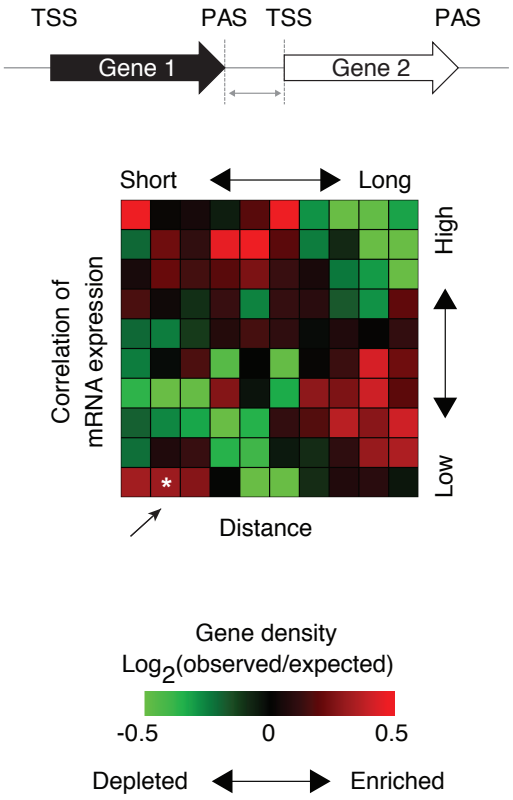
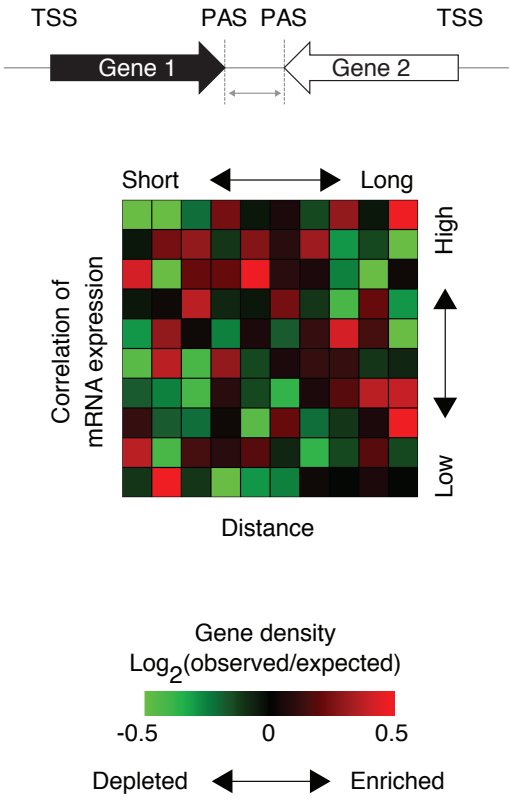
Supplementary Material to:

Telmo Henriques, Zhe Ji, Sue Mei Tan-Wong, Alexandre M. Carmo, Bin Tian, Nicholas J Proudfoot and Alexandra Moreira. Transcription termination between *polo* and *snap*, two closely spaced tandem genes of *D. melanogaster*. Transcription 2012; 3(4); http://dx.doi.org/10.4161/trans.21967; http://www.landesbioscience.com/journals/transcription/article/21967/

