG3 cells, consistent with previous observations for Kc167 and S2R+ cells (Kaul et al., 2014). Gro peaks in BG3 cells have an average peak width of 278 bp and a median peak width of 224 bp (Figure 1C, materials and methods). Thus, the pattern of Gro binding in all 3 cell lines is inconsistent with the historical model in which Gro spreads across the genome from the initial site of recruitment by a transcription factor (Chen and Courey, 2000).

Gene Ontology analysis indicated that the Gro peaks in BG3 cells were associated with transcripts linked to a variety of developmental processes including imaginal disc derived wing morphogenesis, open tracheal system development, and dendrite morphogenesis, in addition to transcripts linked with transcriptional regulation (Figure 1D). The DNA sequences under Gro peaks in BG3 cells (Figure 1E) were enriched for GAGA Factor (GAF) binding

sites in common with Kc167 and S2R+ cells, and l(3)neo38 sites in common with S2R+ cells (Kaul et al., 2014). There was also enrichment for Crooked legs (Crol), Odd-paired (Opa) and Longitudinals lacking (Lola) sites that were not so frequently observed in Kc167 or S2R+ cells (Kaul et al., 2014).

The limited overlap of Gro peaks and association with different profiles of target sequence across these three cell lines reveal that the binding of Gro is largely specific to cell type. However, Gro is predominantly recruited to discrete peaks located in promoter and intronic regions of genes in all three cell types.

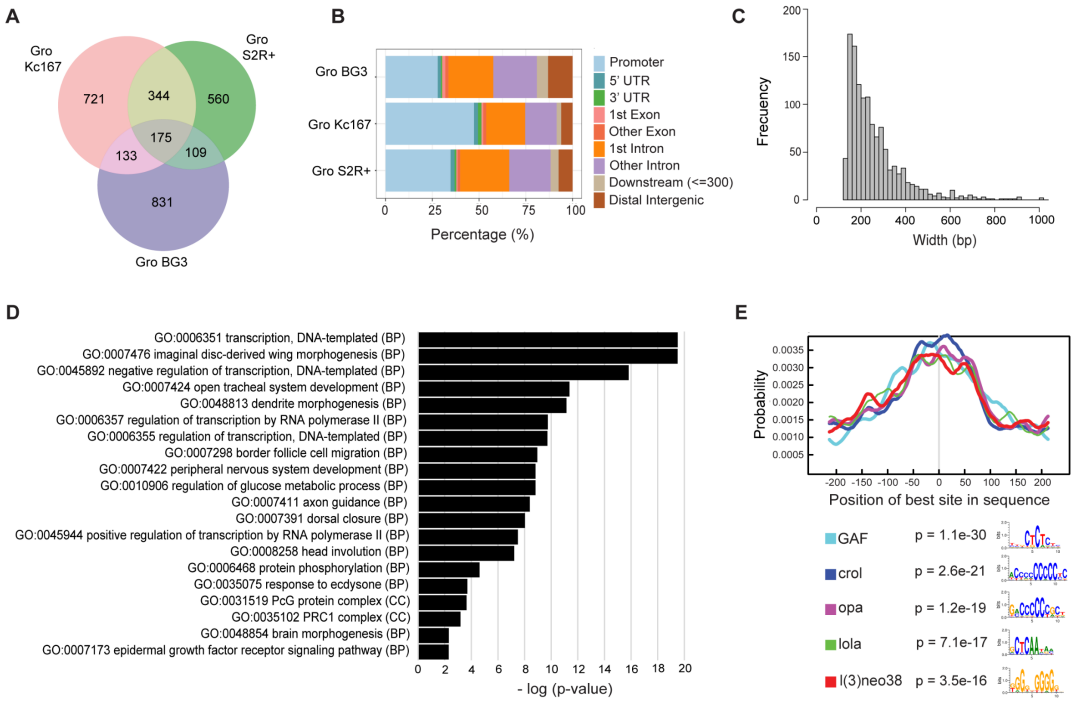


Figure 1: Characterization of high confidence Gro binding sites in BG3 cells and comparison among Gro peaks in three different cell lines.

A) Venn diagram visualizing the overlap among Gro peaks in the Kc167, S2R+ and BG3 cell lines. Overlap was defined by a minimum and reciprocal fraction of 5% overlap across peaks. B) Bar graph showing the percentage of Gro peaks within each genomic annotation in the Kc167, S2R+ and BG3 cell lines. Promoters were defined as ±250 bp with respect to TSSs. C) Histogram showing the frequency of the width peaks of Gro binding sites. D) DAVID Gene Ontology (Huang et al., 2009a, 2009b) analysis of genes associated with Gro binding sites in BG3 cells. Terms plotted were selected by taking the most significant term in a cluster (p-value < 10-5) and the most significant unclustered terms. E) Graph showing the top-scoring five results from the CentriMo analysis at Gro binding sites in BG3 cells: GAF, Crol, Opa, Lola, and l(3)neo38 (Bailey and Machanick, 2012). Each curve represents the probability of the binding location for the named factors at each position of Gro ChIP-seq peak regions (500 bp). The legend shows the motifs and its central enrichment p-value.

Gro is enriched in open "Enhancer" type chromatin

To gain insight into the molecular mechanism underlying Gro-mediated repression, we analysed the chromatin environment associated with Gro peaks. Although it is well established that Gro acts as a repressor, recent studies have indicated that Gro predominantly binds in chromatin that permits transcription. The integrative analysis of the binding profiles of 53 proteins (tagged with DamID) that associate with chromatin in Kc167 cells produced a model in which the *Drosophila* genome contains five principal chromatin types; ''Red'' (active, developmentally regulated), ''Yellow'' (active, housekeeping), ''Blue'' (repressed, by Polycomb Group complexes) ''Green'' (repressed, classic heterochromatin), and ''Black'' (highly repressed) (Filion et al., 2010). We previously observed that Gro recruitment is most highly enriched regions of the genome classed as "red" active chromatin in Kc167 cells by (Filion et al., 2010; Kaul et al., 2014). In addition, Gro recruitment was shown to be associated with DNase I hypersensitivity sites across the genome, indicating that Gro is recruited to chromatin that is open and where to turnover rate of nucleosomes is high, facilitating transcription (Cockerill, 2011). To determine if binding in "active" chromatin is a general feature of Gro recruitment and to evaluate which chromatin marks are associated with Gro binding, we extended this analysis by looking additional data sets in Kc167 cells and by expanding our analysis into BG3 cells.

ATAC-seq has recently become a popular method to assay for regions of open chromatin across the genome (Buenrostro et al., 2013) and acts as an alternative method to DNase I hypersensitivity analysis. Analysis of ATAC-seq data for Kc167cells (Porcelli et al., 2019) confirmed our previously observation that Gro is predominantly recruited to open chromatin (Figure 2A); just 4.4% of Gro peaks did not overlap ATAC-seq peaks (Supp Figure 1). Having established that Gro is recruited to open chromatin in Kc167 cells, we extended our analysis to determine if this holds true across different cell types by analysing DNase I hypersensitivity data from BG3 cells. Akin to Kc167 cells, Gro peaks were also found in regions associated with DNase I hypersensitivity in BG3 cells (Figure 2B). Thus, despite acting to repress transcription, Gro predominantly binds in open chromatin and does not induce the chromatin to compact and become in accessible to the transcription machinery.

To gain further insight into the nature of the chromatin that Gro is recruited to, we profiled Gro peaks with respect to the genome-wide maps of chromatin states in Kc167 and BG3 cells generated by (Skalska et al., 2015). These maps were generated from data from 24 different chromatin modifications and DNaseI hypersensitive sites, and led to the definition of 11 chromatin states using an adaptation of the Hidden Markov model (HMM) approach first published by (Kharchenko et al., 2011). We observed that Gro peaks are most highly

enriched in “Enhancer” (Enh) chromatin (red) in both BG3 and Kc167 cells (Figure 2C and D). Enh chromatin is characterised by the H3K4me1, H3K27ac and H3K56ac modifications and DNAseI hypersensitivity and is one the three defined states of active regulatory chromatin. Around 80% of Gro peaks mapped to Enh chromatin in BG3 cells and over 60% of the superset of high confidence Gro peaks in Kc167 cells defined by (Kaul et al., 2014). Gro binds predominantly in Enh (red) chromatin at the established target genes, *E(spl)mβ-HLH* and *E(spl)m3-HLH* in both cell lines (Figure 2E) (Kaul et al., 2014). There was also some enrichment of Gro peaks in “Active TSS” (orange) in Kc167 cells, however Gro peaks are largely excluded from chromatin considered to be relatively closed and unavailable for transcription including “Polycomb” (blue), “Heterochromatin” (light grey) and “Basal” (black) chromatin (Figure 2C, D).

Taken together, all these observations indicate that Gro is primarily recruited to chromatin that is accessible and permissive for transcription, rather than chromatin that is compacted and repressed. Since Gro does not promote a repressive chromatin state, we proceeded to investigate how Gro can negatively affect the activity of transcription complexes at target genes.

