SUPPLEMENTAL MATERIAL for

An Atlas of Chromatoid Body Components

Oliver Meikar, Vasily V. Vagin, Frédéric Chalmel, Karin Sõstar, Aurélie Lardenois, Molly Hammell, Ying Jin, Matteo Da Ros, Kaja A. Wasik, Jorma Toppari, Gregory J. Hannon, Noora Kotaja

SUPPLEMENTAL MATERIALS AND METHODS

Immunoblotting

Proteins were separated by 10% SDS–PAGE and electroblotted to PVDF membrane. After blocking with 5% (w/v) non-fat milk in PBS, the membranes were incubated with rabbit polyclonal anti-MIWI (Cell Signaling Technology, 2079), anti-MVH (1:1000), or anti-eIF4A3 (1:500). Horseradish peroxidase-conjugated anti-rabbit IgG (GE Healthcare, RPN4301) or anti-rabbit light chain (Millipore , MAB201P) secondary antibodies were used and the signals were detected with ECL Plus Western blotting detection reagents (GE Healthcare) and FujiFilm LAS-4000 camera system.

In situ hybridization

Paraffin-embedded testis sections were rehydrated, post-fixed in 4% PFA and permeabilized with 0.2% Triton X-100. The sections were prehybridized in hybridization buffer [50% formamide (Sigma), 5% SSC buffer (Sigma), 250 µg/ml yeast RNA (Ambion), 1× Denhardt's solution (Fluka) in diethylpyrocarbonate (DEPC, Fluka)-treated water]. Hybridization was done in the same buffer containing 100 µM Cy5-polyT 25mer DNA probe (Exiqon) at 37ºC overnight. Sections were washed in 0.1× SSC and once in 1× SSC and mounted in medium with DAPI.

De novo motif predictions followed by enrichment analysis

We used datasets including genes that are abundant in the CB (CB-High+), genes that are found low level in the CB (CB-Low+), genes that are absent in the CB (CB-), or genes that are differentially expressed in somatic cell (SO), mitotic germ cells (MI), meiotic germ cells (ME) or post-meiotic germ cells (PM), or not differentially expressed (Not DE) during spermatogenesis.

Initial data included 15 clusters of genes/transcripts that were all filtered to include only those genes expressed in round spermatids (rSpt+): 1) CB-High+, SO, 2) CB-High+, MI, 3) CB-High+, ME, 4) CB-High+, PM, 5) CB-High+, Not DE, 6) CB-Low+, SO, 7) CB-Low+, MI, 8) CB-Low+, ME, 9) CB-Low+, PM, 10) CB-Low+, Not DE, 11) CB -, SO, 12) CB -, MI, 13) CB -, ME, 14) CB -, PM, 15) CB -, Not DE. For each cluster, 5'UTR and 3'UTR regions of each transcript were extracted resulting in 30 sequence files with either 5' or 3'UTR regions. Short UTR regions (<8 nucleotides) were removed. A negative set of 30 nucleotide sequence files were created (the 30 5' or 3' UTR sequence files were shuffled). DREME (MEME suite of tools) was used to predict *de novo* nucleotides motifs in each of these 30 sequence files. Briefly, DREME finds discriminative regular expression in a positive set of sequences as compared to a negative set (shuffled sequences). Default parameters were used. 30 predicted *de novo* motif files were generated, each of them containing several predicted *de novo* motifs, and are associated with both: one expression pattern (e.g. CB H+, rSpt+, DET-ME) and one type of UTR region (5' or 3'). MAST (MEME suite of tools) was used to identify the presence of the resulting *de novo* motifs in each of the 30 nucleotide sequence files. With the exception of the -remcorr option used to remove highly correlated motifs from query, default parameters were used. AMEN was used to compute and display the motif enrichment analysis (p-value adjusted with FDR <= 0.001) by comparing 5' or 3'UTR regions of : (i) CB-High+, DET-ME versus CB-, DET-ME transcripts; (ii) CB-High+, DET-PM versus CB-, DET-PM transcripts; (iii) CB-High+, Not DE versus CB-, Not DE transcripts.

Primers for RT-qPCR analysis

Target	**Forward**	**Reverse**
Cuff1279	TGTCAATGTCTGAGCATGTTTCC	TCCTGGGGATAAAAGAGACAAAT
Cuff1289	ATGTCCCCTTAGGTCTCAGC	CATGGACACAGCAATGGATGTA
Cuff1617	CCTCTAGGAAACATGCTCAGACA	CTTAGCCTGGGAACACTGCT
Cuff2242	AGGACACAGATGCATCCTAAATGA	CCTTCCCTGGATGTCACTGG
Cuff2476	AACCACCCAGTCTGCTGAAA	GCGTGTAGAATCTCTGCCCA
Cuff3809	CCACCCCAAGCTTGAACATA	TGTTATGAAGCGTGGAGGCA
Cuff3839	CTGTAGTACCCCACAGTTCAGT	AACTTATGGGCCCCTCTCCA
Cuff4273	CTGAGGTCATCTGGGGAGGA	AGCCTTTGACTGGTCTGTGT
RPL13A	ATGGCGGAGGGGCAGGTTCTG	GTACGACCACCACCTTCCGGC
YWHAZ	CGACCACCCATTGTCCCCGC	ACGTCAAACGCTTCTGGCTGC

REFERENCES

1. Kotaja,N., Kimmins,S., Brancorsini,S., Hentsch,D., Vonesch,J.L., Davidson,I., Parvinen,M. and Sassone-Corsi,P. (2004) Preparation, isolation and characterization of stage-specific spermatogenic cells for cellular and molecular analysis. *Nat. Methods*, 1, 249-254.

