

SUPPLEMENTARY FILES FOR:

Life stage specific poly(A) site selection regulated by *Trypanosoma brucei* DRBD18

Jonathan E. Bard, Brianna L. Tylec, Ashutosh P. Dubey, Natalie A. Lamb,
Donald A. Yergeau, and Laurie K. Read

Contents:

Fig. S1. DRBD18 RNAi leads to increased abundance of many transcripts encoding kinases, phosphatases, and RNA binding proteins (RBP).

Fig. S2. Reproducibility of single cell RNAseq replicates.

Fig. S3. Volcano plot depicting differences in cluster markers between Clusters 0 and 7.

Fig. S4. Reads covering two 5.8S rRNA intergenic regions increased upon DRBD18 depletion.

Fig. S5. Clusters 0 and 7 are enriched for transcripts with proximal poly(A) sites.

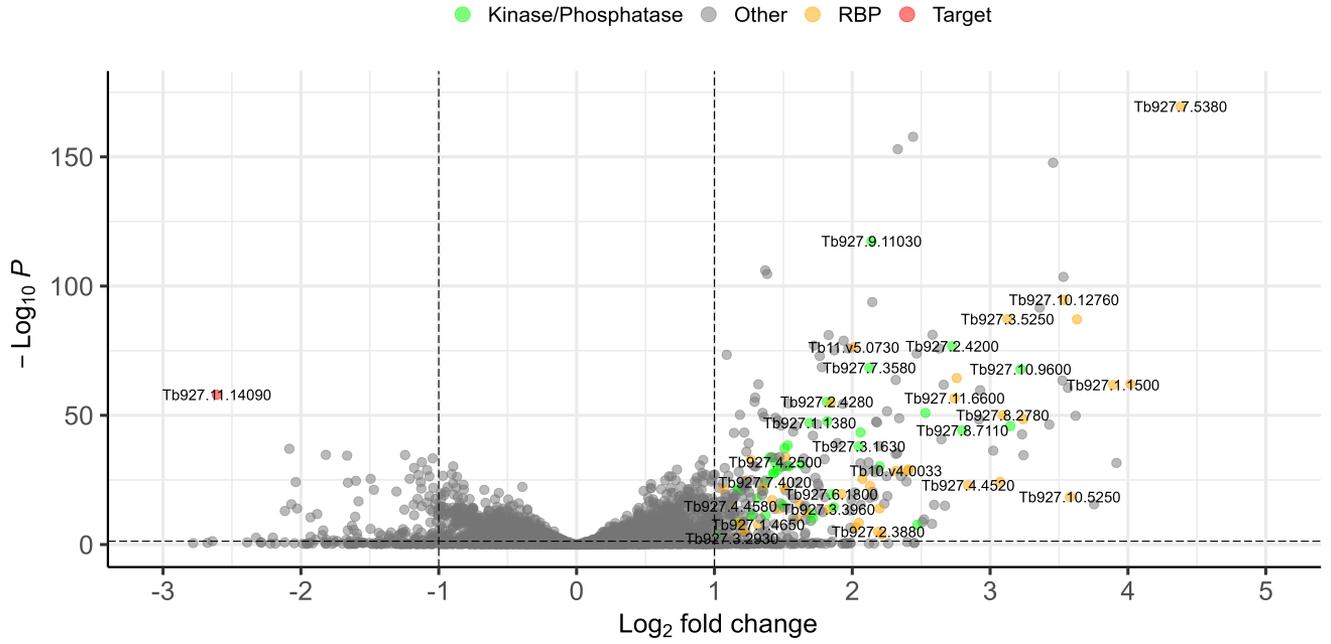
Fig. S6. 3'UTR motifs would be lost upon proximal polyadenylation site generation in DRBD18 depleted cells.

Fig. S7. Upregulation of stumpy BF and MF transcripts with altered polyadenylation sites is reflected at the protein level.

Fig. S8. DRBD18 bound transcripts significantly overlap with numerous gene regulatory parameters.