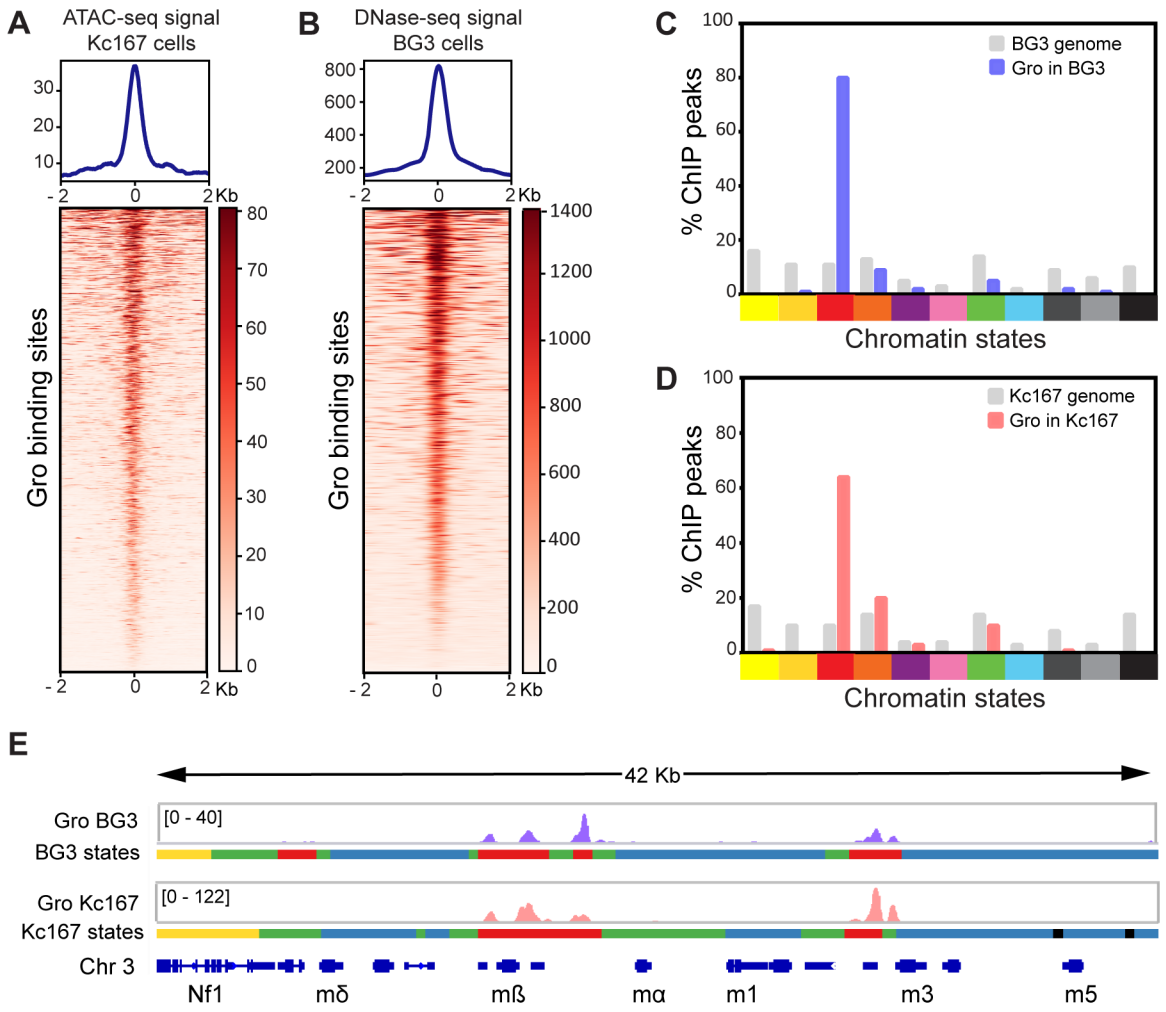


Figure 2: Gro binding sites are present at regions of open chromatin that are enriched in "enhancer" chromatin in Kc167 and BG3 cells.

A) Average profile plot and heatmap showing the ATAC-seq signal at Gro peaks in Kc167 cells. B) Average profile plot and heatmap showing the DNase-seq signal at Gro peaks in BG3 cells. The average plots and the heatmaps are centred at the centre of Gro peaks and extended 2 Kb upstream and downstream of that reference point. C) Distribution of Gro binding sites according to their chromatin type in BG3 and D) Kc167 cells. Chromatin states are coloured according to the scheme defined by (Skalska et al., 2015). Gray bars indicate the proportion of each chromatin signature in the total *Drosophila* genome in each cell line. E) Gro binding profiles in BG3 and Kc167 cells along the *E(spl)-C* locus aligned with the chromatin map colour coded as defined by (Skalska et al., 2015).

Gro co-localises with factors known to regulate RNAP II pausing across the genome

Previous studies have revealed that Gro recruitment is associated with target genes that exhibit promoter proximal RNAP II pausing and depletion of Gro can lead to pause release (Chambers et al., 2017; Kaul et al., 2014). These observations suggest that Gro acts with general factors that promote RNAP II pausing to limit transcription of target genes or by inhibiting the recruitment or activity of factors that activate the release of paused RNAP II. To determine if Gro is recruited to the genome in close association with regulators of RNAP II pausing, we examined whether peaks of Gro binding to chromatin overlap with those of factors known to promote pausing and pause release.

Recruitment of the NELF complex is linked with RNAP II pausing at genes in *Drosophila* (Lee et al., 2008) and acts to promote pausing (Yamaguchi et al., 1999). We examined the overlap of Gro peaks with the Nelf-E peaks generated by ChIP-seq in Kc167 cells by (Nazer et al., 2018). We observe that 64.7 % of Gro peaks overlap with Nelf-E across the genome and that 91.2% of Gro peaks located at TSSs overlap with Nelf-E (Figure 3A), consistent with a model in which Gro acts to promote RNAP II pausing.

GAF is also linked to promoter proximal pausing at many genes in *Drosophila* and acts to promote pausing in part by recruitment of NELF (Li et al., 2013). Using GAF peaks derived from ChIP-seq data generated as part of the modENCODE project, we previously found that 82% of Gro peaks in Kc167 cells located at TSSs overlap with peaks of GAF (Kaul et al., 2014). In BG3 cells, we observe that 51% of Gro peaks overlap with GAF peaks, and this percentage is increased to 56% at TSSs (Figure 3B). This is consistent with the observation that GAF binding sites are frequently found at Gro peaks in both these cell lines [Figure 1E; (Kaul et al., 2014), and a model in which Gro interacts with GAF to promote pausing at many, but not all target genes.

While GAF and NELF promote RNAP II pausing, release of the paused polymerase into productive elongation is driven by the P-TEFb complex, comprised of Cdk9 and Cyclin T (Peterlin and Price, 2006). Cdk9 is a kinase and acts to phosphorylate RNAP II, NELF and Spt5 on specific residues to release paused RNAP II into the elongation phase of transcription. One potential mechanism through which Gro could promote RNAP II pausing is by inhibiting the recruitment of P-TEFb to target genes. In this model, we would expect to see Cdk9 and Cyclin T peaks excluded from TSSs that bind Gro. We examined the overlap of Gro peaks with the Cdk9 ChIP-seq peaks in Kc167 cells (Nazer et al., 2018) and found 75.5 % of Gro peaks overlapped with Cdk9 peaks at TSSs (Figure 3C) and 46.6 % of Gro

peaks overlapped with Cyclin T (Schaaf et al., 2013) at TSSs in BG3 cells (Figure 3D). Thus, Gro does not appear to promote RNAP II pausing primarily by blocking P-TEFb recruitment.

To visualise the profile of pausing factors recruited with Gro at TSSs, we generated an Upset plot of overlapping peaks using published data from Kc167 cells for Nelf-E, Cdk9, GAF and RNAP II (Khan and Mathelier, 2017). Most frequently, Gro peaks overlapped with peaks of all these factors (Figure 3E; 306 of 649 Gro peaks). This combination of peaks is observed at three genes that are established Gro targets; *E(spl)mβ-HLH, E(spl)m3-HLH* and *aos* (Figure 3F) (Housden et al., 2014; Kaul et al., 2014). The second most frequent combination observed included Nelf-E, Cdk9 and RNAP II peaks, but not peaks of GAF (Figure 2E; 90 of 649 Gro peaks).

Taken together, these results indicate that Gro frequently co-localizes with factors previously shown to promote and release RNAP II pausing at TSSs. The co-localization of Gro with Nelf-E in addition to GAF is consistent with our previous observations that Gro can promote RNAP II pausing (Kaul et al., 2014). However, peaks of Cdk9 and Cyclin T are found to overlap Gro peaks at TSSs of known Gro target genes, inconsistent with a model in which Gro promotes RNAP II pausing by blocking P-TEFb recruitment.

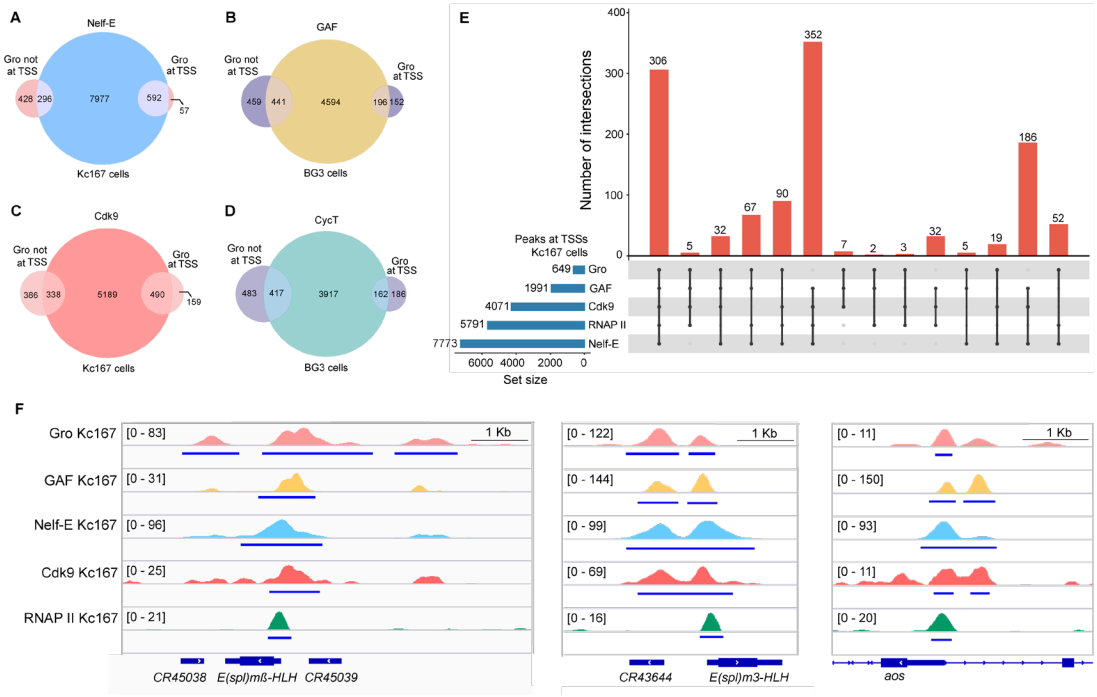


Figure 3: Gro co-occupies genomic locations with Nelf-E, GAF Cdk9, CycT and RNAP II.

A) Venn diagram showing the overlap between Gro and Nelf-E peaks located at TSSs and at all other sites (not at TSSs) in Kc167 cells. B) Venn diagram illustrating the overlap between Gro and GAF peaks located at TSSs and at all other sites (not at TSSs) in BG3 cells. C) Venn diagram showing the overlap between Gro and Cdk9 peaks located at TSSs and at all other sites (not at TSSs) in Kc167 cells. D) Venn diagram illustrating the overlap between Gro and CycT peaks at, and not at, the TSSs in BG3 cells. E) UpSet plot showing the intersections among peaks at TSSs of the following data sets: Gro, Nelf-E, Cdk9, GAF, and total RNAP II (Rpb3) The connected lines in the lower panel of the plot and the upper bar graph (in red) represent respectively the sets of peaks included in each intersection and its corresponding number of overlapping peaks. The horizontal left bar graph represents the total number of peaks in each data set. F) Genome browser views for Gro, GAF, Nelf-E, Cdk9, and total RNAP II (Rpb3) at *E(spl)mb-HLH*, *E(spl)m3-HLH*, and *argos* (*aos*).

