Cell Reports, Volume *25*

Supplemental Information

DAZL Regulates Germ Cell Survival through a Network of PolyA-Proximal mRNA Interactions

Leah L. Zagore, Thomas J. Sweet, Molly M. Hannigan, Sebastien M. Weyn-Vanhentenryck, Raul Jobava, Maria Hatzoglou, Chaolin Zhang, and Donny D. Licatalosi

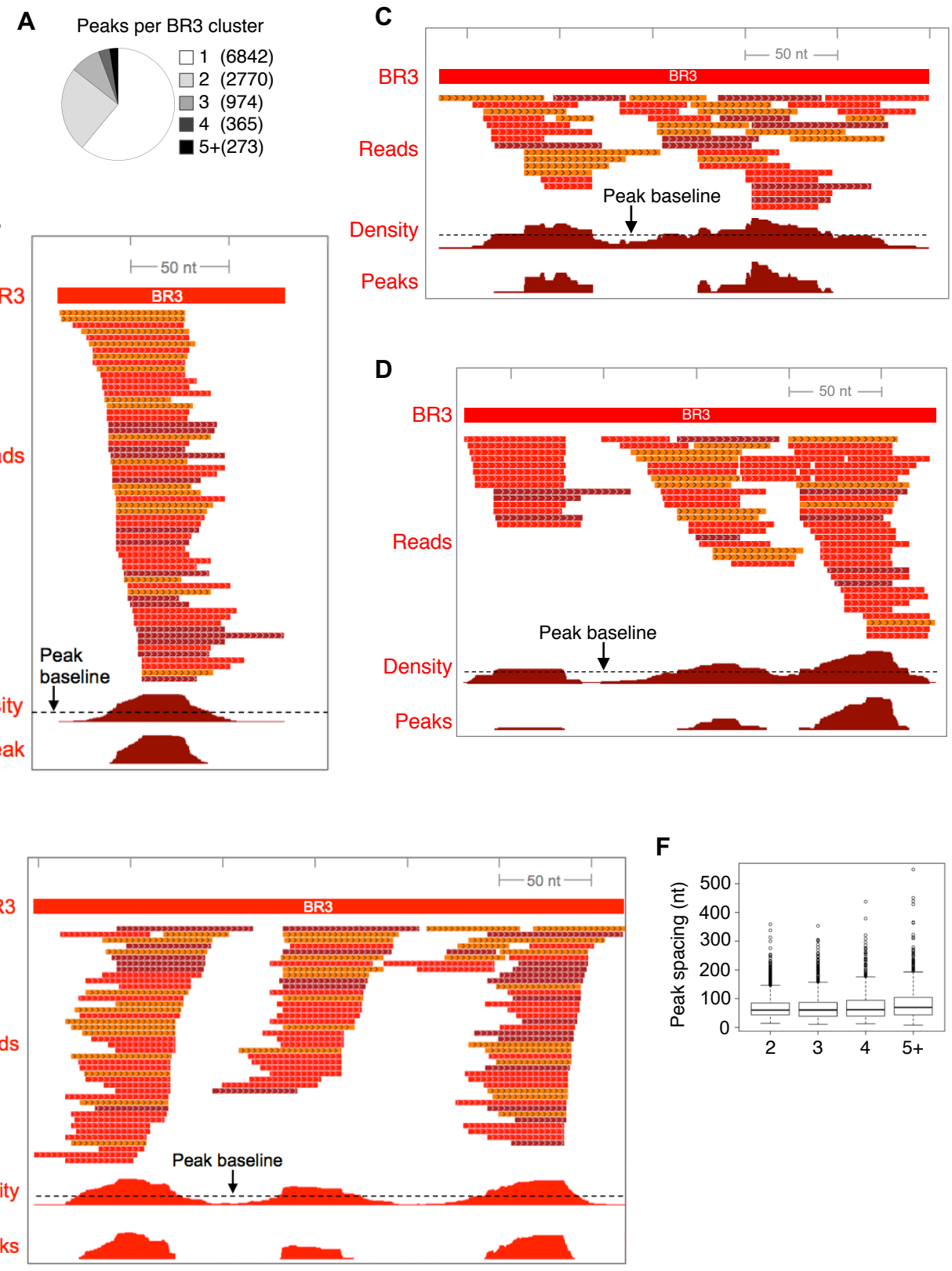

A
Peaks per BR3 cluster
1 (6842)
2 (2770)
3 (974)
4 (365)
5+(273)
B
50 nt
BR3
BR3
Reads
Peak baseline
Density
Peak
C
50 nt
BR3
BR3
Reads
Peak baseline
Density
Peaks
D
50 nt
BR3
BR3
Reads
Peak baseline
Density
Peaks
E
50 nt
BR3
BR3
Reads
Peak baseline
Density
Peaks
F
Peak spacing (nt)
500
400
300
200
100
0
2
3
4
5+

Figure S1. Identification and analysis of DAZL BR3 CLIP peaks, Related to Figure 1. (A) Number of BR3 clusters with different numbers of peaks. BR3 clusters containing one (B), two (C), or three (D, E) peaks are shown. The footprint of the BR3 region is shown at the top. Reads are color coded to reflect each of the three CLIP libraries sequenced. Density indicates the number of overlapping reads per nucleotide within the BR3 region. Peak baseline corresponds to the total number of CLIP bases in the cluster divided by the length of the cluster. Peaks correspond to regions where the CLIP density exceeds the peak baseline. (F) Distances between the midpoints of adjacent peaks in multi-peak clusters.