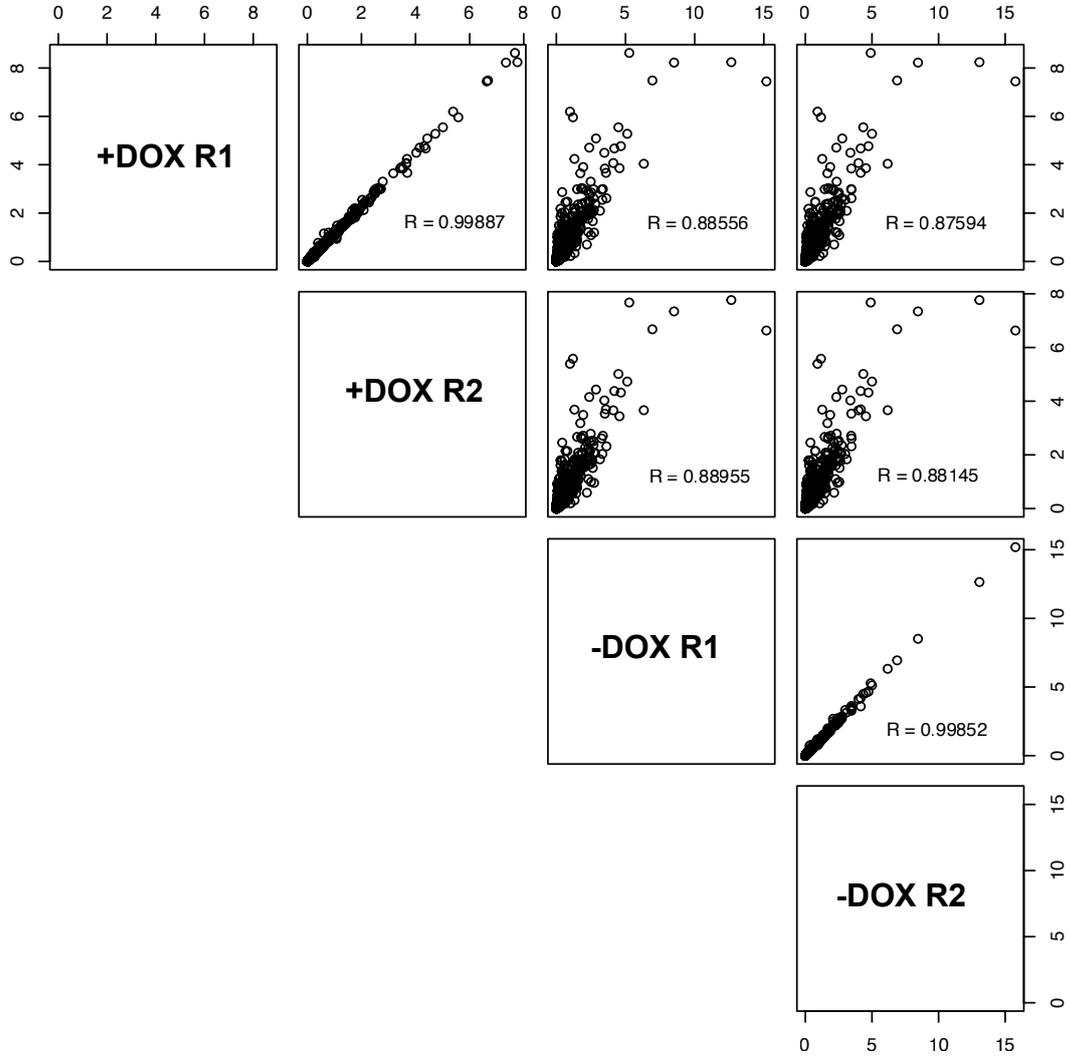
Supplemental Figure 1

DRBD18 RNAi RNA-seq