A



B

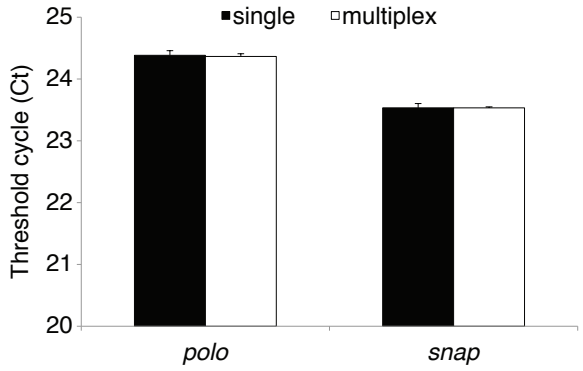


C

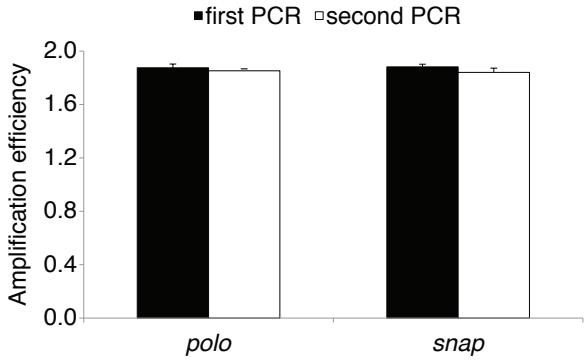
Tissue	*polo* mRNA signal	*snap* mRNA signal
Brain	0 +/- 0	2057 +/- 81
Head	1 +/- 0	954 +/- 61
Eye	0 +/- 0	648 +/- 14
Thoracico abdominal ganglion	3 +/- 1	2365 +/- 82
Salivary gland	5 +/- 1	91 +/- 6
Crop	7 +/- 2	206 +/- 15
Midgut	11 +/- 1	79 +/- 1
Tubule	17 +/- 14	192 +/- 6
Hindgut	5 +/- 2	152 +/- 5
Heart	3 +/- 1	131 +/- 9
Fat body	16 +/- 6	85 +/- 10
Ovary	1260 +/- 15	36 +/- 0
Testis	643 +/- 27	19 +/- 1
Male accessory glands	1 +/- 0	167 +/- 13
Virgin spermatheca	22 +/- 0	196 +/- 20
Mated spermatheca	14 +/- 4	171 +/- 16
Adult carcass	22 +/- 13	138 +/- 5
Larval CNS	746 +/- 21	644 +/- 71
Larval Salivary gland	18 +/- 3	105 +/- 7
Larval midgut	34 +/- 5	80 +/- 6
Larval tubule	16 +/- 1	222 +/- 9
Larval hindgut	55 +/- 4	120 +/- 7
Larval fat body	147 +/- 62	150 +/- 16
Larval trachea	192 +/- 12	73 +/- 15
Larval carcass	121 +/- 14	92 +/- 4
S2 cells (growing)	771 +/- 21	121 +/- 2
Whole fly	613 +/- 8	191 +/- 13