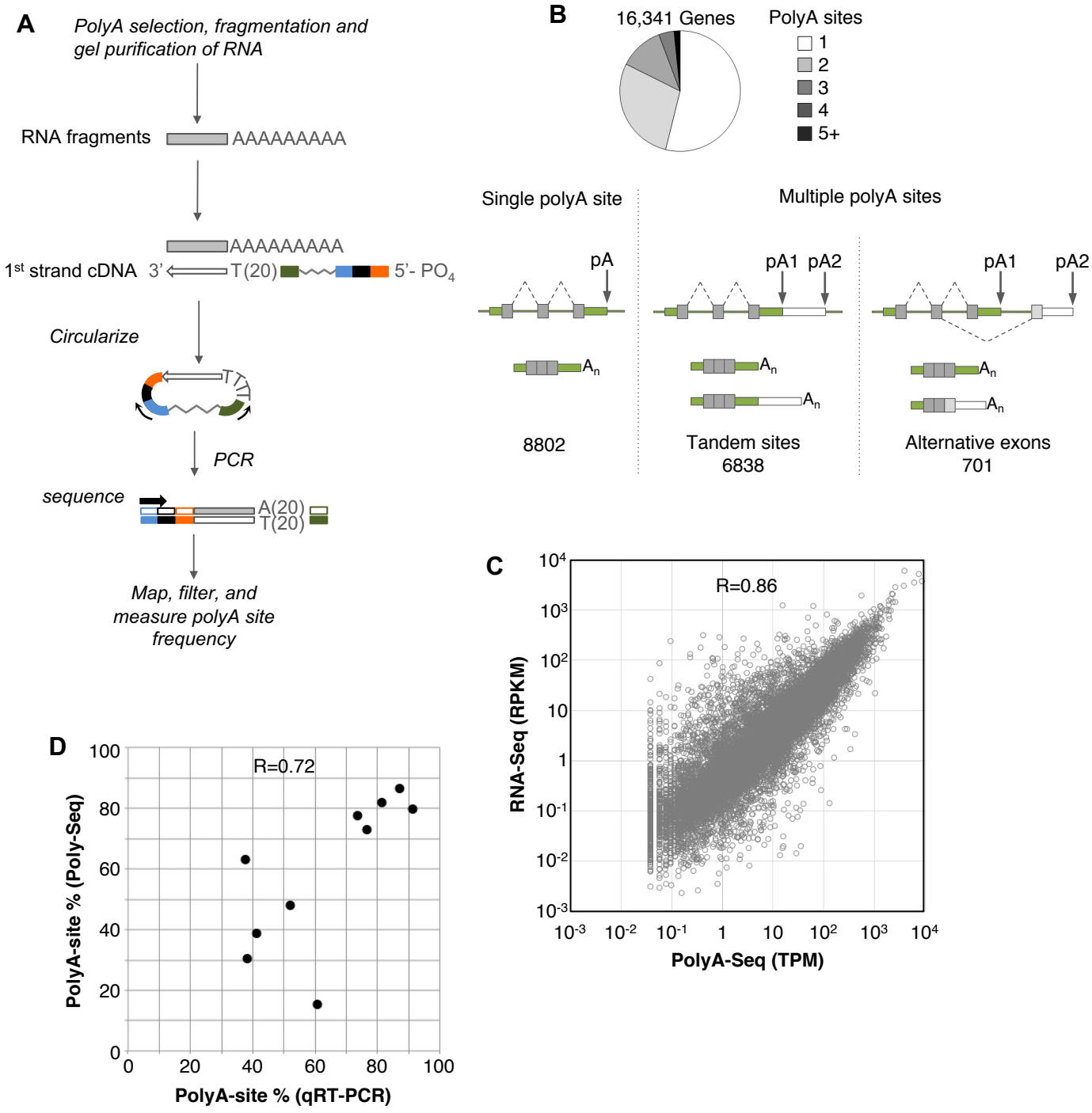
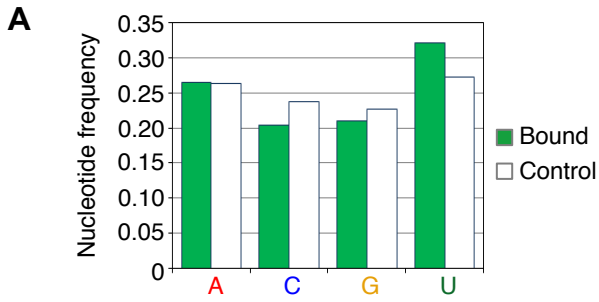

A
PolyA selection, fragmentation and gel purification of RNA
RNA fragments
AAAAAAAAA
AAAAAAAAA
1st strand cDNA
3'
T(20)
5'- PO4
Circularize
PCR
sequence
A(20)
T(20)
Map, filter, and measure polyA site frequency
B
16,341 Genes
PolyA sites
1
2
3
4
5+
Single polyA site
Multiple polyA sites
pA
pA1
pA2
An
8802
Tandem sites
6838
Alternative exons
701
C
R=0.86
RNA-Seq (RPKM)
PolyA-Seq (TPM)
D
R=0.72
PolyA-site % (Poly-Seq)
PolyA-site % (qRT-PCR)

Figure S2. Identification of testis 3'UTRs by PolyA-seq, Related to Figure 2. (A) Overview of the PolyA-Seq library construction strategy. (B) Characterization of polyA site usage in different genes. (C) Comparison of RPKM and TPM values derived from RNA-Seq and PolyA-Seq, respectively, from adult testes. (D) Comparison of proximal polyA site estimates from qRT-PCR and PolyA-Seq.