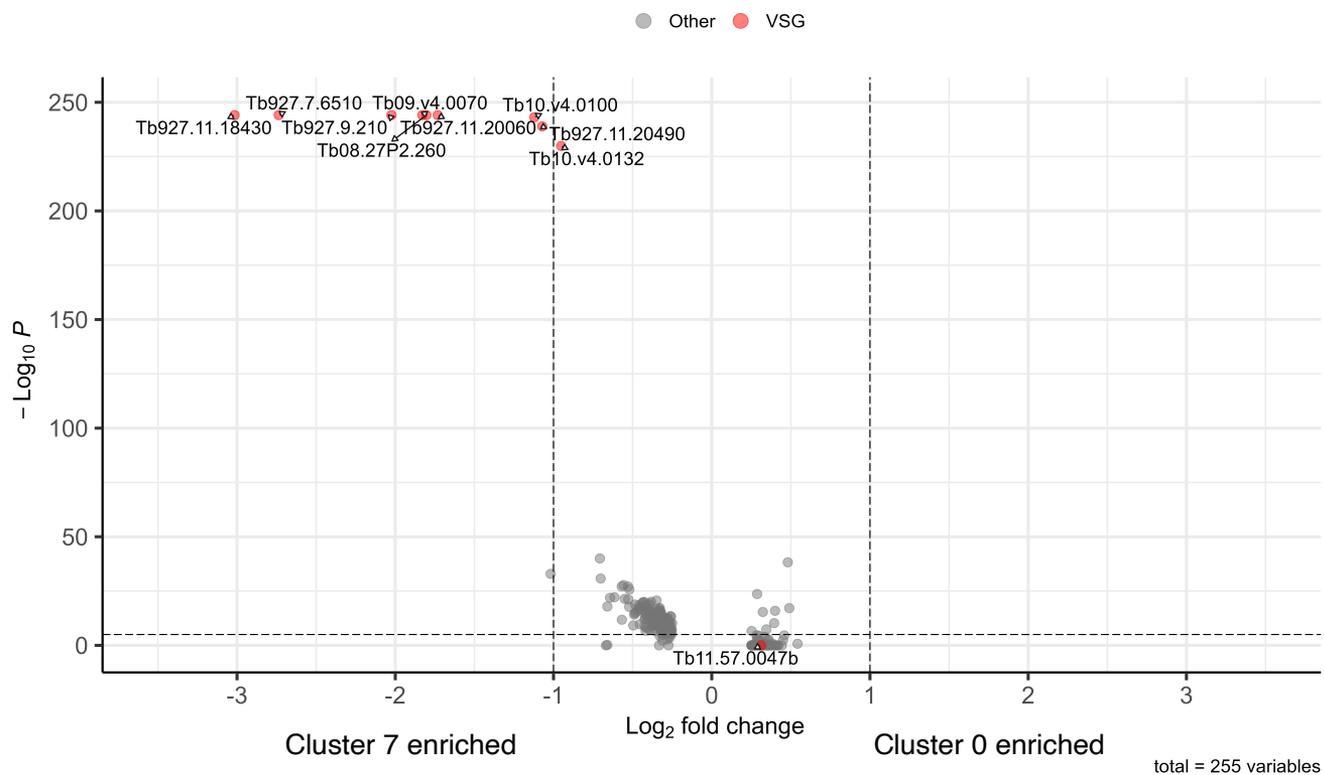
total = 11703 variables

Supplemental Figure 2