data from http://www.flyatlas.org/

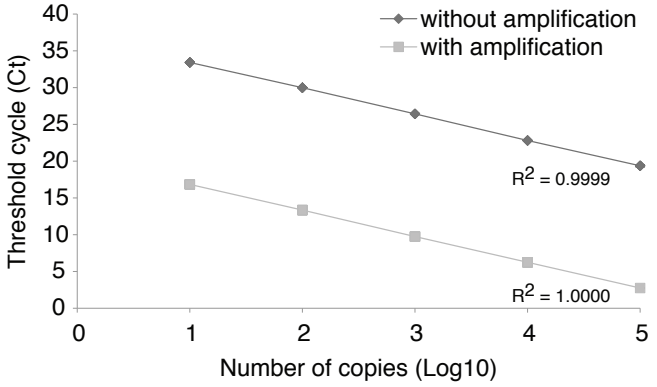
A


single
multiplex
Threshold cycle (Ct)
25
24
23
22
21
20
polo
snap

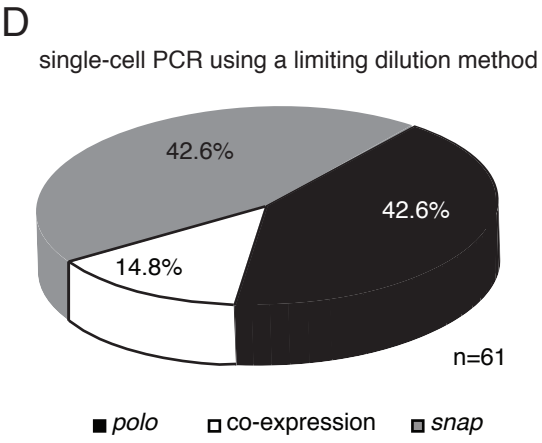
B

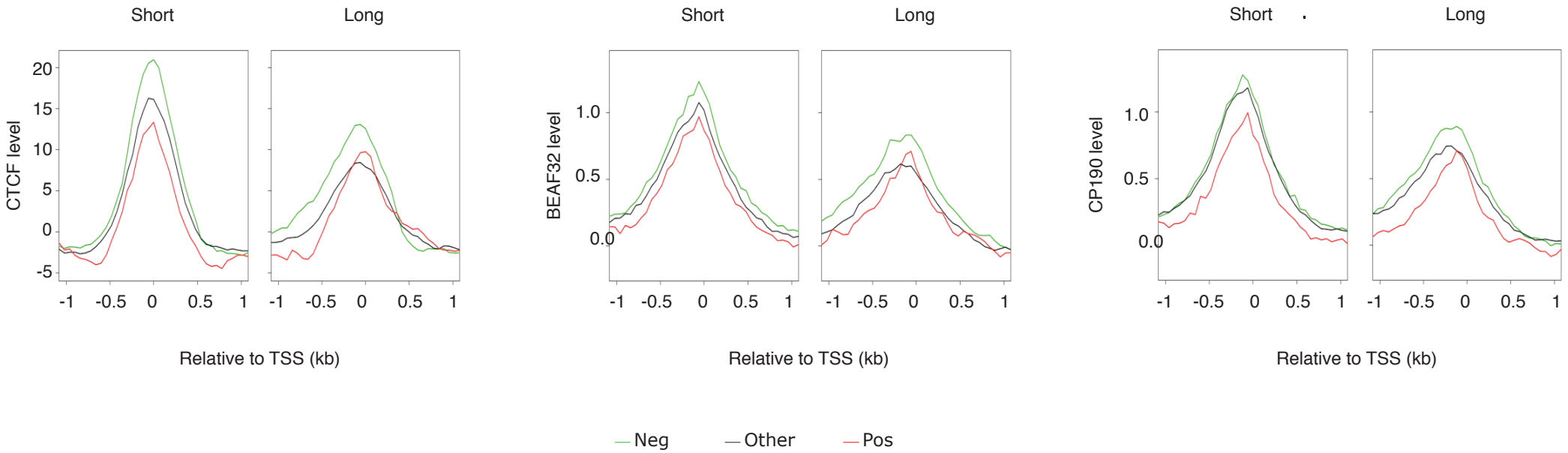

first PCR
second PCR