Establishment of an assay to test for genetic interactions with *gro* during wing development

To test whether the association of Gro with factors that regulate RNAP II pausing in cell culture cells reflects a functional relationship in vivo, we tested for genetic interactions between *gro* and genes encoding factors that regulate pausing using the GAL4/UAS-RNAi system (Heigwer et al., 2018).

We used a combination of *nub-GAL4* and *UAS-RNAi-gro* to knock down *gro* expression in the developing wing (Figure 4). *nub-GAL4* is expressed in the developing wing pouch region of the wing imaginal disc (Csordás et al., 2020). Flies carrying a single copy of each of *nub-GAL4* and *UAS-RNAi-gro* exhibited mild but consistent defects in wing patterning (Figure 4B). We observed ectopic venation above the L2 vein, frequent loss of the anterior cross vein (ACV), and weak longitudinal vein spreading (Figure 4B). As expected, flies carrying 2 copies of *nub-GAL4* and 2 copies of *UAS-RNAi-gr*o exhibited stronger phenotypes than the those carrying single copies (Figure 4C). We observed thickened longitudinal veins that spread out further at the wing margin, ectopic veins in the posterior half of the wing (in addition to the region above L2) and complete loss of the ACV (Figure 4C). The vein thickening and ectopic vein phenotypes are consistent with the phenotype of *gro* loss of function clones in the wing (Celis and Ruiz-Gómez, 1995) and the established model that *gro* acts as an effector of the Notch signalling pathway to repress vein cell fate during wing development (Hasson and Paroush, 2006). The loss of the ACV indicates that Gro functions in ACV patterning, which occurs during pupal development after patterning of the longitudinal veins (Blair, 2007).

To test the ability of the *UAS-RNAi-gro* line to knock down *gro* expression, we tested its ability to rescue the phenotype produced by overexpression of Gro via *UAS-gro*. Flies carrying *nub-GAL4* and *UAS-gro* exhibited small, blistered wings (Figure 4D). This phenotype was largely rescued by a single copy of *UAS-RNAi-gro*, indicating that the *UAS-RNAi-gro* transgene targets *gro* mRNA (Figure 4E). After the initial characterisation the *nub-GAL4; UAS-RNAi-gro* phenotypes, we established a system to classify the severity of the wing phenotypes observed to facilitate our analysis of genetic interactions using GAL4/UAS-RNAi lines (Figure 4F,G).

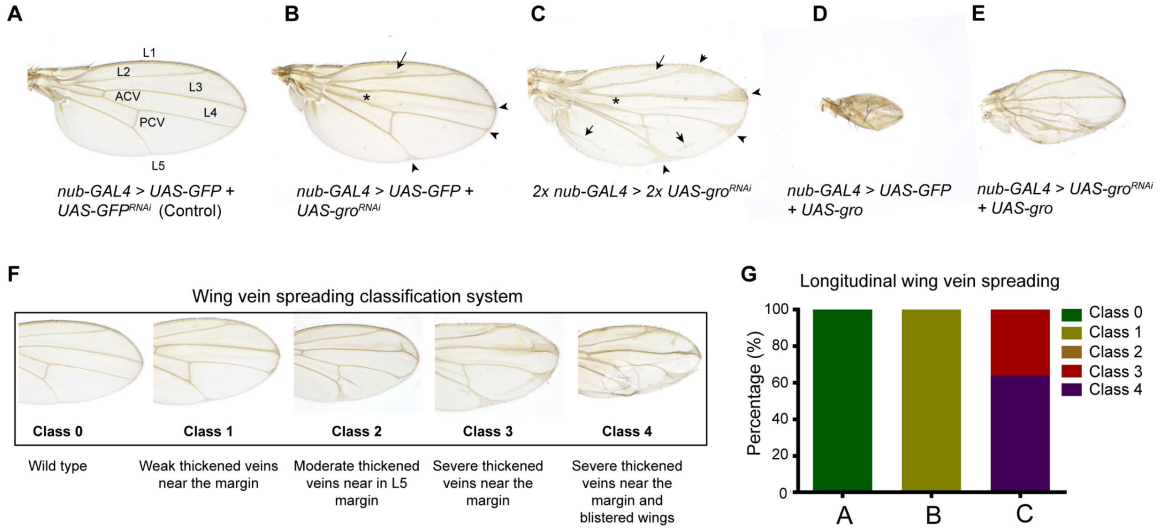


Figure 4: Characterization of *gro* RNAi knockdown during wing development. Representative examples of *Drosophila* adult female wings raised at 25°C of the indicated genotypes. (A) Control wing showing key structures: anterior cross vein (ACV), posterior cross vein (PCV), and longitudinal veins L1 to L5. (B) Heterozygous *gro* RNAi knockdown showing ectopic venation above L2 (arrow), lack of ACV (asterisk) and weak longitudinal vein spreading (arrowheads). (C) Homozygous *gro* RNAi knockdown showing ectopic venation above L2 (arrow), increase of ectopic venation in the posterior half of the wing (arrows), loss of the ACV (asterisk), and severe longitudinal vein spreading (arrowheads). (D) *gro* overexpression showing a small and blistered wing. (E) Simultaneous *gro* knockdown and overexpression showing partial rescue of the wing size and wing vein pattern. (F) Adult wing images that are representative for each of the classes in the classification system and a brief description of their phenotypes. The classification system is based on the severity of the longitudinal vein spreading. There are four classes, which represent increasing degrees of severity (Class 0, 1, 2, 3 and 4). (G) Graph showing the frequency of each class observed for the following phenotypes: control [A], single *gro* knockdown [B], and double *gro* knockdown [C]. At least 70 wings of each genotype were scored for quantification in panel, except for c for which only 25 wings were scored due to the difficulty of obtaining viable flies.

Knock-down of genes encoding factors that promote RNAP II pausing enhance *gro* phenotypes during wing development

Having established a genetic assay to test for interactions with *gro,* we used this strategy to test genes encoding factors known to promote RNAP II pausing. The prediction from the model in which Gro promotes RNAP II pausing is that lowering expression levels of these factors (by *UAS-RNAi*) will lead to an enhancement of the single *nub-GAL4:UAS-RNAi-gro* phenotype.

We have observed that peaks of Gro recruitment to the genome frequently overlap peaks of GAF recruitment in both Kc167 and BG3 cells (Figure 3B)(Kaul et al., 2014), but it is not known if this co-localization has functional significance *in vivo*. The expression of *Trl* (*Trl* encodes GAF) was knocked down alone and in combination with *gro* using the *nub-GAL4>UAS-RNAi* assay (Figure 5A). Consistent with previously published work (Blanch et al., 2015), knock-down of GAF expression in the developing wing by *UAS-RNAi-Trl* resulted in a reduction of the wing blade size but had very little effect on wing vein patterning. However, when both *gro* and *Trl* were knocked down, flies exhibited stronger vein patterning defects than those observed in those where *gro* was knocked down alone (Figure 5A). We observed an increase in longitudinal vein spreading (quantified in Figure 5B), ectopic venation and loss of the ACV (quantified in Figure 5C), among other effects. The phenotype was highly reminiscent of that observed in flies carrying 2 copies of *nub-GAL4* and *UAS-RNAi-gro* (Figure 5B and C). Knockdown of *Trl* enhances the *gro* phenotype, indicating that GAF acts with Gro during vein patterning.

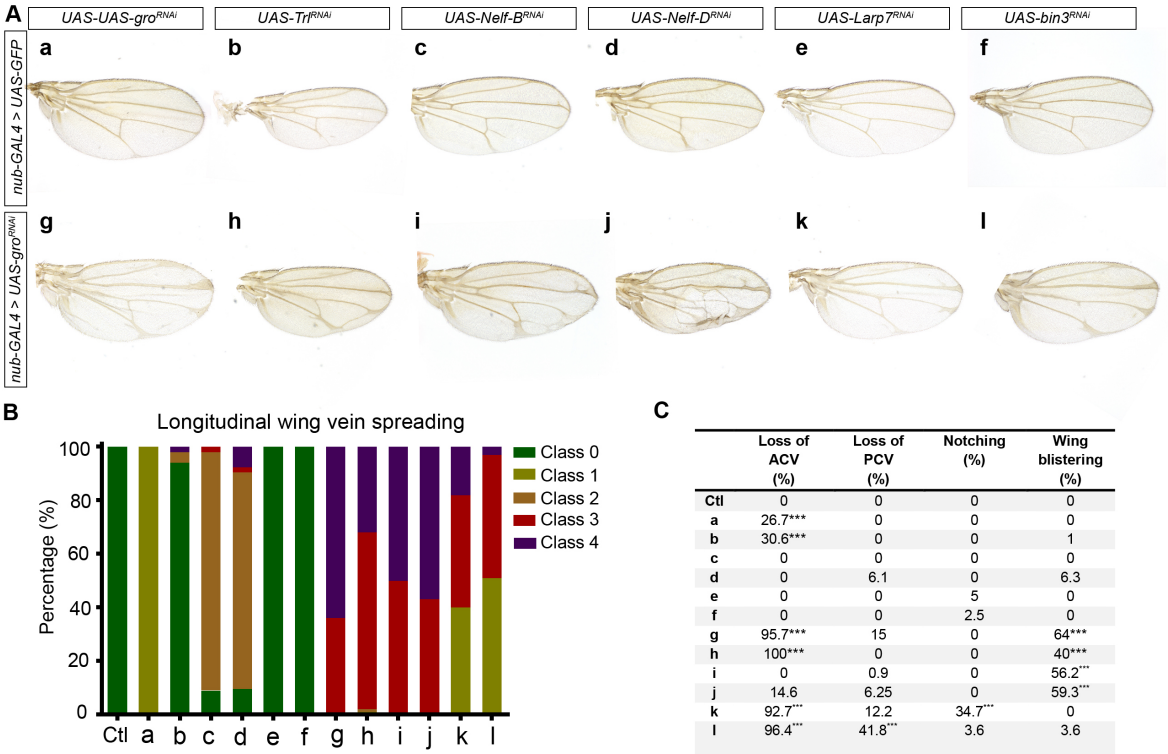
We tested the knock-down of all 4 subunits of *Drosophila* NELF (A, B, D and E) for interactions with *nub-GAL4:UAS-RNAi-gro.* In this assay, the knock-down of either *Nelf-B* or *Nelf-D* alone resulted in weak but consistent phenotypes, and both significantly enhanced the *nub-GAL4:UAS-RNAi-gro* phenotype (Figure 5A, B, C). The knock-down of either *Nelf-A* or *Nelf-E* did not result in a visible phenotype or strong modification of the *nub-GAL4:UAS-RNAi-gro* phenotype (Supp Figure 2). However, we observed a weak increase in vein thickening in the double knock-down *gro* and *Nelf-A.* The different effects of knocking down individual subunits of NELF could reflect differential requirements of each for Gro function or due to different knock down efficiencies of the *UAS-RNAi* lines.