A
Nucleotide frequency
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
A
C
G
U
Bound
Control

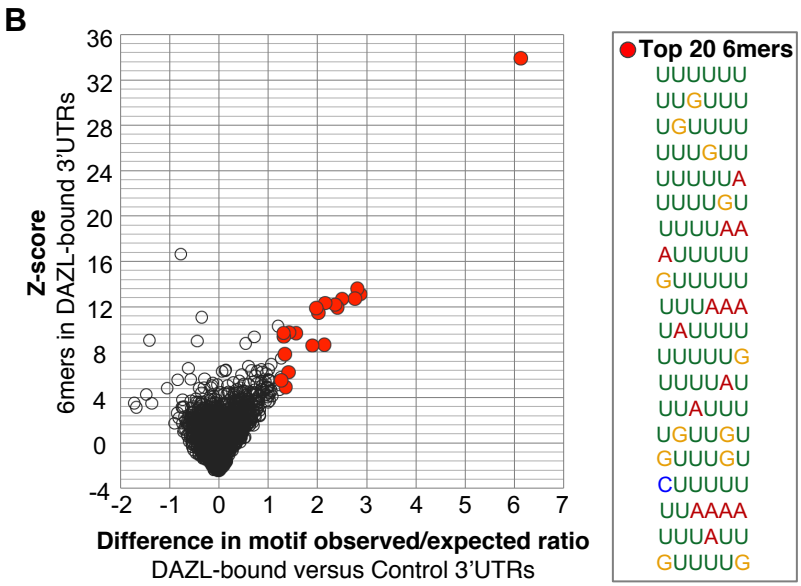

B
Z-score
6mers in DAZL-bound 3'UTRs
36
32
28
24
20
16
12
8
4
0
-4
-2
-1
0
1
2
3
4
5
6
7
Difference in motif observed/expected ratio
DAZL-bound versus Control 3'UTRs
Top 20 6mers
UUUUUU
UUGUUU
UGUUUU
UUUGUU
UUUUUA
UUUUGU
UUUUAA
AUUUUU
GUUUUU
UUUAAA
UAUUUU
UUUUUG
UUUUAU
UUAUUU
UGUUGU
GUUUGU
CUUUUU
UUAAAA
UUUAUU
GUUUUG

Figure S3. Features associated with DAZL-bound and control 3'UTRs, Related to Figure 2. (A) Nucleotide frequency in the last 500 nucleotides of 3'UTRs with DAZL BR3 sites (n=2022) and control 3'UTRs that have no CLIP tags (n=3771). (B) Enrichment of 6-mers present in the 3'UTRs of genes with DAZL-3'UTRs compared to control genes. Each dot represents an individual 6 mer. X-axis corresponds to difference in observed-to-expected ratio for bound versus control 3'UTRs, while y-axis corresponds to the enrichment z-score for the 6-mer in the DAZL-bound 3'UTRs. Shown at right are the top 20 6-mers ranked by difference in the observed-to-expected ratios between bound and control 3'UTRs.