Amplification efficiency
2.0
1.6
1.2
0.8
0.4
0.0
polo
snap

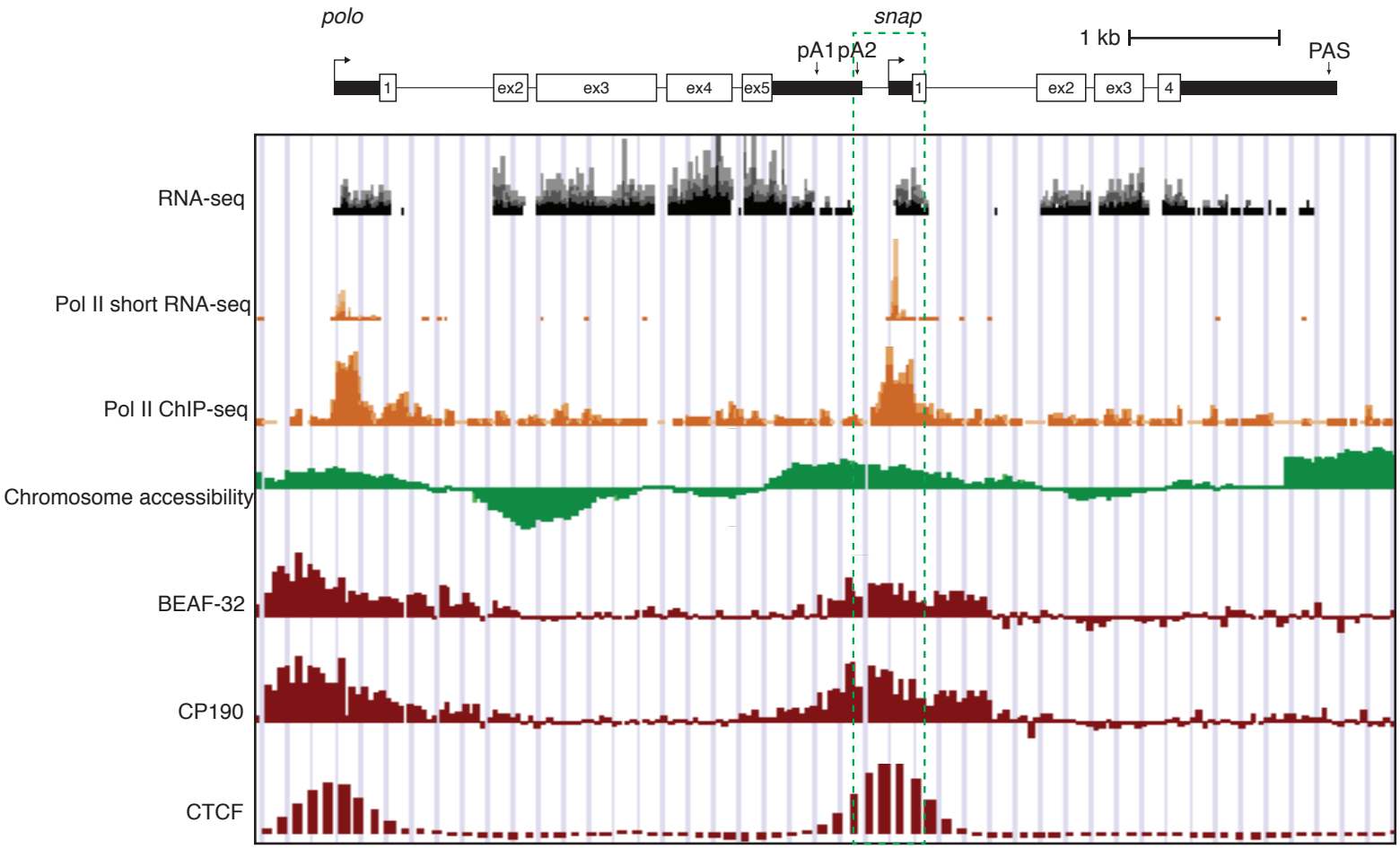
C


without amplification
with amplification
Threshold cycle (Ct)
40
35
30
25
20
15
10
5
0
R2 = 0.9999
R2 = 1.0000
0
1
2
3
4
5
Number of copies (Log10)