2. Barchi,M., Geremia,R., Magliozzi,R. and Bianchi,E. (2009) Isolation and analyses of enriched populations of male mouse germ cells by sedimentation velocity: The centrifugal elutriation. *Methods Mol. Biol.*, 558, 299-321.

3. Chalmel,F., Rolland,A.D., Niederhauser-Wiederkehr,C., Chung,S.S., Demougin,P., Gattiker,A., Moore,J., Patard,J.J., Wolgemuth,D.J., Jegou,B., et al. (2007) The conserved transcriptome in human and rodent male gametogenesis. *Proc. Natl. Acad. Sci. U. S. A.*, 104, 8346-8351.

SUPPLEMENTAL FIGURE LEGENDS

FIGURE S1. Distribution of the newly synthesized RNA in male germ cells. **(A)** Overview of different testicular cell types after 12 h culturing of the stage IV-V seminiferous tubules in the presence of ethynyl uridine (EU, green). Spermatocyte (PSpc) nuclei were most intensively stained. In round spermatids (RS), nuclei and cytoplasmic CBs (arrows) were stained. Elongating spermatids (ES) were negative. DAPI was used to stain the nuclei (blue). Phase contrast microscopy confirmed the CB localization of the EU signal. Scale bar is 20 µm. **(B)** Seminiferous tubule cultures were started at the different stages of the seminiferous epithelial cycle to study the stage-specificity of RNA targeting to the CB. After overnight incubation with EU, round spermatids at all differentiation steps showed the CB localization of the newly synthesized RNA. Examples of CBs are indicated with arrows. Scale bar is 10 µm.

FIGURE S2. Localization of the polyA-containing RNAs in the mouse testis. *In situ* hybridization using polyT probe (red) was performed on paraffin embedded testis sections. DAPI was used to stain nuclei (blue). Cross sections of the seminiferous tubules at representative stages of the epithelial cycle (I-II, II-V, VII-VIII, VIII-IX, X-XI, XII-I) are shown. Signal was found in nuclear patches, cytoplasm and in the CBs of round spermatids during the whole course of round spermatid differentiation. PSpc, pachytene spermatocyte; M, meiotic metaphase; RS, round spermatid; ES, elongating spermatid. Some CBs are indicated by arrows. Scale bar 20 µm.

FIGURE S3. Validation of the CB isolation protocol and the CB samples for RNA profiling. **(A)** Immunofluorescence analysis of the samples collected during the CB isolation. MVH-positive CBs are red. DNA is stained with DAPI (blue). Scale bar is 25 µm. **(B)** Immunoblot analysis using anti-MVH antibody further validated the CB isolation that was used for the CB RNA profiling in Figure 1C. Fractions were collected from the different steps during the isolation process. Arrow points to the specific MVH band and IgG is indicated by an asterisk. **(C)** The presence of another main CB component, MIWI, was also confirmed by anti-MIWI immunoblotting. **(D)** Electron microscopy of a CB next to the nucleus (N) of a round spermatid (left) and two examples of CBs bound to paramagnetic beads (B) after the isolation (right). CBs are marked with arrows. **(E)** ~30 nt RNA band co-immunoprecipitates with the CBs. RNA profiles during the CB-isolation procedure. Note that the ~30 nt band becomes depleted from the pellet fraction after CB immunoprecipitation with MVH-bead complexes. LYS: whole testis lysate after sonication, SUP: supernatant fraction after centrifugation, PEL: CB-containing pellet

fraction after centrifugation, PF: pellet fraction after additional filtration. CB: immunoprecipitation of the PF fraction using anti-MVH antibody, Ctrl: immunoprecipitation of the PF using rabbit IgG. IP leftover: leftover samples of immunoprecipitation after removing the antibody-bead complexes. **(F)** Immunoblot analysis using anti-MVH antibody to validate the CB isolation from juvenile mice testes. **(G)** ^{32}P-labeled CB RNA isolated from adult, 22 dpp and 26 dpp testes. The general RNA profile of early (isolated from 22 dpp testes) and late (isolated from 26 dpp testes) CBs, including the characteristic 30 nt band, did not change significantly during the development. The 5'-CAP molecules were not labeled in this experiment.