The 7SK small non-coding RNA (snRNA) inhibits RNAP II transcription by binding P-TEFb and recruiting an RNA binding protein HEXIM (HEXIM1 or HEXIM2 in mammals) to block Cdk9 activity (Peterlin et al., 2012). The association of P-TEFb with the 7SK snRNP complex is reversible and regulation of this association is a key mechanism to control P-TEFb activity

in cells. We tested the genes encoding La-related protein 7 (Larp7, a component of the 7SKsnRNA complex), Bicoid-interacting protein 3 (Bin3, an RNA methyltransferase that adds a 5'-methyl cap to the product of snRNA:7SK preventing its degradation) and HEXIM for genetic interactions with *gro as UAS-RNAi lines* were available for these genes.

Knock-down of either *Larp7* or *bin3* alone did not lead to any detectable phenotypes in the wing. However, when either of these genes were knocked-down in combination with *gro,* there was a significant enhancement of the phenotypes observed in *nub-GAL4>UAS-RNAi-gro* flies (Figure 5). Knock-down of HEXIM using *nub-GAL4>UAS-RNAi-HEXIM* was lethal, and no flies were recovered to analyse.

Our observation that *gro* genetically interacts with genes encoding factors known to promote RNAP II pausing in the wing are consistent with a model in which Gro mediates repression, at least in part, by promoting RNAP II pausing in vivo.



	Loss of ACV (%)	Loss of PCV (%)	Notching (%)	Wing blistering (%)
Ctl	0	0	0	0
a	26.7***	0	0	0
b	30.6***	0	0	1
c	0	0	0	0
d	0	6.1	0	6.3
e	0	0	5	0
f	0	0	2.5	0
g	95.7***	15	0	64***
h	100***	0	0	40***
i	0	0.9	0	56.2***
j	14.6	6.25	0	59.3***
k	92.7***	12.2	34.7***	0
l	96.4***	41.8***	3.6	3.6

Figure 5: *gro* genetically interacts with *NELF*, *Trl*, *Larp7*, and *bin3* in the developing wing.

(A) Representative examples of *Drosophila* adult female wings raised at 25°C expressing the indicated transgenes. (B) Graph showing the frequency of each class (as defined in Figure 4) observed for the phenotypes shown in A. (C) Table showing the quantification of four wing features [total or partial loss of anterior cross vein (ACV), total or partial loss of posterior cross vein (PCV) notching, and wing blistering] for the phenotypes shown in A. Fisher's exact tests were used for pairwise statistical comparisons. For each feature, the single knockdowns were compared to the control, and double knockdowns were compared to the single knockdowns of the genes in the double. ***$p<0.0001$. At least 70 wings of each genotype were scored for quantification in B and C.

Establishment of an assay to test for genetic interactions with *gro* during eye development

To test whether the genetic interactions between *gro* and genes encoding pausing factors are unique to the wing or reflect a more general property of Gro-mediated repression, we extended our assay to the developing *Drosophila* eye. We used *eyeless-GAL4* (Hazelett et al., 1998) to drive the expression of *UAS-RNAi-gro* (Figure 6). Flies carrying a single copy of each of *ey-GAL4* and *UAS-RNAi-gro* resulted in weak effects on the adult eye, characterized by small regions of necrosis localized in the posterior side of the eye and minor irregularities in the alignment of the ommatidia to each other (Figure 6A). Furthermore, there is a significant increase in the area of the eye in these flies (quantified in Figure 6B). Two copies of *ey-GAL4: UAS-RNAi-gro* resulted in high lethality. This lethality was most likely a consequence of *gro* knockdown in the embryonic central nervous system and/or larval brain, as *eyeless* is also expressed in these tissues (Quiring et al., 1994). The few flies that were recovered had increased necrosis in the posterior part of the eye and irregular ommatidial packing (Figure 6A) in comparison to the single knock-down.

To validate the *gro* knock-down in the eye, we tested the ability of the *UAS-RNAi-gro* transgene to rescue the phenotype generated by the overexpression of *gro* (Figure 6 A,c). The overexpression of *gro* via *UAS-gro* resulted in a severe reduction in the area of the eye area, but there were no obvious irregularities in remaining ommatidia (Figure 6A, quantified in Figure 6B), consistent with previously published results (Choi et al., 2005). The small eye phenotype caused by *gro* overexpression was partially rescued by co-expression of *UAS-RNAi-gro,* providing further evidence that the *UAS-RNAi-gro* is targeting *gro* expression (Figure 6A, quantified in Figure 6B).

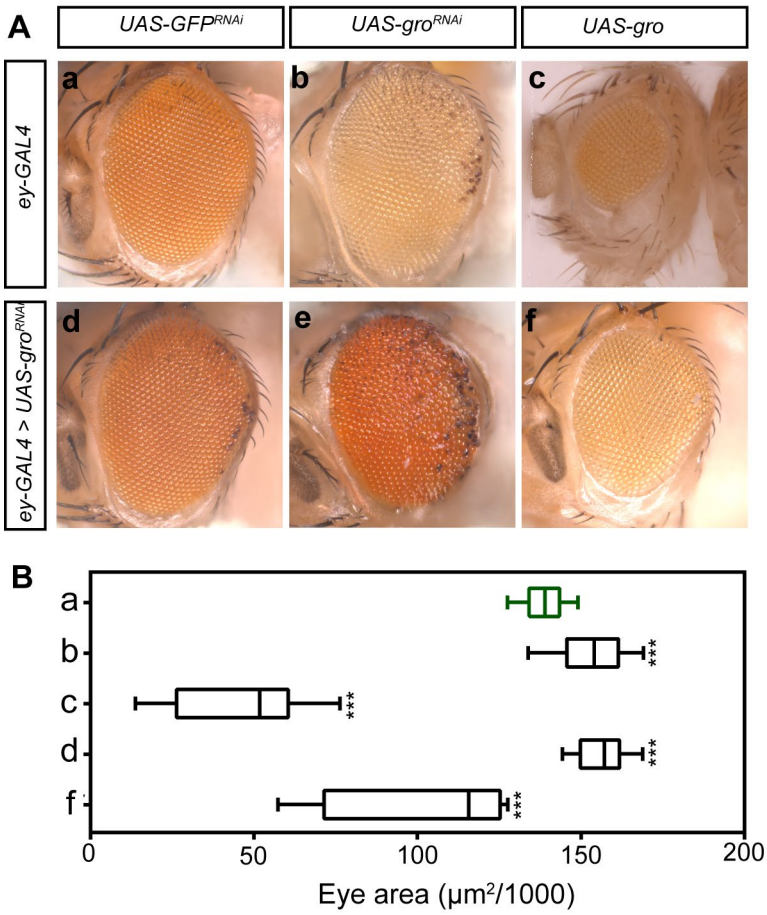


Figure 6: Characterization of *gro* RNAi knockdown during eye development.
(A) Representative examples of *Drosophila* adult female eyes raised at 25°C of expressing the indicated transgenes. (a) Control adult eye (b) single *gro* RNAi knockdown showing an increase in the eye area compared to the control, weak ommatidia disorderliness, and weak cell death in the posterior part of the eye (c) overexpression of *gro* showing a severe reduction in the eye size compared to the control (d) single *gro* RNAi knockdown, additionally expressing a mock RNAi transgene showing an increase in the eye area compared to the control, weak ommatidia disorderliness, and weak cell death in the posterior part of the eye. (e) double *gro* RNAi knockdown showing damage, moderate ommatidia disorderliness, and severe cell death in the posterior part of the eye (arrows). (f) Simultaneous *gro* knockdown and overexpression showing partial rescue of the eye size. (B) Box and whisker plot of eye area (µm2/1000); the Y axis indicates the transgenes expressed in each of the phenotypes under the ey-GAL4 driver. T-test were used for pairwise statistical comparison between the eye area measurements of the indicated genotypes and the and the control (*ey-GAL4/+; UAS-RNAi-GFP/+*). ***$p<0.0001$. At least 10 eyes of each genotype

were measured. The area measurement of the double *gro* knockdown phenotype (e) was excluded from the plot due to the difficulty to obtain viable flies of that genotype.

***gro* genetically interacts with genes encoding regulators of RNAP II pausing in the developing eye.**

We used the genetic assay established in the developing eye to test for interactions between *gro* and the genes we tested in the developing wing: *Trl*, *NELF*, *Larp7* and *bin3*. In addition, we tested *Cdk9*, a component of the P-TEFb complex that is key in the release of the paused RNAP II. We could not test this factor in the wing as the expression of *UAS-RNAi-Cdk9* driven by *nub-GAL4* was lethal.

The individual knockdowns of *Trl*, *Nelf-B*, *Nelf-D*, *Larp7*, and *bin3* (Figure 7A b-f) did not result in any obvious effects on eye morphology or size (quantified in Figure 7B). However, a modest but significant increase in eye size was observed in *ey-GAL4/+; UAS-RNAi-bin3/+* flies. The knock down of *Trl*, *Nelf-B*, *Nelf-D*, *Larp7*, and *bin3* in combination with *gro* resulted in variable phenotypes for each genotype. To simplify the phenotypic analysis, we focused on the flies (≥ 5) exhibiting the most severe eye phenotypes for each condition.

The concurrent knockdown of *gro* and *Trl* did not significantly change the area of the eye in comparison with flies where only expression of *gro* had been knocked down (Figure 7 A, B). In addition, there was no obvious increase in necrosis or ommatidial irregularities. Thus, in contrast to the wing, these genes do not interact in this assay during eye development, indicating that the relationship between *gro* and *Trl* is likely to be tissue specific.