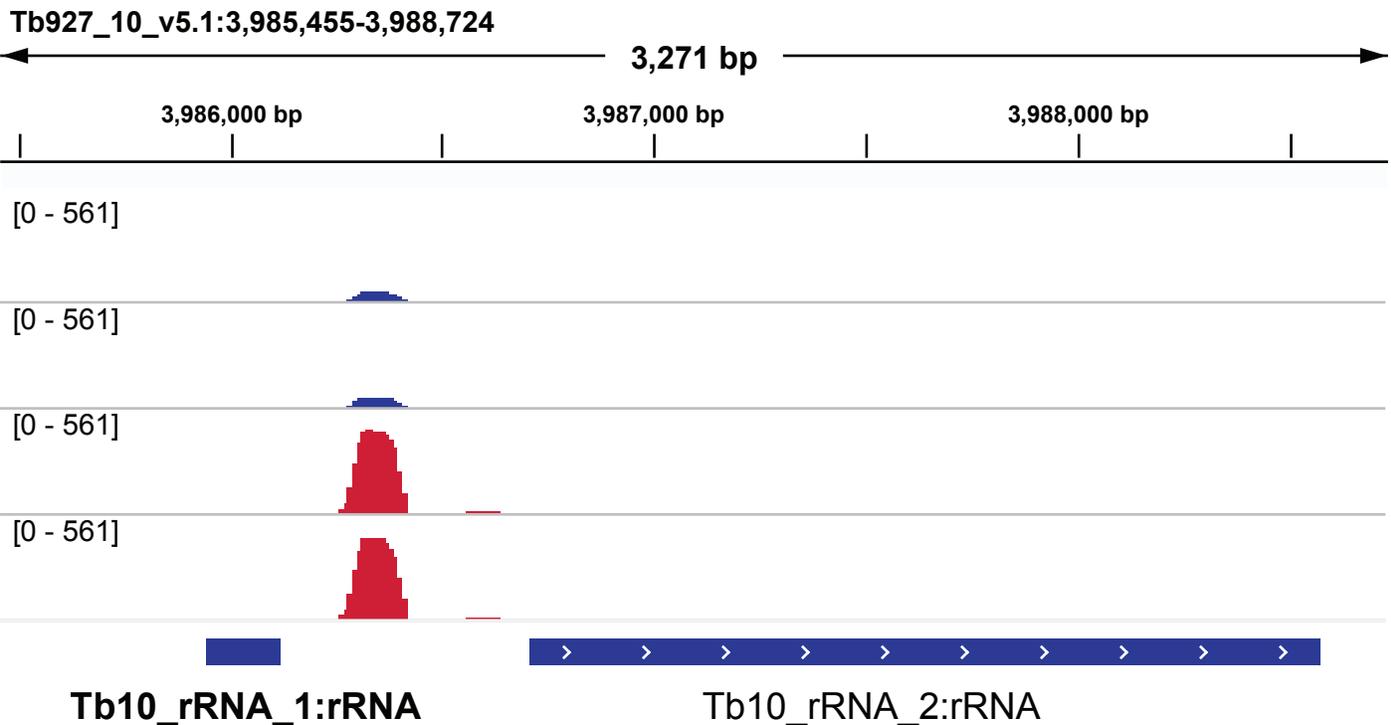
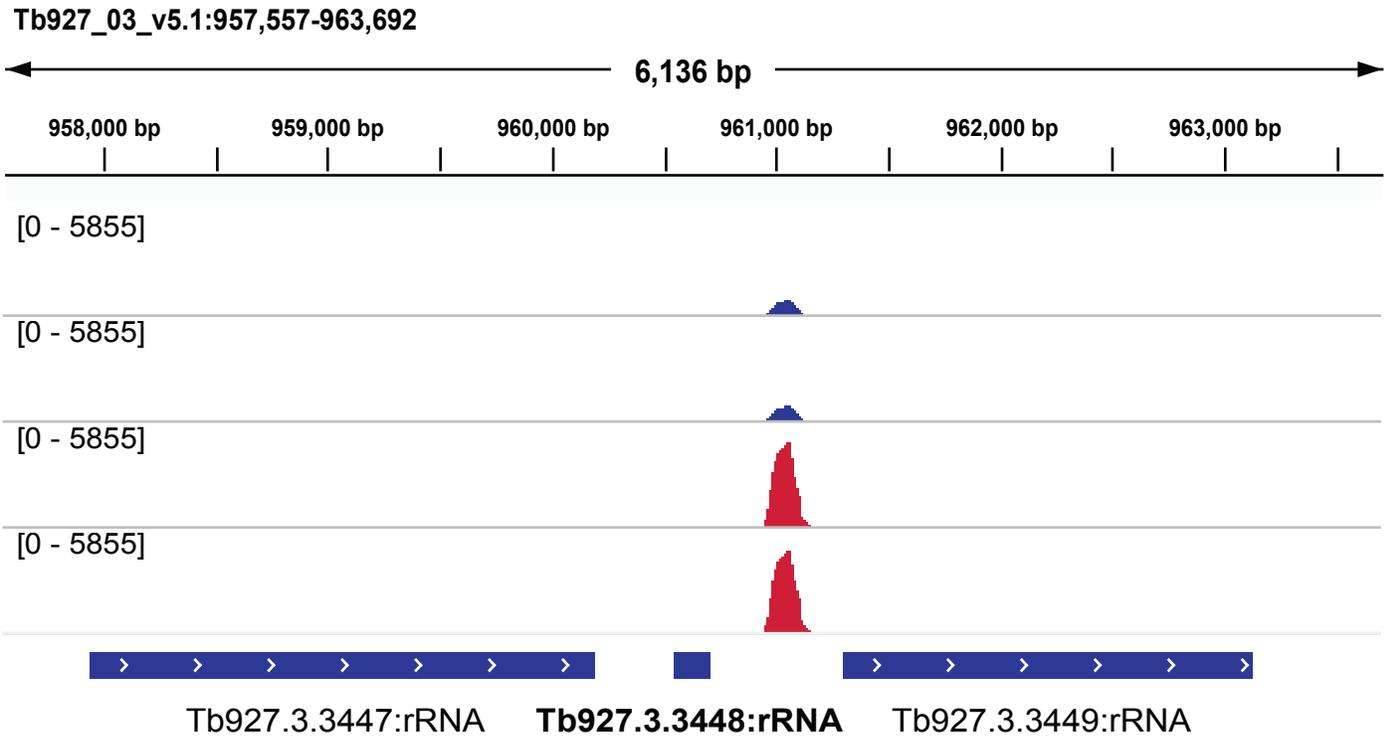