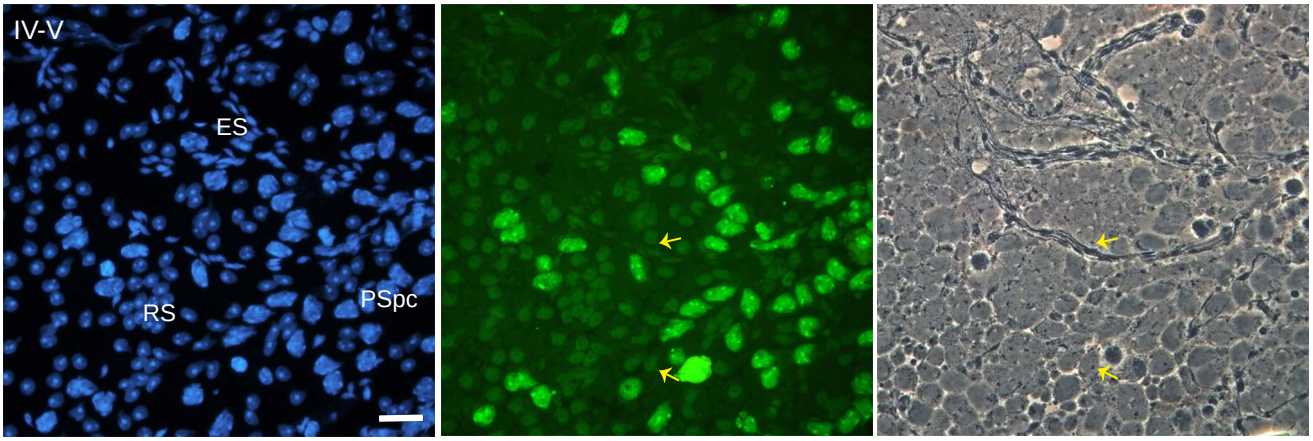
FIGURE S4. Validation of the presence of EJC in the CB. **(A)** Immunoblot analysis of the CB extracts using anti-eIF4A3 antibody. The CB isolation was validated by anti-MVH antibody. The specific immunoreactive protein bands are indicated with arrows and IgG chains are indicated with asterisks. LYS: whole testis lysate after sonication, SUP: supernatant fraction after centrifugation, PEL: CB-containing pellet fraction after centrifugation, PF: pellet fraction after additional filtration, CB: immonoprecipitates from PF fraction using anti-MVH antibody, Ctrl: immunoprecipitates from PF fraction using anti-acrosin (negative control) antibody. **(B)** Immunofluorescence analysis of the ECJ proteins eIF4A3 and RBM8A (red) on paraffin embedded testis sections. CBs were co-labeled with anti-DDX25 antibody (green). Some CBs are indicated with white arrows. DNA was stained with DAPI (blue). Scale bar is 10 µm. Confocal fluorescent microscopy of the co-immunostaining with anti-DDX25 and anti-MVH antibodies validated DDX25 as a CB marker. **(C)** Localization of eIFA3 in the mouse testis. Paraffin embedded testis sections were immunostained with anti-eIF4A3 antibody (green) and counterstained with DAPI (blue) to visualize nuclei. The cross-sections of seminiferous tubules at representative stages (I, VII-VIII, IX-X) are shown. eIF4A3 was expressed in Sertoli cells and in germ cells until step 11 spermatids at stages XI. eIF4A3 localized mostly in the nuclei, but was found in the cytoplasmic granules in late pachytene and diplotene spermatocytes and in CBs in round spermatids. eIF4A3 localized in the CB throughout the whole round spermatid differentiation (steps 1-8) but not in the late CBs in step 9-10 spermatids. PSpc, pachytene spermatocyte; RS, round spermatid; ES, elongating spermatid. Arrows point to CBs. Scale bar is 20 µm

FIGURE S5. Heat map summarizing the expression profiling of the CB-localized mRNAs. Each line represents a gene that was found in the CB at low level (CB-Low+) or at high level (CB-High+) or not found in the CB (CB-) in RNA sequencing. Each column corresponds to a

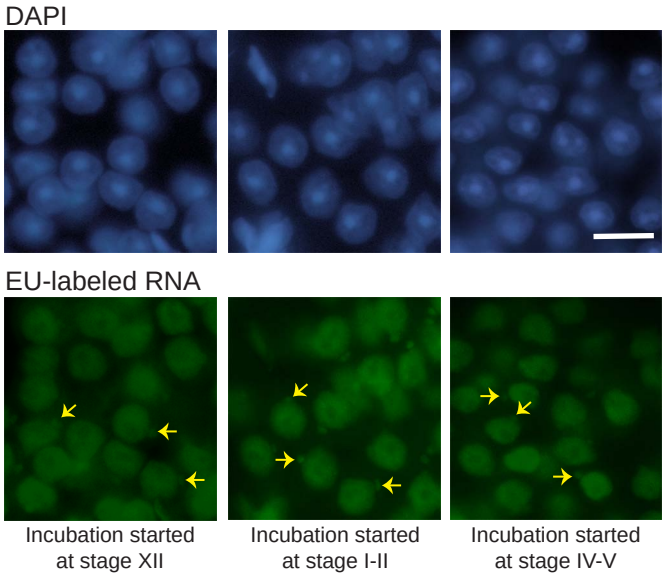
testicular sample [including Sertoli cells (SC), spermatogonia (Spg), pachytene spermatocyte (pSpc), round spermatids (rSpt), seminiferous tubules (Tubules) or total testis] that were hybridized to Affymetrix GeneChip microarrays. Genes were grouped according to their differential expression in the testis: somatic (DET-SO), mitotic (DET-MI), meiotic (DET-ME) or post-meiotic (DET-PM) transcripts differentially expressed in the testis (3). Not-DET represents the group of genes that are detectable in round spermatids but not differentially expressed in testicular cells. rSpt- corresponds to the genes not identified in round spermatids. Low expression signals are displayed in blue and high expression signals in red.