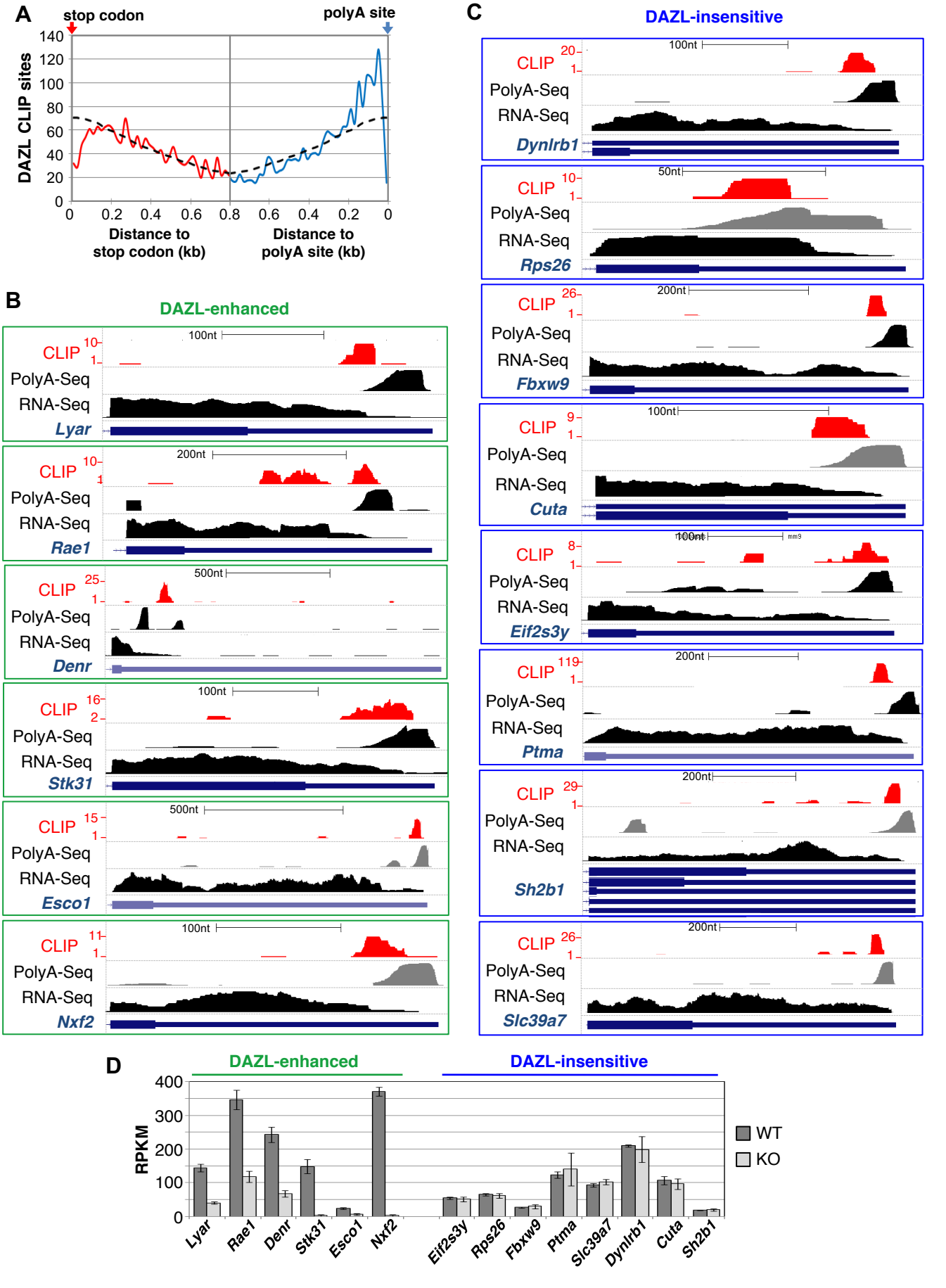
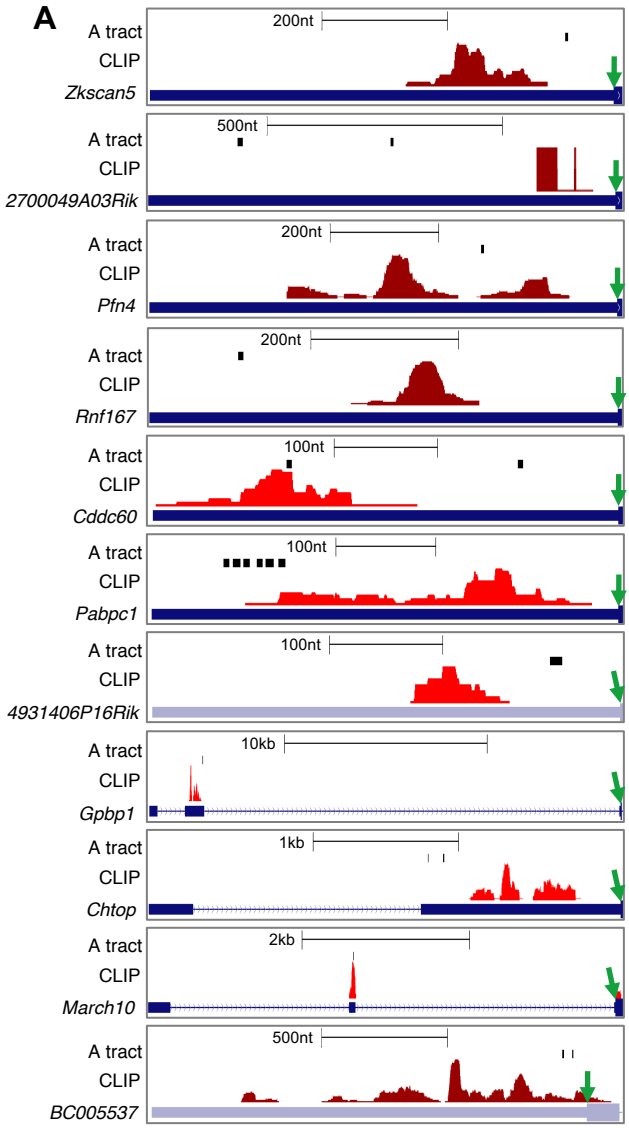
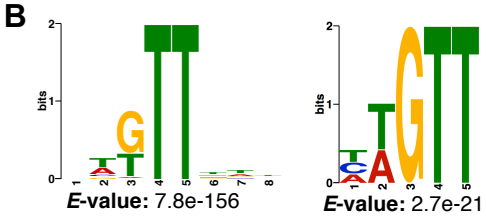
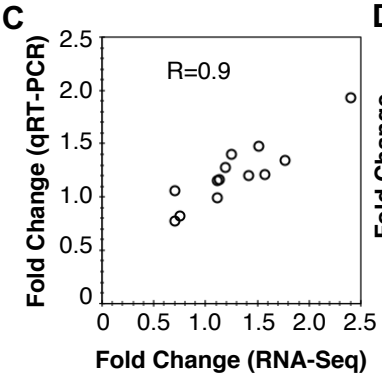
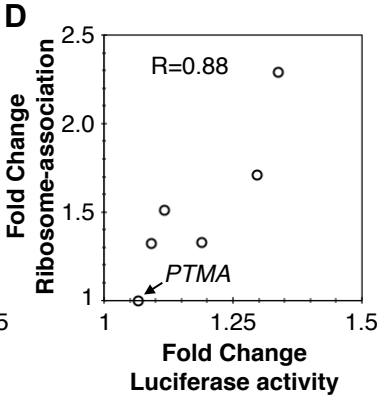

A
stop codon
polyA site
DAZL CLIP sites
Distance to stop codon (kb)
Distance to polyA site (kb)
B
DAZL-enhanced
C
DAZL-insensitive
CLIP
PolyA-Seq
RNA-Seq
Lyar
Rae1
Denr
Stk31
Esco1
Nxf2
Dynlrb1
Rps26
Fbxw9
Cuta
Eif2s3y
Ptma
Sh2b1
Slc39a7
D
DAZL-enhanced
DAZL-insensitive
RPKM
WT
KO