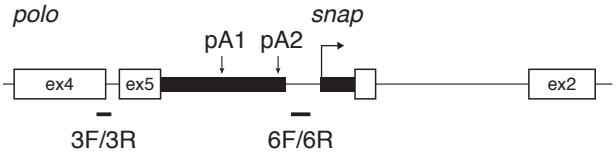
D

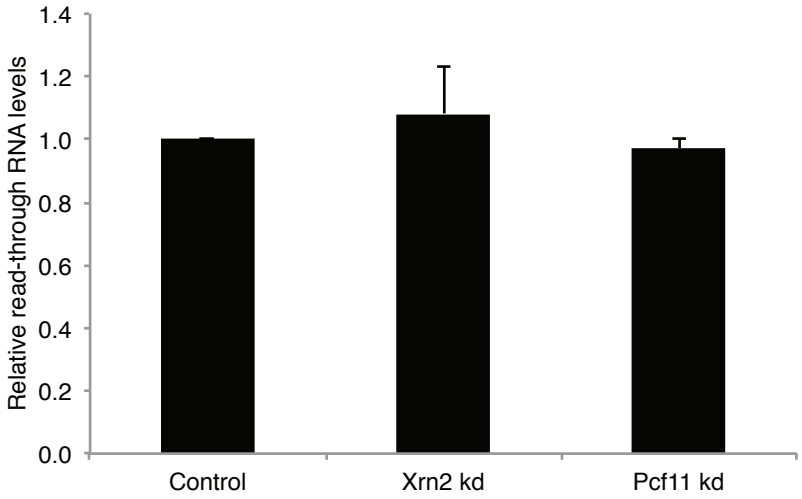

single-cell PCR using a limiting dilution method
42.6%
42.6%
14.8%
n=61
polo
co-expression
snap


Short
Long
CTCF level
20
15
10
5
0
-5
-1
-0.5
0
0.5
1
Relative to TSS (kb)
Short
Long
BEAF32 level
1.0
0.5
0.0
-1
-0.5
0
0.5
1
Relative to TSS (kb)
Short
Long
CP190 level
1.0
0.5
0.0
-1
-0.5
0
0.5
1
Relative to TSS (kb)
Neg
Other
Pos


polo
snap
1 kb
pA1
pA2
PAS
1
ex2
ex3
ex4
ex5
1
ex2
ex3
4
RNA-seq
Pol II short RNA-seq
Pol II ChIP-seq
Chromosome accessibility
BEAF-32
CP190
CTCF

A


polo
snap
pA1
pA2
ex4
ex5
ex2
3F/3R
6F/6R


Relative read-through RNA levels
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Control
Xrn2 kd
Pcf11 kd

B


pA1
pA2
ex4
ex5
3'UTR
AAAAAAA
polo
oligodT
PAS
ex2
ex3
3'UTR
AAAAAAA
snap
oligodT

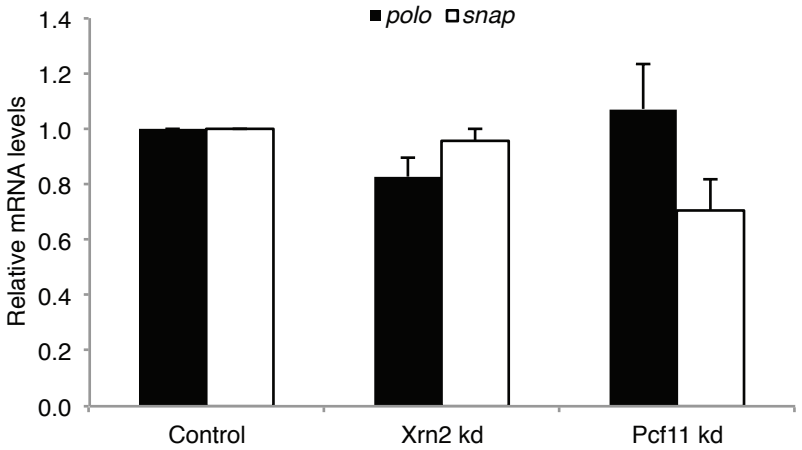

polo
snap
Relative mRNA levels
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Control
Xrn2 kd
Pcf11 kd

Figure S1. *polo-snap* and similar gene pairs show negative correlation of expression. Related to Fig. 1. (A) Schematic representation of the analysed tandem genes across *Drosophila* genome (dm3); heat map representation of correlation between gene distance (tandem only) and correlation of gene expression. All the tandem genes were distributed in a 10x10 table, based on gene distance (x-axis) and correlation of expression (y-axis). Number of genes in each cell of the table was normalized to an expected number using randomized data. Relative gene density, calculated by log2 (observed number/expected number), is presented in a heat map according to the colour scale shown in the figure. The *r* and *P* shown in the figure are based on Spearman correlation. The gene expression correlation between *polo* and *snap* is -0.58. The cell containing *polo-snap* is marked by a white dot. (B) Schematic representation of the analysed convergent gene pairs across the *Drosophila* genome (dm3); heat map representation of correlation between gene distance (convergent only) and correlation of gene expression. (C) Correlation of gene expression between *polo* and *snap*. The data is derived from the FlyAtlas database where gene expression was analysed across 26 fly tissues.