When *gro* was downregulated simultaneously with factors known to promote RNAP II pausing, (*Nelf-B*, *Nelf-D*, *Larp-7*, and *bin3),* there was a noticeable increase in necrosis and irregular packing of ommatidia in comparison with the single *gro* knockdown. (Figure 7A,j-m). The most severe double knock-down phenotype observed was caused by the downregulation of *gro* and *Nelf-D*, which led to lethality. The few flies that hatched had highly irregular ommatidia (Figure 7A,k). In addition, we tested *Nelf-A* and *Nelf-E* in the eye assay. Consistent with the data from the wing, the knockdown of either of these factors alone did not lead to a noticeable phenotype (Supp Figure 3). However, knockdown of *Nelf-A* in combination with *gro* resulted in a slight increase in necrosis on the posterior side of the eye. Expression of all these *UAS-RNAi* lines in combination with *UAS-RNAi-gro* led to rescue of the weak increased eye area phenotype associated with knocking down *gro* (Figure 7B). Thus, knocking down factors that promote pausing in the eye concurrently with *gro* had different effects on patterning (enhancement) and growth (rescue), indicating that the molecular mechanisms underlying Gro function differ in these two processes.

Knockdown of *Cdk9* resulted in a severe reduction of the eye size (Figure 7A,g; quantified in Figure 7B), reminiscent of the phenotype resulting from ectopic expression of *gro* (Figure 6). When *Cdk9* was simultaneously downregulated with *gro* there was a partial rescue of the eye size, and no other obvious effects (Figure 7A,n; quantified in Figure 7B). Thus, Gro and Cdk9 have opposite effects on growth of the adult eye.

The genetic assays indicate that Gro acts with factors that promote RNAP II pausing during patterning and/or differentiation but that they may act to antagonise *gro* function during growth control, suggesting that there is more than one molecular mechanism underlying Gro function. However, our findings from both the bioinformatic and genetic analyses are consistent with a model in which Gro acts to promote RNAP II pausing at many target genes.

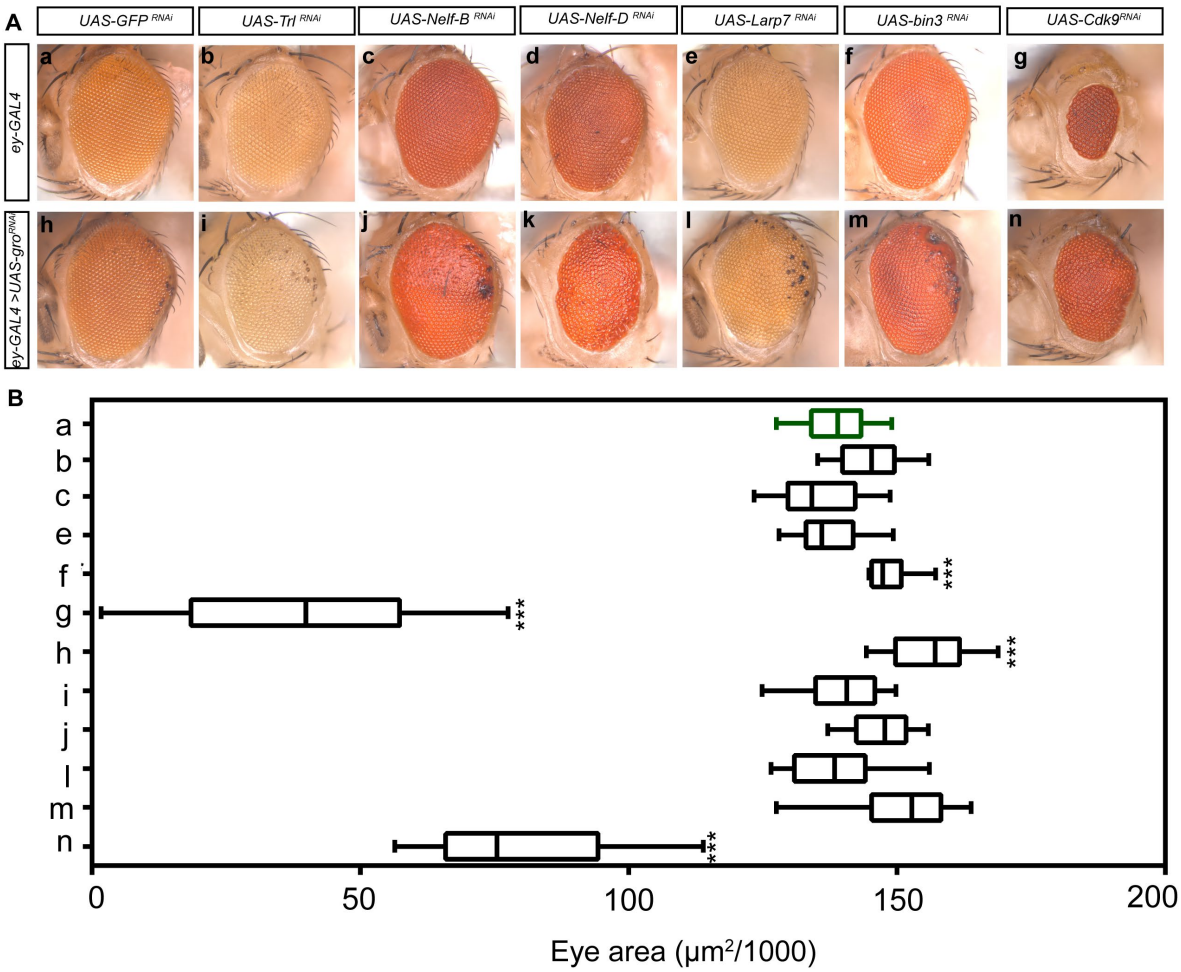


Figure 7: *gro* genetically interacts with *NELF*, *Larp7*, *bin3* and *Cdk9* in the developing eye.

(A) Representative examples of *Drosophila* adult female wings raised at 25ºC expressing the indicated transgenes. (B) Box and whisker plot of eye area (µm2/1000); the Y axis indicates the transgenes expressed in each of the phenotypes under the *ey-GAL4* driver. T-test were used for pairwise statistical comparisons. Eye area measurements of single knockdowns of the indicated genotypes were compared to the control (green in the plot), whereas the double knockdowns were compared to the single knockdown of the genes in the double. ***p<0.0001. At least 10 eyes of each genotype were measured. Measurement of the eye size of *Nelf-D* single and double knockdowns were excluded from the analysis due to the difficulty to obtain double knockdown flies.

DISCUSSION

Gro/TLE proteins have long been recognised as co-repressors for many different transcription factors that act to direct developmental processes. In this paper, we show that while Gro is recruited to genes in a largely cell-type specific manner, there are common attributes to Gro target sites across the genome that give insight into Gro function.

The finding that Gro is predominantly recruited to open chromatin across different cell types is consistent with the long-established model that Gro is recruited by sequence-specific DNA-binding transcription factors. Most transcription factors known to recruit Gro [including the E(spl)-bHLHs, Capicua, and Brinker, involved with vein patterning] are not considered to be pioneer factors as they are unable to bind sequences in highly compacted closed chromatin (Balsalobre and Drouin, 2022; Hasson and Paroush, 2006). The association of Gro with open chromatin is also consistent with observations that Gro repression is rapidly reversible. For example, Gro forms part of the Su(H)/H complex that represses Notch targets in the absence of Notch signalling in *Drosophila* (Barolo et al., 2002; Nagel et al., 2005). Repression by this complex is rapidly reversed (within minutes) when Notch signalling is activated to enable Notch target gene expression (Krejci et al., 2009). The promotion of RNAP II pausing is an example of a repressive mechanism that can be rapidly reversed in response to developmental or environmental cues; one well-established model where this occurs is at heat shock genes (Himanen and Sistonen, 2019; O'Brien and Lis, 1991).

Our analysis of ChIP-seq data reveals that Gro peaks frequently overlap peaks of factors implicated in promoting RNAP II pausing (NELF and GAF), but the presence of Gro does not necessarily block RNAP II or P-TEFb recruitment (Figure 3). For example, we observe overlapping peaks of Cdk9 and Gro binding around the TSS of *E(spl)HLH-mβ* in Kc167 cells (Figure 3F). Depletion of Gro has been shown to induce the release of paused RNAP II and increased transcription of *E(spl)HLH-mβ* in this cell line (Kaul et al., 2014), revealing that Cdk9 is not fully active when Gro is functioning as co-repressor at this locus. Our results support the model in which Gro promotes RNAP II pausing in cooperation with factors that inhibit P-TEFb activity at target loci.

While studies in homogeneous cell populations have given valuable insight into the underlying molecular mechanisms and prevalence of RNAP II pausing, they do not reveal whether regulation of RNAP II pausing plays key roles during animal development or homeostasis. Not surprisingly, null and/or strong loss of function mutations in core factors regulating the transition into elongation are often lethal to the cells that carry them. We took

advantage of the *GAL4:UAS-RNAi* system that knocks down, rather than eliminates gene expression to assay the effects of general factors that regulate RNAP II on development and Gro activity.

Depletion of Gro by RNAi in the developing wing predominantly affects vein development, consistent with Gro's role as a key effector of Notch signalling during vein patterning (Hasson and Paroush, 2006). The knockdown of individual factors known to promote pausing led to indiscernible or weak phenotypes in the wing, but concurrent knockdown of these factors with *gro*, greatly enhanced the *gro* vein phenotypes (Figure 5). These results indicate that like Gro, they also promote repression of target genes in response to Notch signalling during vein development. Gro also appears to cooperate with these factors during eye development, where their simultaneous knockdown enhances necrosis and irregular packing of ommatidia (Figure 7). Conversely, knockdown of general pausing factors did not enhance, but rather rescued the modest increase in the eye size observed with the single knockdown of *gro*. This indicates that Gro does not always require these factors to function and there is probably more than one molecular mechanism underlying Gro function in vivo.

The knockdown of *gro* largely rescued the small eye phenotype caused by knockdown of the pause release factor Cdk9, and ectopic expression of Gro leads to a phenotype reminiscent of the Cdk9 knockdown (Figure 7). Although Gro function does not appear to be dependent on pro-pausing factors to regulate eye growth, Gro and Cdk9 have opposing functions in this context. This indicates that Gro acts by more than one mechanism to inhibit Cdk9 activity. One possibility is that Gro blocks the recruitment or function of factors required to activate P-TEFb; for example, the Super Elongation Complex (SEC; reviewed in (Luo et al., 2012). This mechanism may dominate at the subset of Gro targets that do not show high levels of RNAP II pausing (Chambers et al., 2017; Kaul et al., 2014) as P-TEFb is required for transcription at non-paused genes (Dahlberg et al., 2015).