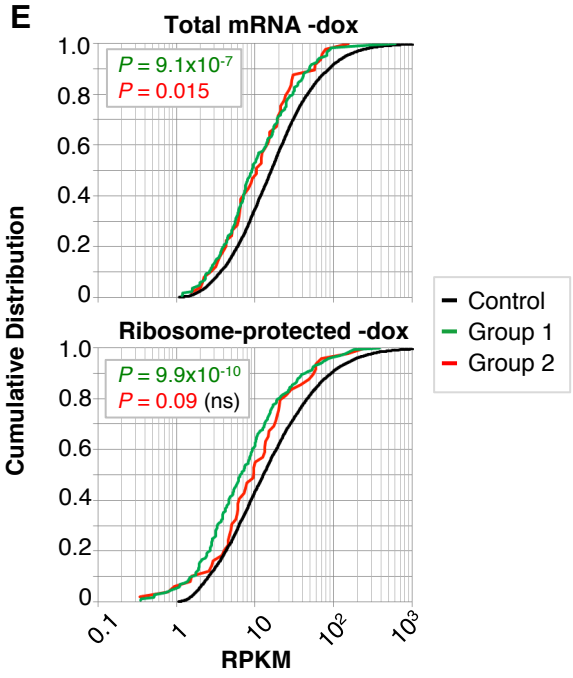
Figure S4. Examination of P6 DAZL-RNA binding in 3'UTRs, Related to Figure 4. (A) Metagene analysis of the distribution of BR3 DAZL-RNA sites (from P6 iCLIP) in 3'UTRs of spermatogonial-expressed 3'UTRs (as in Figure 2C). Genome browser images showing P6 BR3 CLIP sites in 3'UTRs of DAZL-enhanced and DAZL-insensitive genes (B and C, respectively). Adult PolyA-seq and RNA-Seq tracks are also shown. CLIP scale denotes the maximum number of overlapping CLIP tags in the region shown. (D) Average RPKM values and standard deviations for the genes shown in panels (A) and (B).

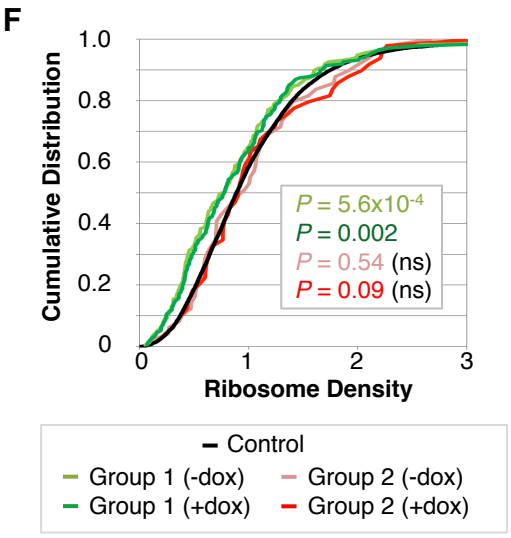

A
A tract
CLIP
Zkscan5
200nt
2700049A03Rik
500nt
Pfn4
Rnf167
Cddc60
100nt
Pabpc1
4931406P16Rik
Gpbp1
10kb
Chtop
1kb
March10
2kb
BC005537


B
bits
E-value: 7.8e-156
E-value: 2.7e-21


C
R=0.9
Fold Change (qRT-PCR)
Fold Change (RNA-Seq)


D
R=0.88
PTMA
Fold Change
Ribosome-association
Fold Change
Luciferase activity


E
Total mRNA -dox
P = 9.1x10-7
P = 0.015
Ribosome-protected -dox
P = 9.9x10-10
P = 0.09 (ns)
Cumulative Distribution
RPKM
Control
Group 1
Group 2


F
Cumulative Distribution
P = 5.6x10-4
P = 0.002
P = 0.54 (ns)
P = 0.09 (ns)
Ribosome Density
Control
Group 1 (-dox)
Group 2 (-dox)
Group 1 (+dox)
Group 2 (+dox)

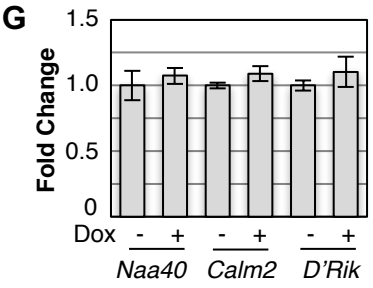
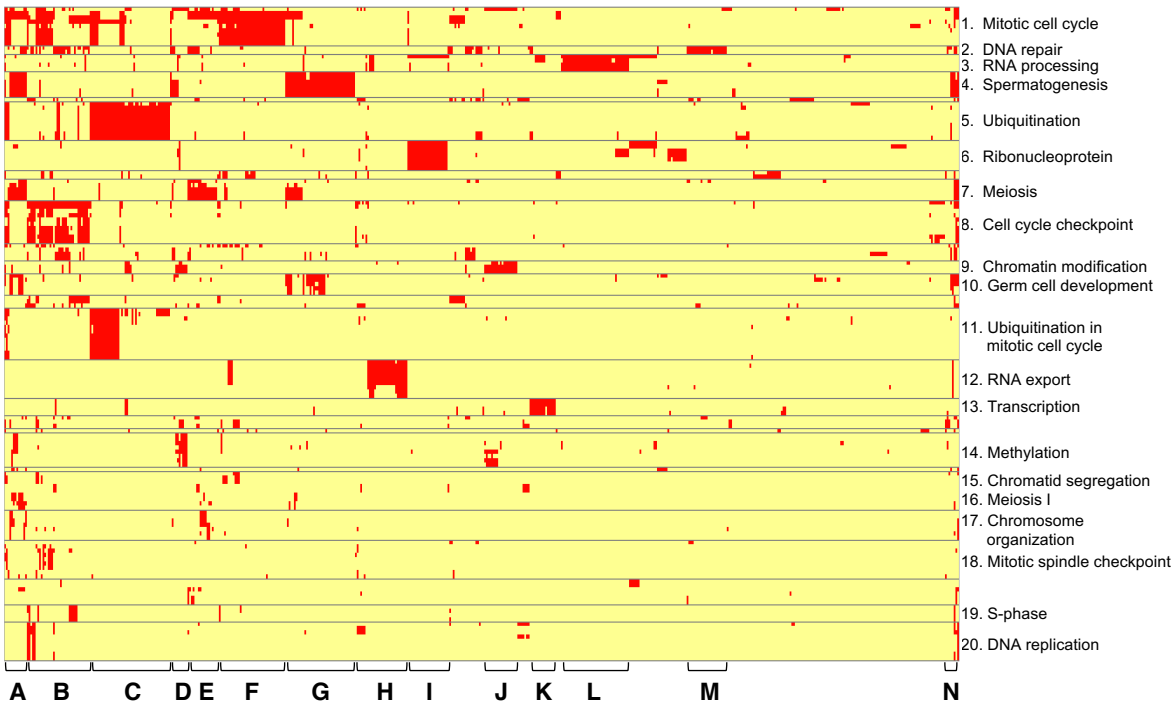