Figure S2. Competition, amplification efficiency and linearity reactions for single-cell PCR. Related to Fig. 1. (A) Competition of PCR amplifications**.** Aliquots of cDNA from S2 cells were amplified separately for each gene, or in multiplex in the first PCR round. Triplicates of these reactions were further amplified in a second real-time PCR. Comparison of threshold cycle mean values (Ct) obtained for each different gene amplified separately (black bars) or in multiplex (white bars). No significant differences were observed between the two amplification conditions for each different gene (t-test, *p*-value >0.1). (B) Aliquots of cDNA from *Drosophila* S2 cells were amplified separately for each gene and each type of PCR reaction to determine PCR efficiency. Slope values from the triplicates were assessed on the exponential phase of the real-time amplification reaction, and PCR efficiencies were determined using LinRegPCR 11.4 software. Mean+/-SEM of PCR efficiencies is shown for the first (black bars) and second (white bars) PCRs. The significance of these differences was evaluated by ANOVA. Within each PCR, all primer combinations had the same efficiency. We also found no significant difference in the variance between the first and second PCR amplifications (ANOVA, *p*-value >0.1). (C) Linearity of double-round PCR amplification. A double-stranded synthesized DNA sequence from rp49 was amplified by a single PCR (◆) or by a two-step PCR of 15 pre-amplification cycles (■). Mean±SEM of triplicates is included,

although SEMs frequently overlap with the symbols as we found little or no variation. Linear regression curves of these amplified standards are indicated and have similar high correlation coefficients (r^2≥0.9999). (D) S2 cells were diluted in Schneider's Insect Medium (Sigma) supplemented with 10% FBS at a concentration of 1cell/ 5ml and 5ml of this suspension was added to each well of a 96-well plate and immediately frozen in liquid nitrogen. Single-cell PCR was performed as in **Fig. 1E-F**.

Figure S3. CTCF, BEAF32 and CP190 levels around the promoter region for different gene sets. Related to Fig. 2. Insulator levels (left, CTCF; middle, BEAF32; right, CP190). Only tandem gene pairs were used, and the Short, Long, Neg, Pos and Other gene pairs are the same as described with Fig. 2. The x-axes of all plots are distance relative to TSS with 0 being TSS.

Figure S4. Molecular features at the *polo* and *snap* transcriptional units. Related to Fig. 2. The several datasets used in this study (see Materials and Methods) were uploaded into the UCSC genome browser and a snapshot around *polo-snap* was taken.

Figure S5. Transcription termination occurs efficiently in the short intergenic region between *polo* and *snap* and is independent of Xrn2 and Pcf11. Related to Fig. 3. (A) RT-qPCR of read-through RNA. The diagram shows the primer positions for reverse transcription and qPCR analysis. The graph shows RT-PCR quantitation of endogenous read-through (ratio between 6F/6R and 3F/3R probes) RNA in S2 cells depleted for Xrn2 and Pcf11 compared to DsRed kd (Control). Primers used for reverse transcription (3R and 6R) and qPCR (3F/3R and 6F/6R) are shown in the diagram. (B) *polo* and *snap* messenger RNA levels upon Xrn2 and Pcf11 depletion. The diagram shows the primer positions for qPCR analysis. The amount of *polo* and *snap* were measured by qPCR. The graph shows the endogenous mRNA levels for *polo* and *snap* in S2 cells where DsRed (Control), Xrn2 and Pcf11 were depleted. OligodT was used for reverse transcription, as depicted in the diagram. For all the panels the value for Control was set at 1 and the error bars show s.e.m.