We propose a model in which Gro acts to inhibit transcription through the P-TEFb dependent checkpoint that regulates the entry of the transcription machinery into the elongation phase. Briefly, pioneer factors such as GAF, enable the opening of chromatin at regulatory regions to facilitate the binding of transcription factors that recruit Gro. Recruitment of Gro stabilises the binding of pro-pausing factors at the TSS and/or blocks the recruitment and function of factors that activate P-TEFb. In either case, Gro impedes the transition of the RNAP II complex into the state where it generates full-length transcripts, leading to attenuation of target gene expression. This repression can be rapidly reversed by the arrival of activating

transcription factors that enable P-TEFb to facilitate the progression of the transcription machinery through the elongation checkpoint to generate full length transcripts.

While it has become well established that many genes in developmental and stimulus-responsive pathways exhibit RNAP II pausing, there has been scarce insight into how RNAP II pausing is established in a gene-specific manner. Here, we present evidence that Gro functions through limiting P-TEFb activity in vivo. Thus, the many different transcription factors in these pathways that recruit Gro have the potential to contribute to the establishment and maintenance of RNAP II pausing at their target genes. Future work will reveal if this mechanism is the predominant mode of Gro action and whether it is used by mammalian transcription factors in conjunction with Gro/TLE proteins to regulate gene expression.

MATERIALS AND METHODS

Fly stocks and maintenance

The conditional knockdown of specific genes was achieved using the GAL4/UAS system (Brand and Perrimon, 1993). The following GAL4 and UAS-RNAi and UAS stocks were obtained from Bloomington Drosophila Stock Centre (BDSC); stock numbers indicated in brackets: *nub-GAL4* (86108), *ey-GAL4* (5534), *UAS-Nelf-D*^{RNAi} (38934), *UAS-Nelf-B*^{RNAi} (34847), *UAS-Nelf-A*^{RNAi} (32897), *UAS-Nelf-E*^{RNAi} (32835), *UAS-Bin3*^{RNAi} (50634), and *UAS-GFP*^{RNAi} (9331), *UAS-GFP* (4775), *UAS-Cdk9*^{RNAi} (35323). UAS-RNAi stocks used were also obtained from Vienna Drosophila Resource Center (VDRC; www.vdrc.at; stock numbers indicated in brackets): *UAS-gro*^{RNAi} (6316), *UAS-Trl* ^{RNAi}(41095), *UAS-Larp7* ^{RNAi} (31635). *UAS-gro* (Orian et al., 2007) was a gift from Ze'ev Paroush.

All the flies were raised on standard cornmeal food and maintained either at 25°C or at 18°C in an incubator. Crosses were set up at 25°C on standard cornmeal media with addition of dry yeast.

Selection of the *UAS-RNAi gro* line

We tested three *Drosophila* transgenic RNAi lines against *gro* from the Vienna Drosophila Resource Center (VDRC) (fly stocks 6316 and 6315 from the GD RNAi collection, and 110546 from the KK RNAi collection), and one from the Bloomington Drosophila Stock Center (fly stock 91407). We excluded the fly line from the KK RNAi collection (110546) as it expresses ectopic *tiptop* (Vissers et al., 2016), as well as the one from Bloomington as it gave no phenotype when driven by *nub-GAL4.* The stocks from the VDRC GD RNAi collection (6316 and 6315) gave a very similar phenotype to each other when driven by *nub-GAL4*. However, the 6315 line proved difficult to maintain as an expanded stock so we used 6316 in our experiments, which we refer to as *UAS-gro*^{RNAi}*.* We included *UAS-GFP* in our crosses to maintain a constant number of UAS elements in the single and double knockdowns to ameliorate any effects due changes in the ratio of GAL4 to UAS binding sites across the genome.

Imaging of adult fly wings and eyes

Wings were dissected in ethanol, then rinsed in isopropanol before being mounted in Euparal and imaged using a Zeiss AxioZoom.V16 microscope mounted with an objective Apo Z 0.5x/0.125 FWD 114 MM and Axiocam 506 colour (D) camera.

2-to-3-day old flies were collected and immobilized by freezing at -80ºC. On the day of imaging, flies were thawed, and heads were removed with forceps and mounted on double side tape. Eyes were immediately imaged using a Zeiss AxioZoom.V16 microscope mounted with an objective Apo Z 1.0x/0.25 FWD 60 MM and Axiocam 506 colour (D) camera. Images shown in figures are maximum projections of consecutive optical z-sections. Eye size area was measured using ImageJ software (Schneider et al., 2012).

Analysis of ChIP-seq data

Sequences were aligned to the dm6.26 reference genome using bowtie2 v2.3.3 (Langmead and Salzberg, 2012) with the default parameters, specifying the options "--no-unal" and "-local" to remove unmapped reads and trim adapters. Multimapping reads were then removed using the "view" program of SAMtools v1.7 (Li et al., 2009) with the parameter "-q 20". The program "markdup" was also used to remove PCR duplicates from mapped reads, with the parameter "-r".

For the analysis of Gro ChIP-seq data in BG3, two biological replicates were used to obtain a high confidence set of peaks. Peaks were called using the MACS2 v2.1.1.2 software providing input and immunoprecipitated samples simultaneously with default parameters. Peaks present in both biological replicates (with an FDR less than 10% in at least one of the samples and a p-value less than 0.001 in the other) were selected (Gaspar, 2018).

ChIP-seq peaks for the rest of the factors analysed in this work were retrieved from public databases (see Table 1). Peaks from the data bases that were aligned to the dm5 version of the *Drosophila* genome were converted to dm6.26 using the coordinate converter from Flybase (Larkin et al., 2020). Where ChIP-seq peaks from two replicates were provided, BEDTools intersect was used to obtain a single set of high confidence peaks using a minimum and reciprocal fraction of overlap of 5%.

The Bioconductor library ChIPseeker v1.24.0 (Yu et al., 2015)was used to annotate the peaks by identifying nearby genes, and to visualize their genomic locations. Promoters were defined -250 to 250 bp from the TSSs.The breadth of Gro peaks in BG3 cells was analysed and plotted using the Bioconductor library ChIPpeakAnno v3.22.2 (Zhu et al., 2010). For optimal visualization, the histogram was divided into 500 bins and the two Gro peaks that were greater than 4000 bp were excluded from the histogram.

Open chromatin regions at Gro binding sites were analysed using a DNA-seq dataset in BG3 from modEncode (accession number: 5528), and an ATAC-seq dataset in Kc167 cells

(Porcelli et al., 2019). DNA-seq and ATAC-seq signals at Gro binding sites were plotted in an average plot and a heatmap using computeMatrix and plotHeatmap from deepTools v3.4.3 (Ramírez et al., 2014).The plots were cantered at the centre of Gro peaks and extended 2 Kb upstream and downstream of the reference point.

CentriMo v5.4.1 (Bailey and Machanick, 2012) software from the MEME suite was used to identify sequence motifs that were enriched in 500 bp sequences centred on the Gro binding peak summits in BG3 identified by MACS2. The Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 (Huang et al., 2009a, 2009b) was used for the functional annotation of genes present at Gro binding sites or nearby identified by ChIPseeker.

The chromatin states files for the *Drosophila* genome in Kc167 and BG3 cells files were provided by Sarah Bray (Skalska et al., 2015).To analyse the chromatin states at Gro binding sites in those cell lines, BEDTools intersect was used with a minimum overlap of 50%. Percentage of Gro peaks within each chromatin state (or colour) were then plotted using GraphPad Prim v6.

Co-localization study of ChIP-seq peaks

BEDTools intersect or Intervene (Khan and Mathelier, 2017) were used for the analysis of overlapping binding sites between Gro and the rest of the factors. A minimum and reciprocal fraction of 5% of overlap was required to consider it an overlap. The Venn diagrams shown in the figures were created using the matplotlib-venn package (Hunter, 2007). Upset plots were generated directly from the Intervene module UpSet.

BigWig files were generated for the visualization of the genome tracks of ChIP-seq signal in the Integrate genomics viewer (Thorvaldsdóttir et al., 2013). BigWig files were generated from BAM files using bamCoverage from deepTools v3.4.3 using a bin size of 10, and normalization by counts per million reads mapped (CPM). BigwigCompare was then used to normalize each file against its input using a bin size of 10 and operation subtract as parameters.

Table 1

Deposited data		GSE accession number	Reference
Gro ChIP-seq	Kc167	Array Express E-MTAB-2316	(Kaul et al., 2014)
Gro ChIP-seq	BG3	Array Express E-MTAB-12108	
Gro ChIP-seq	S2R+	Array Express E-MTAB-2316	(Kaul et al., 2014)
GAF ChIP-seq	Kc167	modEncode ID: 3245	(Celniker et al., 2009)
GAF Chip-Chip	BG3	modEncode ID: 2651	(Celniker et al., 2009)
Nelf-E ChIP-seq	Kc167	GSE116889	(Nazer et al., 2018)
Cdk9 ChIP-seq	Kc167	GSE116889	(Nazer et al., 2018)
Pol II isoforms ChIP-seq	Kc167	GSE116889	(Nazer et al., 2018)
DNA-seq	BG3	GSE41351	(Lubelsky et al., 2014)
ATAC-seq	Kc167	GSE122575	(Porcelli et al., 2019)
CycT Chip-Chip	BG3	GSE42399	(Schaaf et al., 2013)

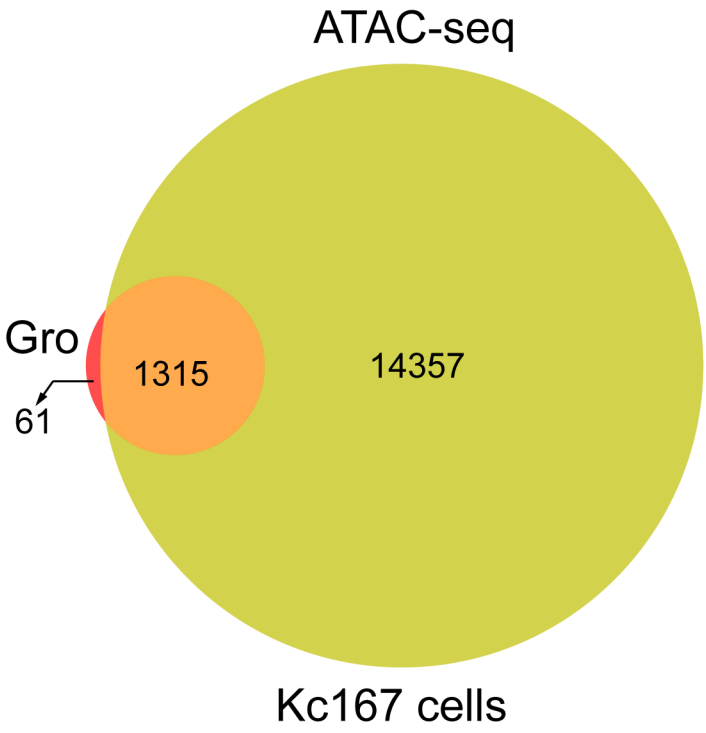
ACKNOWLEDGEMENTS

We would like to thank Aamna Kaul for generating the Gro ChIP-seq data from BG3 cells and Christine Ashton for all the technical support with maintaining fly stocks and ordering. The microscopy was performed with support from the Oxford Brookes University Bioimaging Centre. In addition, many thanks to Alex Buffry, David Ish-Horowicz, Sebastian Kittelmann, Gabrielle Mastrolonardo and Ze'ev Paroush for thoughtful comments on the manuscript and Rob White and Saad Arif for a valuable discussion during MLMQ's PhD thesis defence.