G
Fold Change
Dox
Naa40
Calm2
D'Rik

Figure S5. DAZL binding in 5'UTRs and consequences of DAZL expression in GC-1 spg cells, Related to Figures 5 and 6. (A) Genome browser images of 11 5'UTRs with DAZL BR3 CLIP clusters (adult testis) and polyA tracts of 5 A's or more. PolyA tracts indicated by black boxes. Green arrows denote the start codon. (B) The most significant motifs (identified by the programs MEME and DREME, left and right respectively) in 10nt regions centered at the 5'end of DAZL BR3 CLIP sites from GC-1 spg cells. (C) Comparison of dox-induced RNA level fold changes determined by qRT-PCR and RNA-Seq (13 genes represented). (D) Comparison of dox-induced fold changes in luciferase activity for mRNA reporters bearing different 3'UTRs (x-axis) versus fold changes in ribosome-association for the corresponding endogenous transcripts determined by deep sequencing (y-axis). (E) Cumulative distribution of RPKM values for control, Group 1 (enhanced), and Group 2 (repressed) genes (black, green, red lines, respectively) in untreated GC-1 spg cells. *P* values (Wilcoxon rank sum test) correspond to pairwise comparisons of Group 1 (green) or Group 2 (red) RPKM values to control. (F) Cumulative distribution of ribosome density values for Group 1 and 2 genes before and after dox-treatment. Ribosome density is the ratio of RPKM values from ribosome-protected fragments to total RNA for each gene. (G) qRT-PCR analysis of pre-mRNA levels following dox treatment relative to pre-mRNA levels in untreated cells. Error bars represent standard deviations for three technical replicates per gene.

Zagore_Fig_S6



A
Ccnb1
Mad2l1
Plk1
Sycp3
Ehmt2
Mael
Piwil2
Tdrd9
Spo11
Dmc1
Trip13
Rad51c
Tex15

B
Tipin
Cdt1
Rad17
Clspn
Nae1
Ddca5
Chmp1a
Atm
Zw10
Apc
Bub1
Kntc1
Pcid2
Cenpe
Cep192
Mre11a
Ercc3
Dtl
Usp28
Ints3
Uimc1
Xpc
Mapk14
Obfc2a
Apbb1
Crlf3
Rb1
Apex1
Dbf4
Fbxo31
Ccnd1
Hras1
Rfwd3
Rbm38
Dgkz
Hinfp

C
Ube2c
Bub1b
Cdc20
Anapc1
Anapc4
Psmd13
Psme3
Psmd5
Psmd2
Psmd11
Psmd1
Psmc5
Psmc1
Psmb7
Psmb3
Psmb1
Psmb2
Chfr
Ccnf
Usp9x
Taf1
Taf9
Hdac6
Usp3
Arrb2
Pan2
Gclc
Ube4b
Usp14
Uchl1
Usp34
Usp24
2700078e11Rik
Cul5
Usp7
Usp26
Siah1b
Uba6
Usp19
Wwp2
Usp5
Rnf216
Rnf180
Nub1
Cdc34
Fbxo2

D
Herc2
Fancd2
Prkdc
Kdm1b
Prmt7
Suv39h1
Smarca4
Baz2a
Dnmt1
Dnmt3b

E
Msh6
Topbp1
Exo1
Pms2
Rpa1
Rad50
Rad51
Tex11
Cpeb1
Syce2
Sycp1
Sgol2
Stag3
Fanca
Smc1b
Klhdc3
Psmc3ip

F
Cenpf
Rcc1
Hells
Kif18a
Espl1
Nek2
Ahctf1
Nup37
Nup43
Ncapg2
Ncapg
Ncapd2
Ncapd3
Timeless
Nek1
Papd5
Kif2c
Ndc80
Kif20b
Kif11
Haus4
Haus8
Nudc
Klhl21
Dsn1
Incenp
Cdca8
Cdc6
Aurkb
Zc3hc1
Wee1
Rcc2
Nek3
Mastl
Fsd1
Cep55
Cdca2
Aurka
Ccna2

G
Dazl
Ccnb1ip1
Mei1
Piwil1
Ccna1
Ddx4
Fkbp6
Clgn
Asz1
Boll
Celf1
Rad18
Dnd1
Tdrd7
Kit
Zbtb16
Taf7l
Mast2
Sox3
Hook1
Strbp
Adad1
Rnf17
Htt
Hmga1
Zfp541
Txnrd3
Tle3
Spata5
Sohlh2
Sbf1
Rnf114
Nr6a1
Nanos3
Morc1
Mea1
Gag2
Gsr
Cadm1
D1pas1