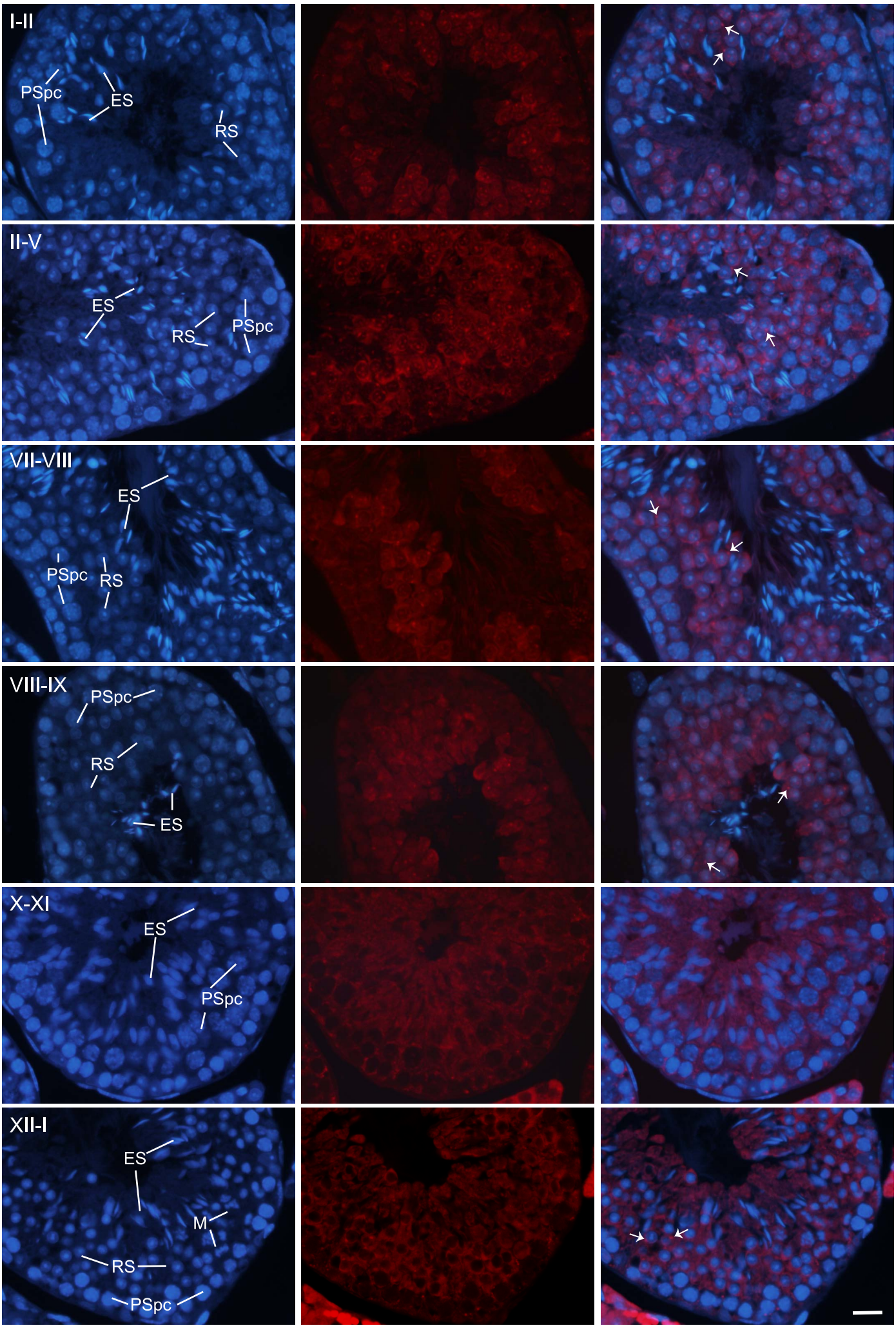
FIGURE S6. *De novo* motif predictions and motif enrichment analysis. The table shows the motifs that were predicted by *de novo* motif prediction and were found to be enriched (indicated in red and bold font) in the specific clusters of transcripts. CB-High+, transcripts highly abundant in the CB; CB-, non-CB-localized transcripts; ME, transcripts induced in meiotic cells; PM, transcripts induced in post-meiotic cells; Not-DET, transcripts that are not differentially expressed in testis but detectable in round spermatids.