Supplemental Figure 3

Cluster 0 vs Cluster 7 Markers

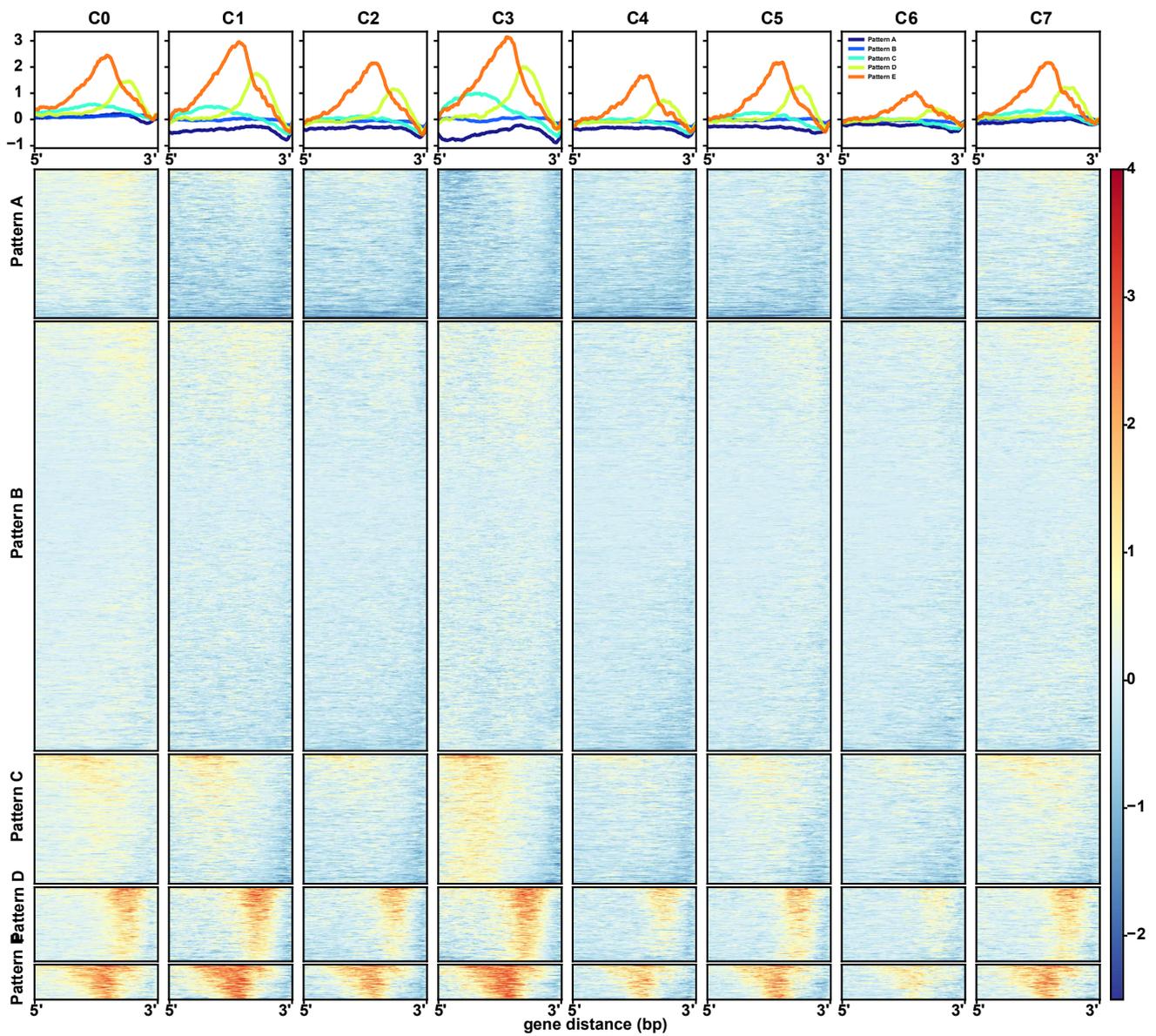


Supplemental Figure 4



Supplemental Figure 5

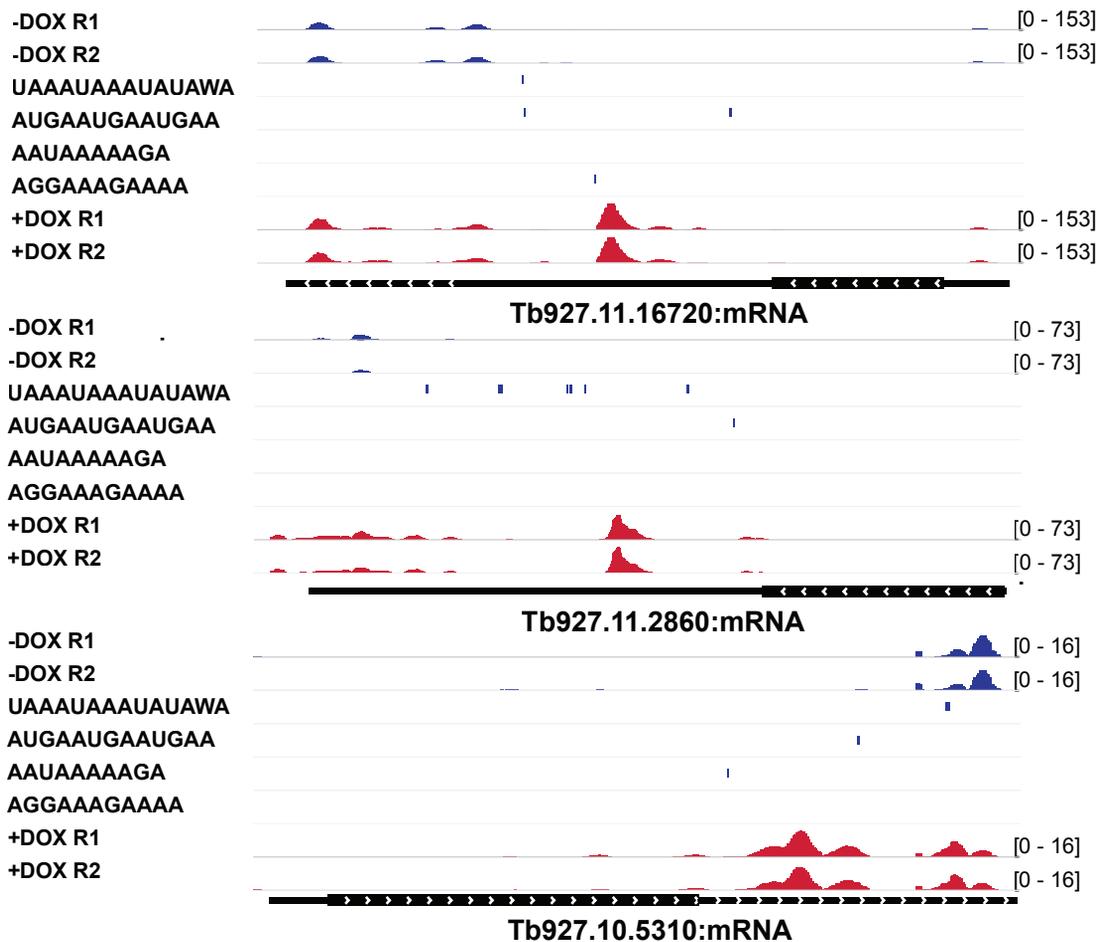
A



B

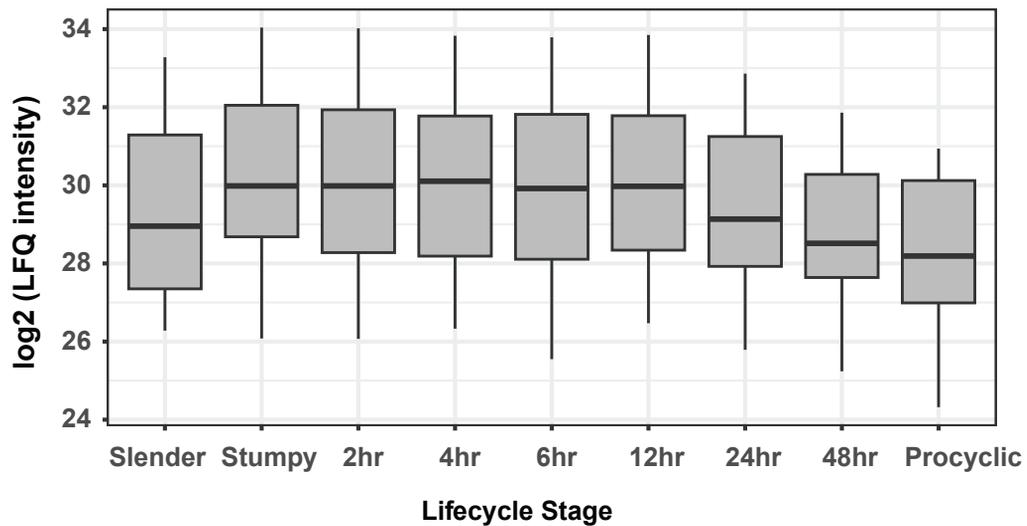
	Pattern A 2081 genes	Pattern B 6016 genes	Pattern C 1815 genes	Pattern D 1039 genes	Pattern E 488 genes
Cluster 0 236 genes	2 overlap $p = 1$	7 overlap $p = 1$	25 overlap $p = 0.99$	65 overlap $p = 4.35e-17$	133 overlap $p = 1.22e-124$
Cluster 1 2 genes	2 overlap $p = 0.03$	0 overlap $p = 1$	0 overlap $p = 1$	0 overlap $p = 1$	0 overlap $p = 1$
Cluster 2 84 genes	43 overlap $p = 3.75e-12$	21 overlap $p = 1$	10 overlap $p = 0.86$	8 overlap $p = 0.48$	2 overlap $p = 0.87$
Cluster 3 2 genes	1 overlap $p = 0.33$	1 overlap $p = 0.77$	0 overlap $p = 1$	0 overlap $p = 1$	0 overlap $p = 1$
Cluster 4 74 genes	36 overlap $p = 1.32e-09$	21 overlap $p = 1$	7 overlap $p = 0.96$	8 overlap $p = 0.34$	2 overlap $p = 0.82$
Cluster 5 402 genes	93 overlap $p = 0.01$	158 overlap $p = 1$	79 overlap $p = 0.016$	55 overlap $p = 8.99e-4$	17 overlap $p = 0.52$
Cluster 6 175 genes	56 overlap $p = 3.96e-06$	68 overlap $p = 1$	28 overlap $p = 0.48$	13 overlap $p = 0.80$	7 overlap $p = 0.61$
Cluster 7 167 genes	5 overlap $p = 1$	37 overlap $p = 1$	22 overlap $p = 0.84$	37 overlap $p = 1.58e-07$	64 overlap $p = 3.13e-45$