Table S1. Datasets used in this study

Datasets	GEO accession #	Sample source	References
Chromosome accessibility	GSM441282	S2 cells	27
RNA-seq	GSM480160	S2 cells	63
Insulator-CTCF200	GSM499649	S2 cells	28

Insulator-BEAF32	GSM409067	E0-12h embryos	28
Insulator-CP190	GSM409068	E0-12h embryos	28
Pol II ChIP-seq	GSM463297	S2 cells	28
Short RNA-seq	GSM463298	S2 cells	28
Gene expression across tissues	GSE7763	26 fly tissues*	26
TFIIB and RNAP II clusters	NA	Stage 4-5 embryos	43

*: 26 fly tissues include Brain, Head, Eye, Thoracicoabdominal ganglion, Salivary gland, Crop, Midgut, Tubule, Hindgut, Heart, Fat body, Ovary, Testis, Male accessory glands, Virgin spermatheca, Mated spermatheca, Adult carcass, Larval CNS, Larval Salivary gland, Larval midgut, Larval tubule, Larval hindgut, Larval fat body, Larval trachea, Larval carcass, and S2 cells (growing).

Table S2. Oligonucleotides

single cell RT-PCR

polo ex4 F	CGCGTATGCCCCATTTACAC
polo exon4 F	CCGTACAACATGTGCCGTAG
polo exon5 R	CTTTAGACACGCCGTTCTCC
snap ex2 F	CAAGAAGGTGGACGTGGAGA
snap exon2 F	CAACTGCCTGATGAAGTCCA
snap exon3 R	GACCGACTCTTCACCCTTGA
rp49 ex2 F	CCAGCATACAGGCCCAAGAT
rp49 exon2 F	AAGAAGCGCACCAAGCACT
rp49 exon3 R	GTTCGATCCGTAACCGATGT
rp49 ex3 T7R	taatacgactcactatagggagaGTTCGATCCGTAACCGATGT

BrUTP-NRO/ ChIP

polo and *snap*

1F	GCTTTGTGCTTGGTTTTCGT
1R	TTTACTACGGACTGCCCCTTT
2F	GTTCTTGCCCAGCTCTTGTC
2R	AGATTGGCCTTGAGGAAGGT
3F	CCGTACAACATGTGCCGTAG
3R	CCAGATGTACATGATGCCGA
4F	AAGGCCGAATGTTAGTTTAACG
4R	TTTCGATATGAAGGGGAAGG
5F	ACGTGTTTCGAAATGCCTAT
5R	ACACTTAAACACTTTGCAGCAG
6F	CTGAAAAAAGATTGTTTGCAACG
6R	GGGCTGTAACGTTTGTTCGT
7F	ACCACTTCTGCGGTCACAC

7R	GTCACCCATGTCGAGGATTT
8F	GGACGCCATCGAGTGCTA
8R	TGGACTTCATCAGGCAGTTG
9F	GCAGTACCCAGCATTCCA
9R	CCCTCGATGTTCTGTTCCTCTA

CG30046

1F	GGAATTTTCGCTCAGTTTGC
1R	CAGGGACCCAAGCAGAAGTA
2F	CCCACAGTTCGGATTATCGT
2R	TCGTTTCAAGCATGGAGTCA
3F	TATCCGGAACTCCCATGAAA
3R	AAGGTCTTCCCCACTAAAGTCC
4F	ACACCTGGAGGAGATGTGGA
4R	ACTCATCCGCACAGAAGTTG
5F	GCAACTATGGCGATTCCAG
5R	GGAACACTTCTCGCATACCC
6F	TGCAGACATGGAATTTGCAC
6R	TCATTCGAGCTACGCAGAGA
7F	GTTAAAGCGCCTCGACTTTG
7R	GAATGATGCAGGCTGACTGA
8F	TTTTGGTATGCATCTCCCATT
8R	TTGTTGGTGTGCTGTTTGGT
9F	GCGAGCAATCGGTTTTCTAT
9R	CAGGATCAGCAGGCTCATAA