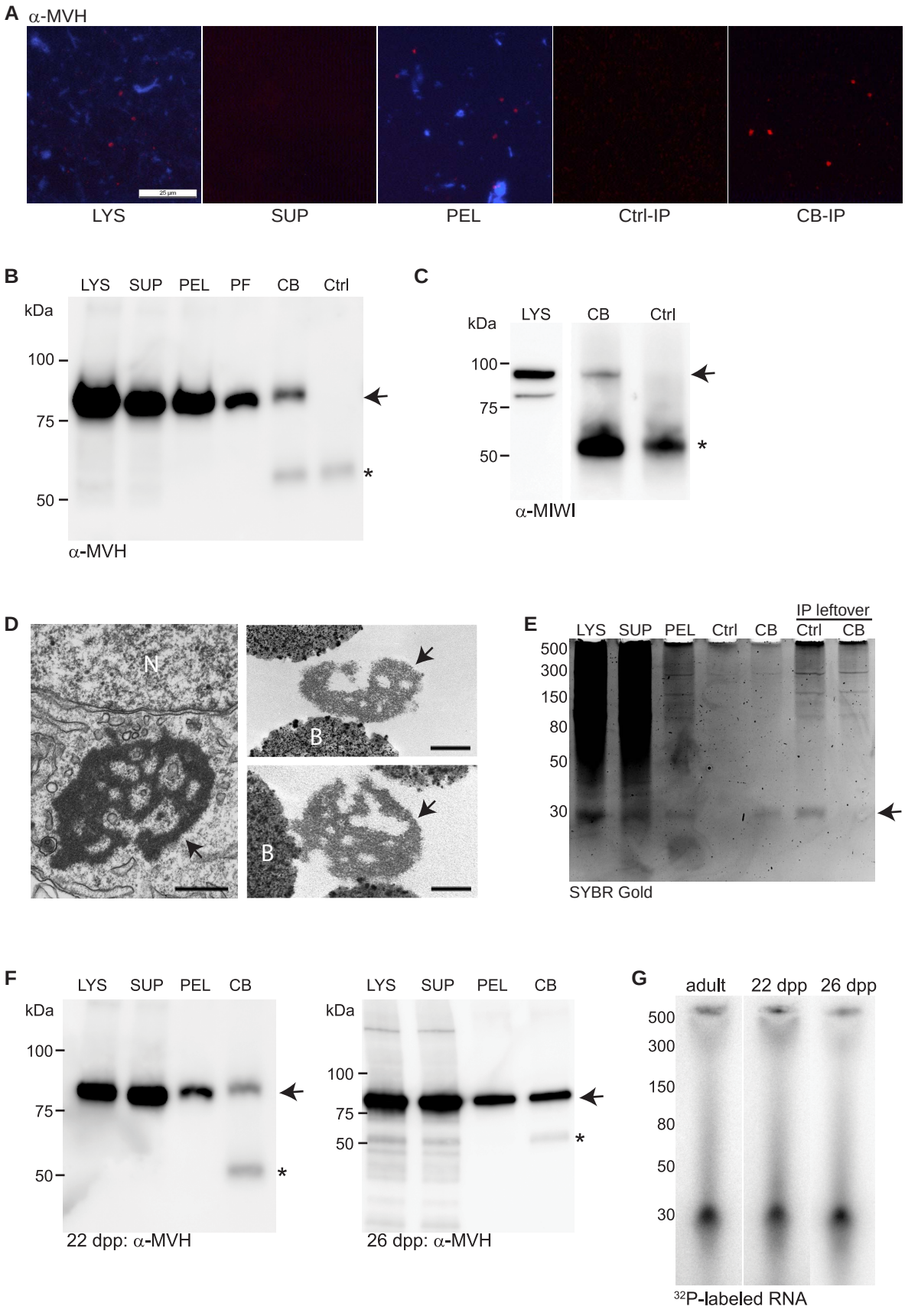
A

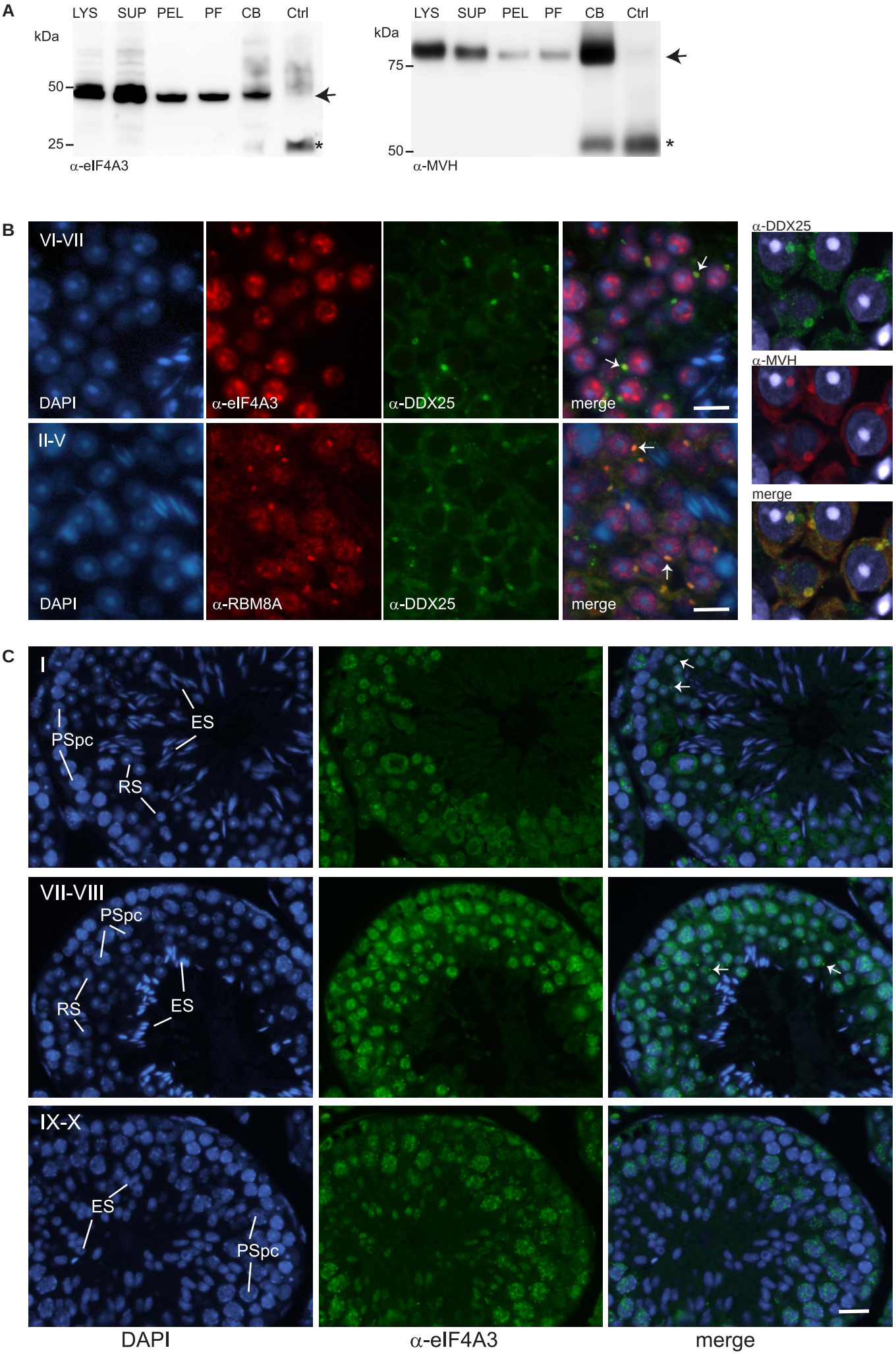

IV-V
ES
RS
PSpc
DAPI
EU-labeled RNA
phase contrast

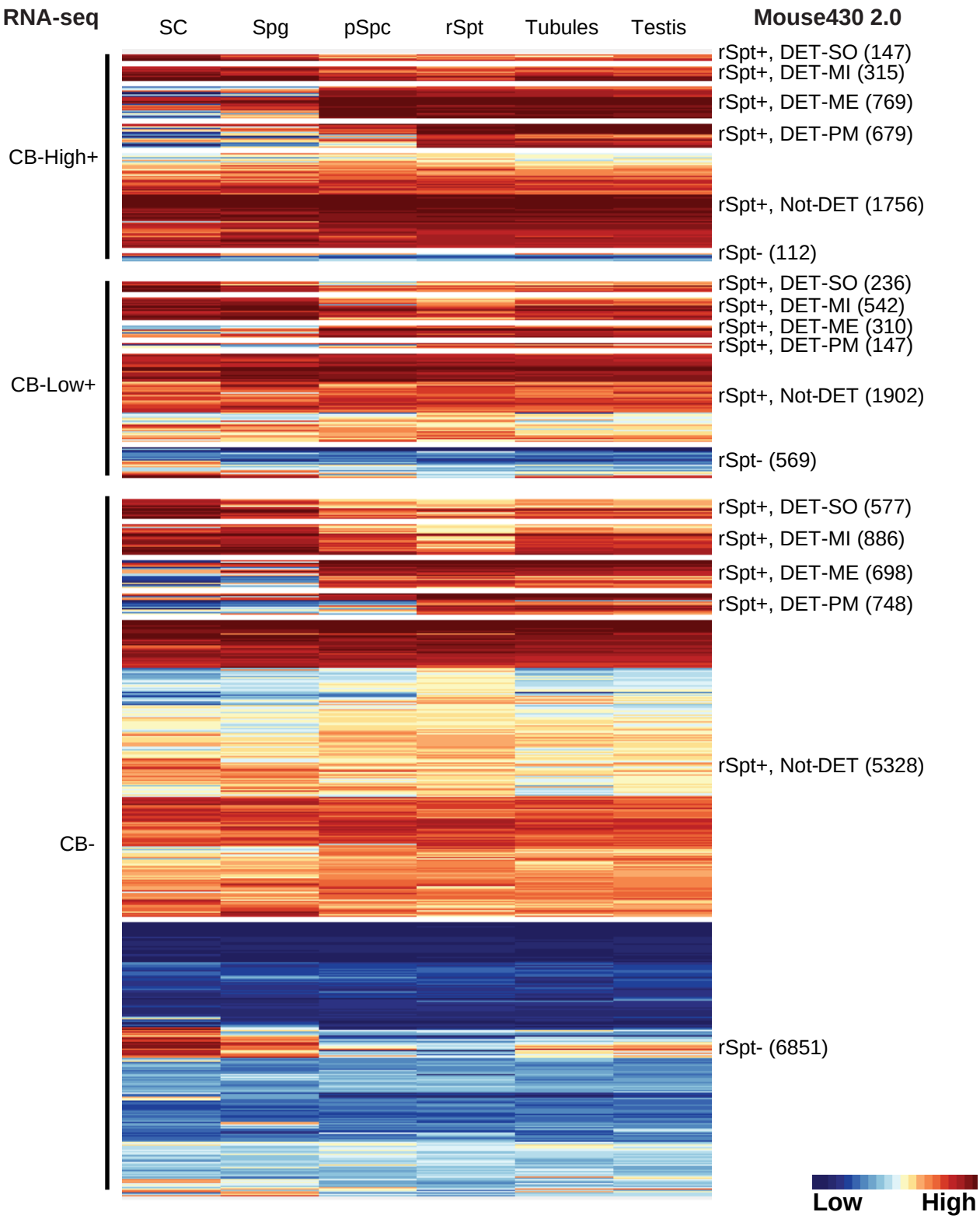
B


DAPI
EU-labeled RNA
Incubation started at stage XII
Incubation started at stage I-II
Incubation started at stage IV-V


I-II
PSpc
ES
RS
II-V
ES
RS
PSpc
VII-VIII
ES
PSpc
RS
VIII-IX
PSpc
RS
ES
X-XI
ES
PSpc
XII-I
ES
M
RS
PSpc
DAPI
polyT
merge


A
α-MVH
LYS
SUP
PEL
Ctrl-IP
CB-IP
B
LYS SUP PEL PF CB Ctrl