FUNDING

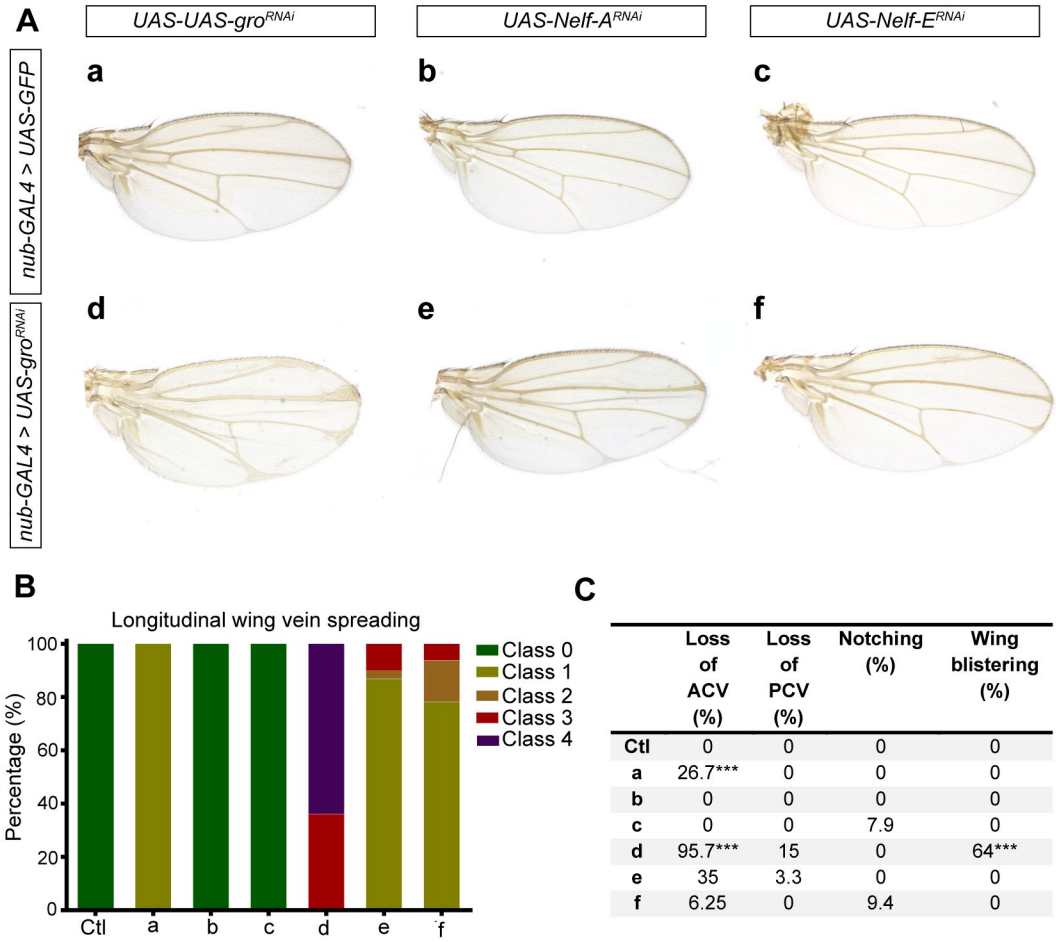
This work was was funded by a Project Grant from the Wellcome Trust (WT090021MA) to BHJ (www.wellcome.ac.uk), an MRC Career Development Award to BHJ (G0800310) and a Nigel Groome Studentship to MLMQ.

Supplementary Figures and Legends



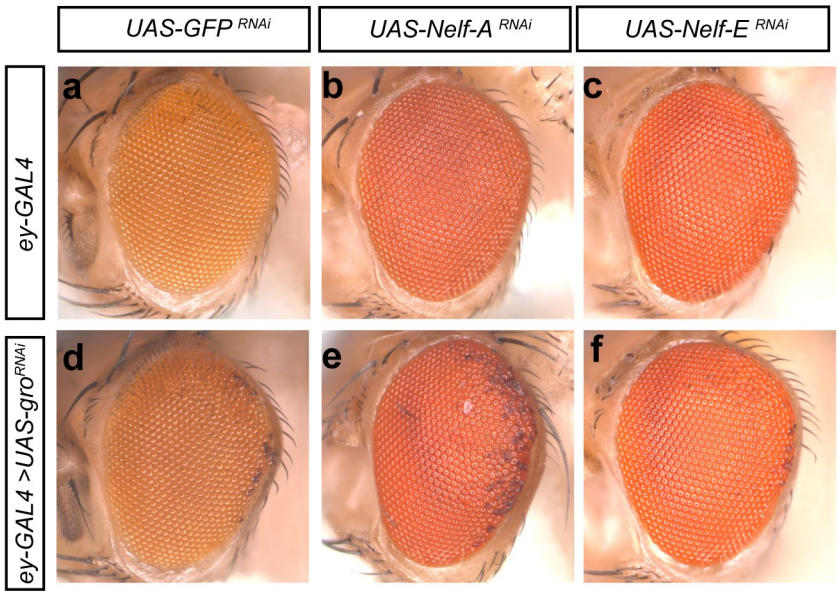
Supp Figure 1: Gro peaks overlap ATAC-seq peaks in Kc167 cells.

Venn diagram showing the overlap between Gro and ATAC-seq peaks in Kc167 cells.



	Loss of ACV (%)	Loss of PCV (%)	Notching (%)	Wing blistering (%)
Ctl	0	0	0	0
a	26.7***	0	0	0
b	0	0	0	0
c	0	0	7.9	0
d	95.7***	15	0	64***
e	35	3.3	0	0
f	6.25	0	9.4	0

Supp Figure 2: RNAi Knockdown *gro*, and *Nelf A/E* in the developing wing.
(A) Representative examples of *Drosophila* adult female wings raised at 25°C expressing the indicated transgenes. (B) Graph showing the frequency of each class (as defined in Figure 4) observed for the phenotypes shown in A. (C) Table showing the quantification of four wing features [total or partial loss of anterior cross vein (ACV), total or partial loss of posterior cross vein (PCV) notching, and wing blistering] for the phenotypes shown in A. Fisher's exact tests were used for pairwise statistical comparisons. For each feature, the single knockdowns were compared to the control, and double knockdowns were compared to the single knockdowns of the genes in the double. ***$p<0.0001$. At least 30 wings of each genotype were scored for quantification in B and C.



Supp Figure 3: RNAi Knockdown *gro*, and *Nelf A/E* in the developing eye.
Representative examples of *Drosophila* adult female wings raised at 25°C expressing the indicated transgenes.

References

Agarwal, M., Kumar, P., and Mathew, S.J. (2015). The Groucho/Transducin-like enhancer of split protein family in animal development. IUBMB Life *67*, 472–481. https://doi.org/10.1002/iub.1395.

Bailey, T.L., and Machanick, P. (2012). Inferring direct DNA binding from ChIP-seq. Nucleic Acids Research *40*, e128–e128. https://doi.org/10.1093/nar/gks433.

Balsalobre, A., and Drouin, J. (2022). Pioneer factors as master regulators of the epigenome and cell fate. Nat Rev Mol Cell Bio *23*, 449–464. https://doi.org/10.1038/s41580-022-00464-z.

Barolo, S., Stone, T., Bang, A.G., and Posakony, J.W. (2002). Default repression and Notch signaling: Hairless acts as an adaptor to recruit the corepressors Groucho and dCtBP to Suppressor of Hairless. Genes & Development *16*, 1964–1976. https://doi.org/10.1101/gad.987402.

Blackledge, N.P., and Klose, R.J. (2021). The molecular principles of gene regulation by Polycomb repressive complexes. Nat Rev Mol Cell Bio *22*, 815–833. https://doi.org/10.1038/s41580-021-00398-y.

Blair, S.S. (2007). Wing Vein Patterning in Drosophila and the Analysis of Intercellular Signaling. Annu Rev Cell Dev Bi *23*, 293–319. https://doi.org/10.1146/annurev.cellbio.23.090506.123606.

Blanch, M., Pineyro, D., and Bernues, J. (2015). New insights for Drosophila GAGA factor in larvae. Royal Society Open Science *2*, 150011–150011. https://doi.org/10.1098/rsos.150011.

Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Dev Camb Engl *118*, 401–415. .

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods *10*, 1213–1218. https://doi.org/10.1038/nmeth.2688.

Celis, J.F. de, and Ruiz-Gómez, M. (1995). groucho and hedgehog regulate engrailed expression in the anterior compartment of the Drosophila wing. Development *121*, 3467–3476. .

Celniker, S.E., Dillon, L.A.L., Gerstein, M.B., Gunsalus, K.C., Henikoff, S., Karpen, G.H., Kellis, M., Lai, E.C., Lieb, J.D., MacAlpine, D.M., et al. (2009). Unlocking the secrets of the genome. Nature *459*, 927–930. https://doi.org/10.1038/459927a.

Chambers, M., Turki-Judeh, W., Kim, M.W., Chen, K., Gallaher, S.D., and Courey, A.J. (2017). Mechanisms of Groucho-mediated repression revealed by genome-wide analysis of Groucho binding and activity. BMC Genomics *18*, 1–17. https://doi.org/10.1186/s12864-017-3589-6.

Chen, G., and Courey, A.J. (2000). Groucho/TLE family proteins and transcriptional repression. Gene *249*, 1–16. .

Cherbas, L., Willingham, A., Zhang, D., Yang, L., Zou, Y., Eads, B.D., Carlson, J.W., Landolin, J.M., Kapranov, P., Dumais, J., et al. (2011). The transcriptional diversity of 25 Drosophila cell lines. Genome Research *21*, 301–314. https://doi.org/10.1101/gr.112961.110.