Supplemental Figure 6

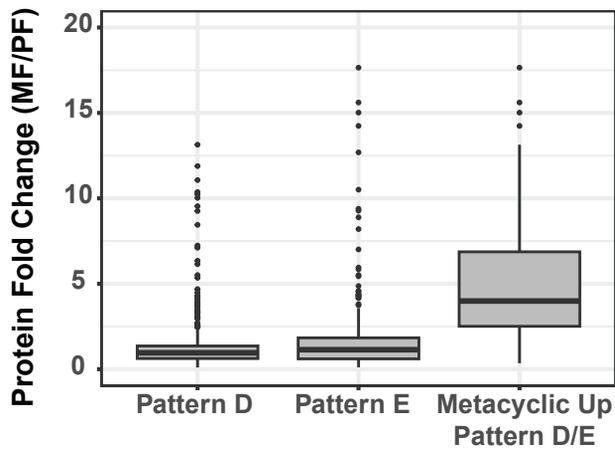


Supplemental Figure 7

A



B



Supplemental Figure 8

A

	Slender 57 genes	Stumpy 103 genes	Early PF 34 genes	Late PF 13 genes	MF Up 703 genes	MF Down 609 genes
DRBD18 Bound 579 genes	1 overlap p = 0.96	32 overlap p = 1.42e-17	0 overlap p = 1	0 overlap p = 1	141 overlap p = 3.39e-50	13 overlap p = 1

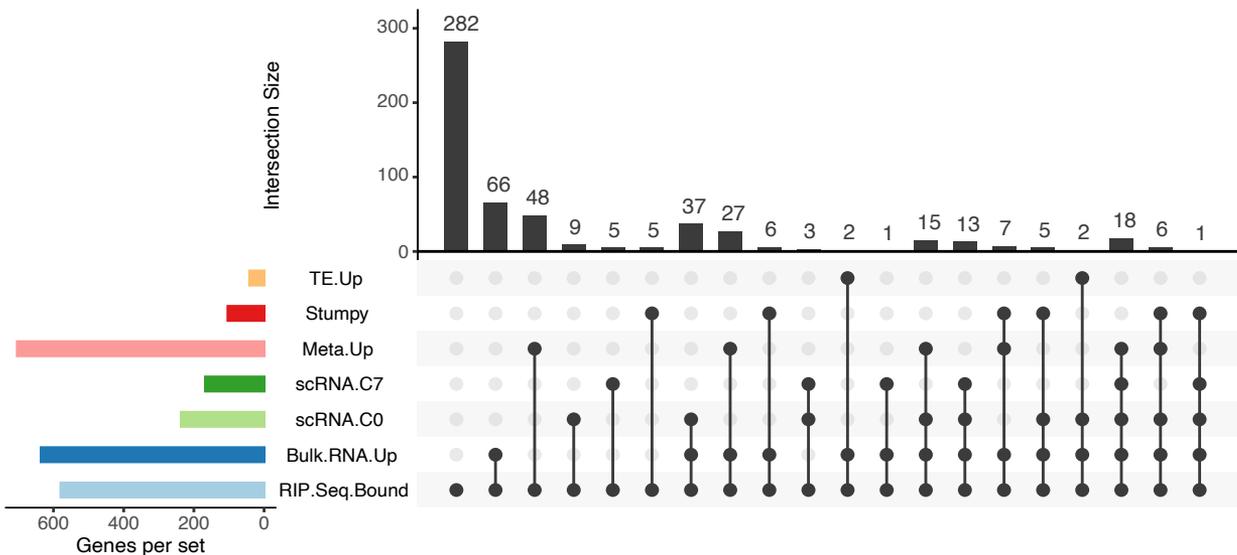
B

	Export impaired 95 genes	TE up 41 genes	TE down 61 genes	RNAseq Upregulated 635 genes	RNAseq Downregulated 363 genes
DRBD18 Bound 579	2 overlap p = 0.95	10 overlap p = 2.40e-05	0 overlap p = 1	218 overlap p = 2.41e-134	15 overlap p = 0.81