kDa
100
75
50
*
α-MVH
C
LYS CB Ctrl
kDa
100
75
50
*
α-MIWI
D
N
B
B
E
IP leftover
LYS SUP PEL Ctrl CB Ctrl CB
500
300
150
80
50
30
SYBR Gold
F
LYS SUP PEL CB
kDa
100
75
50
*
22 dpp: α-MVH
LYS SUP PEL CB
kDa
100
75
50
*
26 dpp: α-MVH
G
adult 22 dpp 26 dpp
500
300
150
80
50
30
32P-labeled RNA


A
LYS SUP PEL PF CB Ctrl
kDa
50
25
α-eIF4A3
LYS SUP PEL PF CB Ctrl
kDa
75
50
α-MVH
B
VI-VII
DAPI
α-eIF4A3
α-DDX25
merge
II-V
DAPI
α-RBM8A
α-DDX25
merge
α-DDX25
α-MVH
merge
C
I
PSpc
ES
RS
VII-VIII
PSpc
RS
ES
IX-X
ES
PSpc
DAPI
α-eIF4A3
merge


RNA-seq
SC
Spg
pSpc
rSpt
Tubules
Testis
Mouse430 2.0
CB-High+
CB-Low+
CB-
rSpt+, DET-SO (147)
rSpt+, DET-MI (315)
rSpt+, DET-ME (769)
rSpt+, DET-PM (679)
rSpt+, Not-DET (1756)
rSpt- (112)
rSpt+, DET-SO (236)
rSpt+, DET-MI (542)
rSpt+, DET-ME (310)
rSpt+, DET-PM (147)
rSpt+, Not-DET (1902)
rSpt- (569)
rSpt+, DET-SO (577)
rSpt+, DET-MI (886)
rSpt+, DET-ME (698)
rSpt+, DET-PM (748)
rSpt+, Not-DET (5328)
rSpt- (6851)
Low
High