H
Xpo1
Thoc6
Ddx39
Thoc1
Xpo7
Eif5a2
Ranbp2
Ranbp17
Nupl1
Nup93
Nup88
Nup50
Nup210
Nup188
Lrpprc
Nup155
Xpot
Eif5a
Nxf2
Nup160
Nup133
Fyttd1
Gle1

I
Pes1
Pdcd11
Ftsj3
Bop1
Wdr55
Utp20
Utp18
Tbl3
Rrp9
Rrp1b
Rbfa
Pop5
Nop2
Nop14
Nol9
Nhp2
Imp4
Exosc8
Emg1
Dis3
Ddx56
Dcaf13
Ddx51

J
Eed
Jarid2
Ash2l
Cxxc1
Suv39h2
Whsc1
Ezh1
Wdr5
Eya2
Hat1
Huwe1
Taf6l
Noc2l
Myst2
Kat2a
Yeats4
Phf17
Actl6a
Phf16

K
Ccnh
Gtf2h3
Taf6
Polr2g
Polr2e
Polr2b
Polr2a
Gtf2f1
Gtf2f2
Ell
Taf5
Taf4b
Gtf2e2
Gtf2a1
Gtf2e1

L
Mov10
Srpk1
Tfip11
Sugp1
Srsf9
Snrnp40
Snrnp25
Sf3b3
Sf3a3
Sf3a1
Puf60
Prpf3
Ppp1r8
Ppil1
Plrg1
Phf5a
Naa38
Lsm4
Lsm3
Hnrnpu
Esrp2
Esrp1
Eftud2
Ddx41
Cstf2
Cstf1
Cpsf3
Aqr
Cpsf1
Cpsf4
Rod1
Wdr77
Lsm2
Zrsr2
Snrpe
Sip1
Gemin5
Ddx20
Gemin4

M
Ung
Lig1
Rfc4
Rfc5
Bard1
Parg
Esco1
Esco2
Palb2
Tdp2
Setx
Tdp1
Smc5
Zswim7
Cib1
Uhrf1
Trp53bp1
Fanci
Brip1
Eef1e1
Parp1
Chd1l
Mum1

N
Bax
Ddx25
Brca2
Msh2
Stra8

Figure S6. Clusters of enriched GO terms associated with *Dazl*-enhanced genes, Related to Figure 3. (A) Hierarchical clustering of enriched GO terms (P<0.01) and genes associated with the set of 1462 genes that have decreased RNA levels in *Dazl* KO cells. Twenty clusters of related terms are shown at left, and different groups of genes (A-N) shown at the bottom.