Choi, C.Y., Kim, Y.H., Kim, Y.-O., Park, S.J., Kim, E.-A., Riemenschneider, W., Gajewski, K., Schulz, R.A., and Kim, Y. (2005). Phosphorylation by the DHIPK2 Protein Kinase Modulates the Corepressor Activity of Groucho. Journal of Biological Chemistry *280*, 21427–21436. https://doi.org/10.1074/jbc.m500496200.

Cockerill, P.N. (2011). Structure and function of active chromatin and DNase I hypersensitive sites. Febs J *278*, 2182–2210. https://doi.org/10.1111/j.1742-4658.2011.08128.x.

Core, L., and Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. Gene Dev *33*, 960–982. https://doi.org/10.1101/gad.325142.119.

Csordás, G., Grawe, F., and Uhlirova, M. (2020). Eater cooperates with Multiplexin to drive the formation of hematopoietic compartments. Elife *9*, e57297. https://doi.org/10.7554/elife.57297.

Dahlberg, O., Shilkova, O., Tang, M., Holmqvist, P.-H., and Mannervik, M. (2015). P-TEFb, the Super Elongation Complex and Mediator Regulate a Subset of Non-paused Genes during Early Drosophila Embryo Development. PLoS Genetics *11*, e1004971-25. https://doi.org/10.1371/journal.pgen.1004971.

Dollinger, R., and Gilmour, D.S. (2021). Regulation of promoter proximal pausing of RNA polymerase II in metazoans. J Mol Biol *433*, 166897. https://doi.org/10.1016/j.jmb.2021.166897.

Filion, G.J., Bemmel, J.G. van, Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., Castro, I.J. de, Kerkhoven, R.M., Bussemaker, H.J., et al. (2010). Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells. CELL *143*, 212–224. https://doi.org/10.1016/j.cell.2010.09.009.

Gaertner, B., and Zeitlinger, J. (2014). RNA polymerase II pausing during development. Development *141*, 1179–1183. https://doi.org/10.1242/dev.088492.

Gaspar, J.M. (2018). Improved peak-calling with MACS2. Biorxiv 496521. https://doi.org/10.1101/496521.

Hasson, P., and Paroush, Z. (2006). Crosstalk between the EGFR and other signalling pathways at the level of the global transcriptional corepressor Groucho/TLE. British Journal of Cancer *94*, 771–775. https://doi.org/10.1038/sj.bjc.6603019.

Hazelett, D.J., Bourouis, M., Walldorf, U., and Treisman, J.E. (1998). decapentaplegic and wingless are regulated by eyes absent and eyegone and interact to direct the pattern of retinal differentiation in the eye disc. Dev Camb Engl *125*, 3741–3751. .

Heigwer, F., Port, F., and Boutros, M. (2018). RNA Interference (RNAi) Screening in Drosophila. Genetics *208*, 853–874. https://doi.org/10.1534/genetics.117.300077.

Himanen, S.V., and Sistonen, L. (2019). New insights into transcriptional reprogramming during cellular stress. J Cell Sci *132*, jcs238402. https://doi.org/10.1242/jcs.238402.

Housden, B.E., Terriente-Felix, A., and Bray, S.J. (2014). Context-Dependent Enhancer Selection Confers Alternate Modes of Notch Regulation on argos. Molecular and Cellular Biology *34*, 664–672. https://doi.org/10.1128/mcb.01045-13.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1–13. https://doi.org/10.1093/nar/gkn923.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44–57. https://doi.org/10.1038/nprot.2008.211.

Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Comput Sci Eng *9*, 90–95. https://doi.org/10.1109/mcse.2007.55.

Jennings, B.H. (2013). Pausing for thought: Disrupting the early transcription elongation checkpoint leads to developmental defects and tumourigenesis. BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology *35*, 553–560. https://doi.org/10.1002/bies.201200179.

Jennings, B.H., and Ish-Horowicz, D. (2008). The Groucho/TLE/Grg family of transcriptional co-repressors. Genome Biol *9*, 205. https://doi.org/10.1186/gb-2008-9-1-205.

Jennings, B., Preiss, A., Delidakis, C., and Bray, S. (1994). The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the Drosophila embryo. Dev Camb Engl *120*, 3537–3548. .

Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nature Reviews Molecular Cell Biology *16*, 167–177. https://doi.org/10.1038/nrm3953.

Kaul, A., Schuster, E., and Jennings, B.H. (2014). The Groucho Co-repressor Is Primarily Recruited to Local Target Sites in Active Chromatin to Attenuate Transcription. Plos Genet *10*, e1004595. https://doi.org/10.1371/journal.pgen.1004595.

Khan, A., and Mathelier, A. (2017). Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. Bmc Bioinformatics *18*, 287. https://doi.org/10.1186/s12859-017-1708-7.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480–485. https://doi.org/10.1038/nature09725.

Krejci, A., Bernard, F., Housden, B.E., Collins, S., and Bray, S.J. (2009). Direct Response to Notch Activation: Signaling Crosstalk and Incoherent Logic. Science Signaling *2*, ra1–ra1. https://doi.org/10.1126/scisignal.2000140.

Lagha, M., Bothma, J.P., Esposito, E., Ng, S., Stefanik, L., Tsui, C., Johnston, J., Chen, K., Gilmour, D.S., Zeitlinger, J., et al. (2013). Paused Pol II Coordinates Tissue Morphogenesis in the Drosophila Embryo. CELL *153*, 976–987. https://doi.org/10.1016/j.cell.2013.04.045.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

Larkin, A., Marygold, S.J., Antonazzo, G., Attrill, H., Santos, G.D., Garapati, P.V., Goodman, J.L., Gramates, L.S., Millburn, G., Strelets, V.B., et al. (2020). FlyBase: updates to the Drosophila melanogaster knowledge base. Nucleic Acids Res https://doi.org/10.1093/nar/gkaa1026.

Lee, C., Li, X., Hechmer, A., Eisen, M., Biggin, M.D., Venters, B.J., Jiang, C., Li, J., Pugh, B.F., and Gilmour, D.S. (2008). NELF and GAGA Factor Are Linked to Promoter-Proximal Pausing at Many Genes in Drosophila. Molecular and Cellular Biology *28*, 3290–3300. https://doi.org/10.1128/mcb.02224-07.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, 1000 Genome Project Data Processing (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Li, J., Liu, Y., Rhee, H.S., Ghosh, S.K.B., Bai, L., Pugh, B.F., and Gilmour, D.S. (2013). Kinetic Competition between Elongation Rate and Binding of NELF Controls Promoter-Proximal Pausing. Molecular Cell *50*, 711–722. https://doi.org/10.1016/j.molcel.2013.05.016.

Luo, Z., Lin, C., and Shilatifard, A. (2012). The super elongation complex (SEC) family in transcriptional control. Nature Publishing Group *13*, 543–547. https://doi.org/10.1038/nrm3417.

Nagel, A.C., Krejci, A., Tenin, G., Bravo-Patiño, A., Bray, S., Maier, D., and Preiss, A. (2005). Hairless-mediated repression of notch target genes requires the combined activity of Groucho and CtBP corepressors. Molecular and Cellular Biology *25*, 10433–10441. https://doi.org/10.1128/mcb.25.23.10433-10441.2005.

Nazer, E., Dale, R.K., Palmer, C., and Lei, E.P. (2018). Argonaute2 attenuates active transcription by limiting RNA Polymerase II elongation in Drosophila melanogaster. Nature Publishing Group *8*, 1–10. https://doi.org/10.1038/s41598-018-34115-1.

O'Brien, T., and Lis, J.T. (1991). RNA polymerase II pauses at the 5' end of the transcriptionally induced Drosophila hsp70 gene. Mol Cell Biol *11*, 5285–5290. https://doi.org/10.1128/mcb.11.10.5285.

Orian, A., Delrow, J.J., Nieves, A.E.R., Abed, M., Metzger, D., Paroush, Z., Eisenman, R.N., and Parkhurst, S.M. (2007). A Myc–Groucho complex integrates EGF and Notch signaling to regulate neural development. Proc National Acad Sci *104*, 15771–15776. https://doi.org/10.1073/pnas.0707418104.

Paroush, Z., Finley, R.L., Kidd, T., Wainwright, S.M., Ingham, P.W., Brent, R., and Ish-Horowicz, D. (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell *79*, 805–815. https://doi.org/10.1016/0092-8674(94)90070-1.

Peterlin, B.M., and Price, D.H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. Mol Cell *23*, 297–305. https://doi.org/10.1016/j.molcel.2006.06.014.

Peterlin, B.M., Brogie, J.E., and Price, D.H. (2012). 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. Wiley Interdiscip Rev Rna *3*, 92–103. https://doi.org/10.1002/wrna.106.

Porcelli, D., Fischer, B., Russell, S., and White, R. (2019). Chromatin accessibility plays a key role in selective targeting of Hox proteins. Genome Biology *20*, 1–19. https://doi.org/10.1186/s13059-019-1721-4.

Quiring, R., Walldorf, U., Kloter, U., and Gehring, W.J. (1994). Homology of the eyeless Gene of Drosophila to the Small eye Gene in Mice and Aniridia in Humans. Science *265*, 785–789. https://doi.org/10.1126/science.7914031.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res *42*, W187–W191. https://doi.org/10.1093/nar/gku365.

Schaaf, C.A., Kwak, H., Koenig, A., Misulovin, Z., Gohara, D.W., Watson, A., Zhou, Y., Lis, J.T., and Dorsett, D. (2013). Genome-Wide Control of RNA Polymerase II Activity by Cohesin. PLoS Genetics *9*, e1003382-13. https://doi.org/10.1371/journal.pgen.1003382.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

Skalska, L., Stojnic, R., Li, J., Fischer, B., Cerda-Moya, G., Sakai, H., Tajbakhsh, S., Russell, S., Adryan, B., and Bray, S.J. (2015). Chromatin signatures at Notch-regulated enhancers reveal large-scale changes in H3K56ac upon activation. Embo J *34*, 1889–1904. https://doi.org/10.15252/embj.201489923.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178–192. https://doi.org/10.1093/bib/bbs017.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to Repress RNA Polymerase II Elongation. Cell *97*, 41–51. https://doi.org/10.1016/s0092-8674(00)80713-8.

Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382–2383. https://doi.org/10.1093/bioinformatics/btv145.

Zhu, L.J., Gazin, C., Lawson, N.D., Pagès, H., Lin, S.M., Lapointe, D.S., and Green, M.R. (2010). ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. Bmc Bioinformatics *11*, 237. https://doi.org/10.1186/1471-2105-11-237.