		rSpt+					
		CB-High+			CB-		
		ME	PM	Not-DET	ME	PM	Not-DET
5'UTR motifs (2 terms)	3703	286	304	333	212	220	645
GYTGGACA	25	1 / 2	3 / 2	1 / 2	1 / 1	**15 / 2**	1 / 4
GGTGCWG	43	3 / 3	3 / 4	1 / 4	4 / 3	**14 / 3**	5 / 8
3'UTR motifs (43 terms)	2886	226	207	335	145	127	568
TCTAGAAC	19	1 / 2	0 / 1	1 / 2	**14 / 1**	**15 / 1**	0 / 4
CTTBAAGA	54	2 / 4	4 / 4	3 / 6	**19 / 3**	**17 / 2**	10 / 11
RACAGACA	36	1 / 3	2 / 3	1 / 4	**15 / 2**	**14 / 2**	2 / 7
CTCAAGTG	48	2 / 4	1 / 3	2 / 6	**17 / 2**	**14 / 2**	8 / 9
TTTCWGAA	36	2 / 3	1 / 3	1 / 4	**15 / 2**	**14 / 2**	3 / 7
AGCACAVA	54	1 / 4	2 / 4	5 / 6	**18 / 3**	**15 / 2**	1 / 11
CAAGGRAA	31	3 / 2	1 / 2	2 / 4	**14 / 2**	**14 / 1**	3 / 6
RAACCATT	31	0 / 2	3 / 2	2 / 4	**14 / 2**	**14 / 1**	2 / 6
CCATTGCT	44	1 / 3	0 / 3	5 / 5	**16 / 2**	**15 / 2**	8 / 9
AAATGTS	40	2 / 3	4 / 3	5 / 5	**15 / 2**	**14 / 2**	1 / 8
CAAAGVC	62	5 / 5	2 / 4	4 / 7	**18 / 3**	**19 / 3**	8 / 12
CCTCACWT	62	5 / 5	3 / 4	6 / 7	**18 / 3**	**15 / 3**	5 / 12
CAGGTKGA	64	1 / 5	4 / 5	8 / 7	**18 / 3**	**16 / 3**	2 / 13
GWCCAGAC	52	1 / 4	1 / 4	5 / 6	**16 / 3**	**17 / 2**	5 / 10
CCTGVWG	62	2 / 5	1 / 4	6 / 7	**17 / 3**	**16 / 3**	8 / 12
CCCASCC	170	13 / 13	11 / 12	17 / 20	**28 / 9**	**19 / 8**	20 / 34
TTTCCTWG	102	7 / 8	4 / 7	8 / 12	**21 / 5**	**17 / 5**	14 / 20
CTBCCYT	77	6 / 6	2 / 6	8 / 9	**18 / 4**	**15 / 3**	10 / 15
AGGCCAS	86	8 / 7	4 / 6	11 / 10	**19 / 4**	**15 / 4**	12 / 17
GGRAGTCT	72	7 / 6	1 / 5	9 / 8	**17 / 4**	**16 / 3**	8 / 14
GCCATCCW	90	5 / 7	4 / 7	5 / 10	**19 / 5**	**17 / 4**	12 / 18
ARAWGGAA	59	4 / 5	3 / 4	7 / 7	**15 / 3**	**15 / 3**	7 / 12
ACTTTBTA	52	3 / 4	3 / 4	7 / 6	**14 / 3**	**15 / 2**	7 / 10
ARAATGTT	79	2 / 6	6 / 6	5 / 9	**17 / 4**	**18 / 4**	5 / 16
TYTYTCCT	88	2 / 7	3 / 6	10 / 10	**18 / 4**	**18 / 4**	10 / 17
CMTTCTCC	66	4 / 5	1 / 5	11 / 8	**15 / 3**	**16 / 3**	7 / 13
AGCATCCT	159	11 / 13	11 / 11	16 / 19	**24 / 8**	**20 / 7**	20 / 31
AGCAGWTG	172	8 / 14	12 / 12	17 / 20	**25 / 9**	**22 / 8**	24 / 34
AGGAGCCC	77	2 / 6	4 / 6	8 / 9	**16 / 4**	**15 / 3**	13 / 15
TKGCCAAG	139	5 / 11	6 / 10	11 / 16	**22 / 7**	**22 / 6**	21 / 27
CTGACCTB	115	6 / 9	2 / 8	13 / 13	**19 / 6**	**20 / 5**	19 / 23
MCATCTGT	107	8 / 8	2 / 8	17 / 12	**18 / 5**	**16 / 5**	18 / 21
WGCTGCT	154	16 / 12	6 / 11	16 / 18	**22 / 8**	**22 / 7**	23 / 30
RGGACCA	144	4 / 11	10 / 10	14 / 17	**21 / 7**	**20 / 6**	19 / 28
CMCCTCC	111	7 / 9	8 / 8	8 / 13	**18 / 6**	**17 / 5**	12 / 22
ARGYCACT	88	7 / 7	6 / 6	8 / 10	**15 / 4**	**19 / 4**	15 / 17
TKTGCACT	91	5 / 7	3 / 7	12 / 11	**15 / 5**	**21 / 4**	8 / 18
TGGAAAWG	93	4 / 7	3 / 7	12 / 11	**15 / 5**	**17 / 4**	16 / 18
CATGKGG	143	9 / 11	7 / 10	21 / 17	**19 / 7**	**21 / 6**	19 / 28
TTTDAAAA	108	2 / 9	10 / 8	16 / 13	**16 / 5**	**21 / 5**	18 / 21
ARCCCATG	141	8 / 11	6 / 10	14 / 16	17 / 7	**22 / 6**	19 / 28
KGAGGCC	204	13 / 16	11 / 15	24 / 24	22 / 10	**22 / 9**	33 / 40
CASTGKG	179	14 / 14	7 / 13	24 / 21	20 / 9	**20 / 8**	27 / 35

Depleted **p-value** **Enriched**

1 **0.95** **0.05** **1e-5** **0**