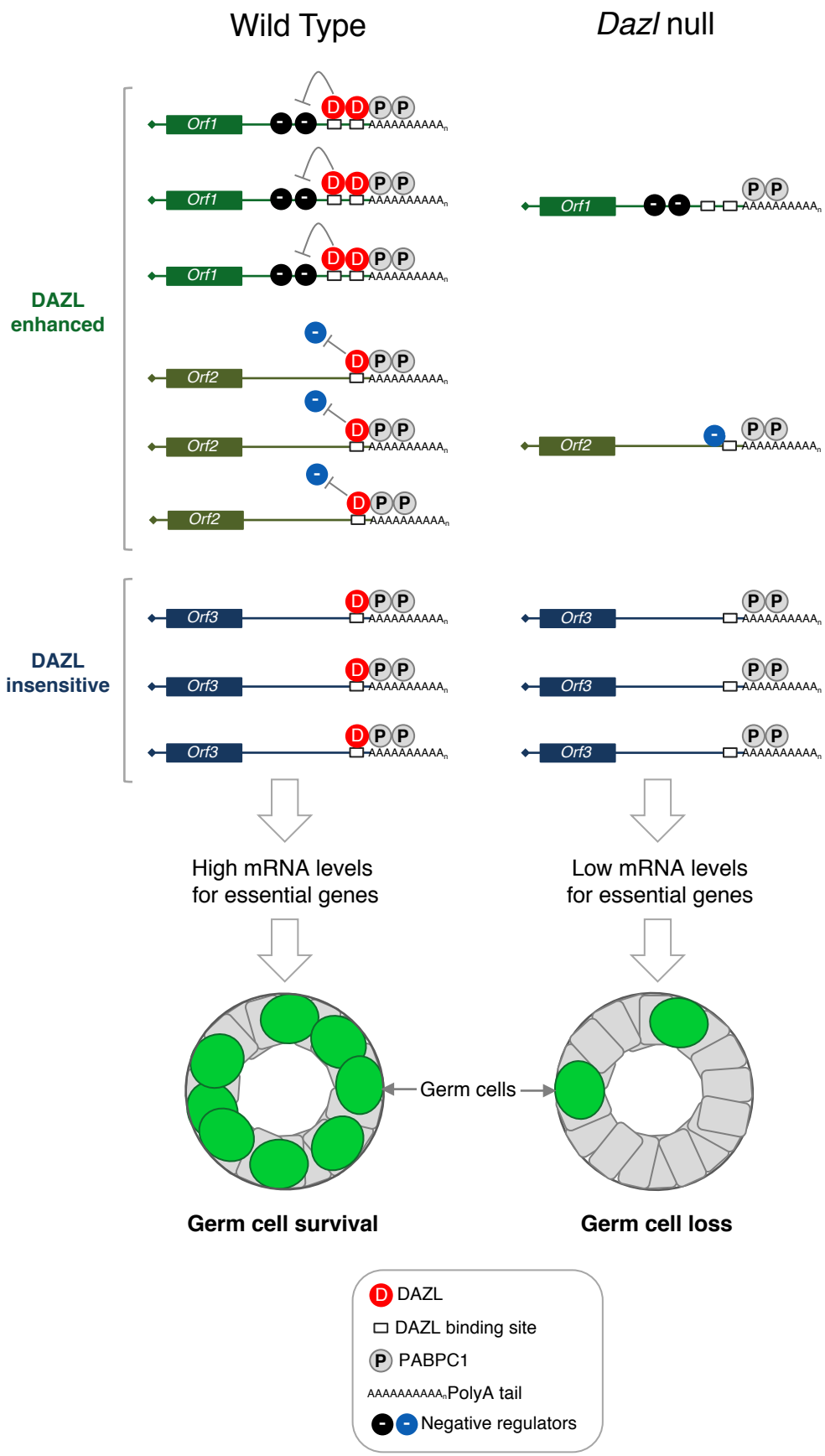

Wild Type
Dazl null
DAZL enhanced
DAZL insensitive
Orf1
Orf2
Orf3
High mRNA levels for essential genes
Low mRNA levels for essential genes
Germ cells
Germ cell survival
Germ cell loss
DAZL
DAZL binding site
PABPC1
PolyA tail
Negative regulators

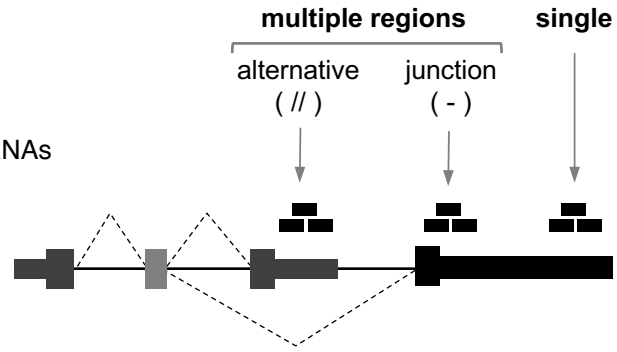
Figure S7. Proposed model for *Dazl*-dependent germ cell maintenance, Related to Figures 1 to 7. DAZL binds a broad set of mRNAs via polyA-proximal GUU interactions facilitated by Pabpc1-polyA interactions. DAZL maintains high mRNA levels for select targets (*Orf1* and *Orf2*), potentially by blocking the function or binding of negative regulatory factors. In the absence of DAZL, *Orf1* and *Orf2* mRNAs are reduced, whereas *Orf3* mRNAs (not subject to negative regulation) are unaffected.

9524 / 11297 map to annotated genes in RefSeq

237 map to non-coding RNAs
("NR_" RefSeq IDs)

225 map to overlapping coding and non-coding RNAs
("NR_" and "NM_" RefSeq IDs)

9062 map to protein coding genes
("NM_" RefSeq IDs only)



Genic location	BR3 Clusters	% BR3	5'UTR	3'UTR	CDS	Intron
3UTR	6397	70.59		6397		
3UTR // 5UTR	3	0.03	3	3		
3UTR // CDS-3UTR	10	0.11		10	10	
3UTR // CDS-INTRON	1	0.01		1	1	1
3UTR // INTRON	157	1.73		157		157
3UTR // INTRON-3UTR	5	0.06		5		5
3UTR // INTRON-CDS-3UTR	3	0.03		3	3	3
3UTR-INTRON	1	0.01		1		1
5UTR	57	0.63	57			
5UTR // CDS	1	0.01	1		1	
5UTR // INTRON	7	0.08	7			7
5UTR-CDS	37	0.41	37		37	
5UTR-CDS // INTRON	3	0.03	3		3	3
5UTR-CDS // INTRON-5UTR-CDS	3	0.03	3		3	3
5UTR-CDS-INTRON	3	0.03	3		3	3
5UTR-INTRON	4	0.04	4			4
CDS	741	8.18			741	
CDS // 3UTR	1	0.01		1	1	
CDS // INTRON	16	0.18			16	16
CDS // INTRON-CDS	1	0.01			1	1
CDS-3UTR	775	8.55		775	775	
CDS-3UTR // CDS-3UTR-INTRON	2	0.02		2	2	2
CDS-3UTR // CDS-INTRON	4	0.04		4	4	4
CDS-3UTR // INTRON	15	0.17		15	15	15
CDS-3UTR-INTRON	5	0.06		5	5	5
CDS-INTRON	226	2.49			226	226
INTRON	340	3.75				340
INTRON-3UTR	10	0.11		10		10
INTRON-3UTR // INTRON-CDS-3UTR	1	0.01		1	1	1
INTRON-5UTR	2	0.02	2			2
INTRON-5UTR // 5UTR	1	0.01	1			1
INTRON-5UTR-CDS	6	0.07	6		6	6
INTRON-5UTR-CDS-INTRON	3	0.03	3		3	3
INTRON-5UTR-INTRON	1	0.01	1			1
INTRON-CDS	126	1.39			126	126
INTRON-CDS // 3UTR	2	0.02		2	2	2
INTRON-CDS // INTRON-CDS-3UTR	1	0.01		1	1	1
INTRON-CDS-3UTR	62	0.68		62	62	62
INTRON-CDS-3UTR-INTRON	1	0.01		1	1	1
INTRON-CDS-INTRON	28	0.31			28	28

Supplemental Table 1. Annotation of 11,297 BR3 clusters in RefSeq fragments, Related to Figure 1. Clusters that overlap two different fragments in the same transcript (such as coding and UTR sequence) are denoted with "-". Clusters that map to positions with alternative annotations (such as intron in one isoform and exon in another) are denoted with "/".