qPCR

7SL RNA F	TTGGCTAAGGAGGGATGAAC
7SL RNA R	CTACTGCCTACCACGGGAAC
dPcf11 F	GCATCAGCTAAAGCACCAAA
dPcf11 R	TATAATCGGCGGACTTGGAG
dXrn2 F	GTTCAAGGCCAGGGACAAG
dXrn2 R	AAGGATCCGTGGTCCTGATT
TFIIB F	CACAAGCATCCGAGCATAAG
TFIIB R	GTAGGACTGTCTGATGGTGACG
eGFP F	TATATCATGGCCGACAAGCA
eGFP R	GTTGTGGCGGATCTTGAAGT
polo exon3 F	GCGATATCGAGAGCCTGTACC
polo exon4 F	CCGTACAACATGTGCCGTAG
polo exon5 R	CTTTAGACACGCCGTTCTCC
rp49 exon2 F	AAGAAGCGCACCAAGCACT
rp49 exon3 R	GTTCGATCCGTAACCGATGT
snap exon2 F	CAACTGCCTGATGAAGTCCA
snap exon3 R	GACCGACTCTTCACCCTTGA
snap exon4 R	CCCTCGATGTTCTGTTCCTCTA

RNAi

DsRed T7F	taatacgactcactatagggagaCTTCAAGGTGCGCATGGAG
DsRed T7R	taatacgactcactatagggagaGGACTTGAACTCCACCAGGTAGTG
dPcf11 T7F	taatacgactcactatagggagaGCGAAGTGGCTTTCCTAGTG
dPcf11 T7R	taatacgactcactatagggagaTCTCCCAAAAGGAATGATGC
dXrn2 T7F	taatacgactcactatagggagaATCCATCCGTGCACGCATC
dXrn2 T7R	taatacgactcactatagggagaGTTGCAGATGTCGCAGGGA
TFIIB T7F	taatacgactcactatagggagaCTCCGCTCATCGAGGACTAC
TFIIB T7R	taatacgactcactatagggagaACGTCCCGGTACAATATCCA

***polo* and *snap* cloning**

intergenic polo HindIII F	TAGTAGaagcttGGTAACGGCCATAGAATTTG
snap intergenic SpeI R	TAGTAGactagtCCAAGTGACAGGTGCAAA
snap intron1 SpeI R	TAGTAGactagtAGTCGCACTTTTGCGCTATT
polo ex3R	CTTTCCGTTGATAAGATTAGTGAGC
polo int4F	GTTTGTATAATTGTACAGCATCGGC
snap ex2R	CAGATTATTGGGATCACTTTCGTA
snap int3F	GTGGGTATCAGTTTGCTTAAGTTTT
pnoMT F	CAAGTGAATCATCTCAGTGCAACT
pnoMT R	GGCCTCGTGATACGCCTATT

3C analysis

1F	GCGGTTAAACACAGTGTTGCGCA
1R	GCGTCAGCTGAATTGGTGTGCC
2F	GCTCAGAACGAGGAGTATCGCCG
2R	GCCGAAGAACCGCATGCG
ApoI 2R	GGAGCCCCTTCCAAGAAAAT
3F	CTCCTTCGAGGTGGACATCTGGTC
3R	CGCCGCCGGTTTCCTTAAGTAG
4F	GCCCAGAAAGTATCCTCCTCTCCCATC
4R	AGGCATTTCGAAACACGTTTGCTTAG
anchor F	TGTAATATTGTGCGTTCGTAGTGCGCT
anchor R	AGCTGCCCTAAACGTTGCAAACA
polo exon4 F	CCGTACAACATGTGCCGTAG
polo exon5 R	CTTTAGACACGCCGTTCTCC
rp49 1F	CGACGTATCGATGTTTTTTCACACCA
rp49 1R	GACTAACGCAGTTCAACTCGAAACCG
rp49 2F	CAGTGGGTCAGTGCACTAATGGCTACT
rp49 2R	CACAATCCTCGTTGGCACTCACC

3’ RACE analysis

3’ RACE adapter primer	GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTTVN
anchor primer	GACCACGCGTATCGATGTCGAC
polo 3F	CCGTACAACATGTGCCGTAG