C

	Cluster 0 236 genes	Cluster 1 2 genes	Cluster 2 84 genes	Cluster 3 2 genes	Cluster 4 74 genes	Cluster 5 402 genes	Cluster 6 175 genes	Cluster 7 167 genes
DRBD18 Bound 579	117 overlap p = 1.02e-90	0 overlap p = 1	6 overlap p = 0.24	0 overlap p = 1	8 overlap p = 0.03	7 overlap p = 1	5 overlap p = 0.94	57 overlap p = 5.7e-33

D



Supplemental Figure Legends

Fig. S1. DRBD18 RNAi leads to increased abundance of many transcripts encoding kinases, phosphatases, and RNA binding proteins (RBP). Volcano plot depicts differential expression analysis of bulk RNAseq from two published DRBD18 RNAi experiments (1, 2) comparing a total of five replicate uninduced and induced (DRBD18 depleted) samples. Vertical dotted lines indicate two-fold changes in transcript abundance upon DRBD18 RNAi; horizontal dotted line indicates $p_{adj} < 0.05$.

Fig. S2. Reproducibility of single cell RNAseq replicates. Correlation analysis of average gene expression across all genes, between each pairwise sample.

Fig. S3. Volcano plot depicting differences in cluster markers between Clusters 0 and 7. Cutoffs as in Fig. S1. Red, VSG genes.

Fig. S4. Reads covering two 5.8S rRNA intergenic regions increase upon DRBD18 depletion. Normalized total read counts from two -DOX and two +DOX replicates are shown.

Fig. S5. Clusters 0 and 7 are enriched for transcripts with proximal poly(A) sites. (A) Heatmap \log_2 ratio profiles (top) and gene expression across the gene body for each DeepTools pattern, across all single cell clusters. Positive values (red) indicate increased abundance in the +DOX sample, while negative values (blue) indicate increased abundance in the -DOX sample. (B) Hypergeometric test of the overlap between scRNAseq clusters and distinct 3' UTR pattern.

Fig. S6. 3'UTR motifs would be lost upon proximal polyadenylation site generation in DRBD18 depleted cells. STREME was used identify enriched motifs in 3'UTRs (from the stop codon to the 5' end of the downstream gene) of Patterns D and E transcripts, with Pattern A transcripts serving as negative controls. Shown are examples of motifs that would be lost upon proximal polyadenylation site generation when DRBD18 is depleted.

Fig. S7. Upregulation of stumpy BF and MF transcripts with altered polyadenylation sites is reflected at the protein level. (A) The 56 stumpy BF specific transcripts that exhibited altered polyadenylation upon DRBD18 knockdown (Patterns D and E) were compared to the corresponding protein levels over a published differentiation time course (3), establishing that increased transcript levels in stumpy BF are reflected in increased protein levels. (B) The 198 MF specific transcripts that exhibited altered polyadenylation upon DRBD18 knockdown (Patterns D and E), as well as the entire Patterns D and E datasets, were compared to the published MF/PF protein ratio (4). This establishes that increased transcript levels in MF are reflected in increased protein levels.

Fig. S8. DRBD18 bound transcripts significantly overlap with numerous gene regulatory parameters. (A) Hypergeometric tests demonstrate the significance of the overlap between DRBD18 bound transcripts and life cycle specific transcripts. (B) As in A, but showing overlap with transcripts whose export, translation, or abundance is regulated by DRBD18. (C) As in A, but showing overlap with scRNAseq cluster markers. (D) UpsetR plot illustrating the intersections between multiple gene regulatory groups.

Supplemental References

1. A. Mishra *et al.*, Selective nuclear export of mRNAs is promoted by DRBD18 in *Trypanosoma brucei*. *Molecular microbiology* **116**, 827-840 (2021).
2. K. Lott *et al.*, Arginine methylation of DRBD18 differentially impacts its opposing effects on the trypanosome transcriptome. *Nucleic acids research* **43**, 5501-5523 (2015).
3. M. Dejung *et al.*, Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of *Trypanosoma brucei*. *PLoS pathogens* **12**, e1005439 (2016).
4. R. Christiano *et al.*, The proteome and transcriptome of the infectious metacyclic form of *Trypanosoma brucei* define quiescent cells primed for mammalian invasion. *Molecular microbiology* **106**, 74-92 (2017).