

Supplementary Information

Bombyx Vasa sequesters transposon mRNAs in nuage via phase separation requiring RNA binding and self-association

Yamazaki et al.

Table of contents

Supplementary Figures

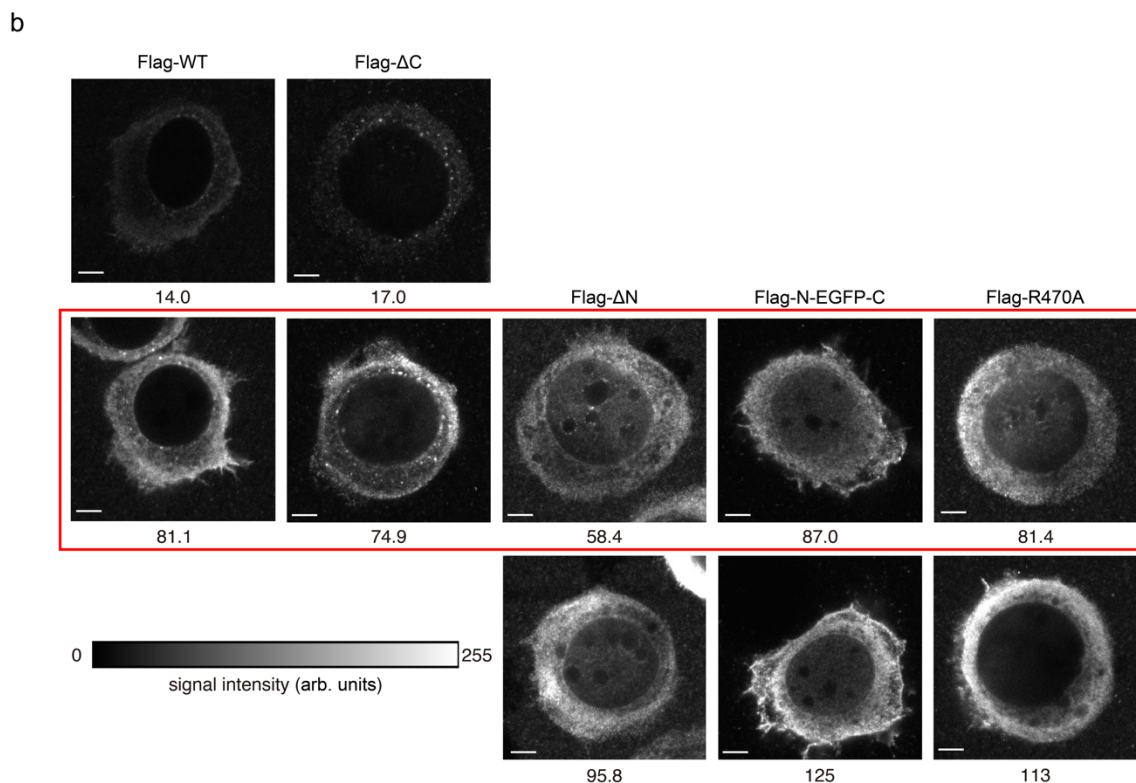
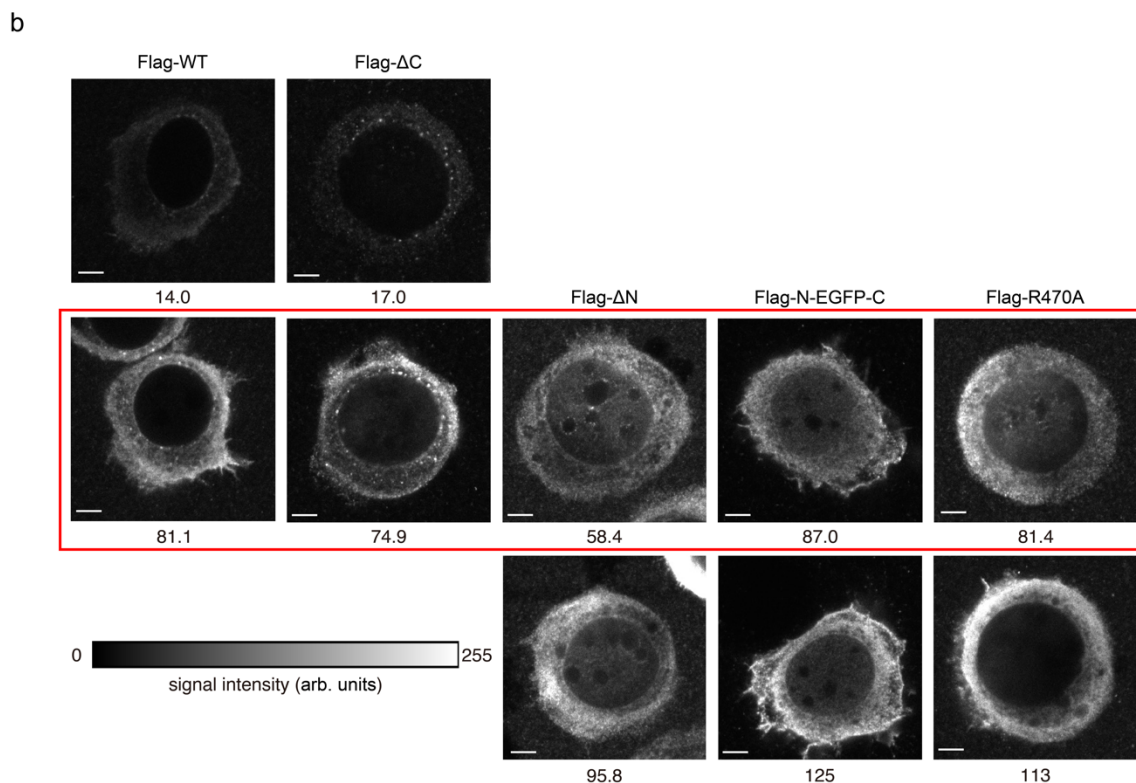
Supplementary Figure.1 Unique trait of each domain in BmVasa.

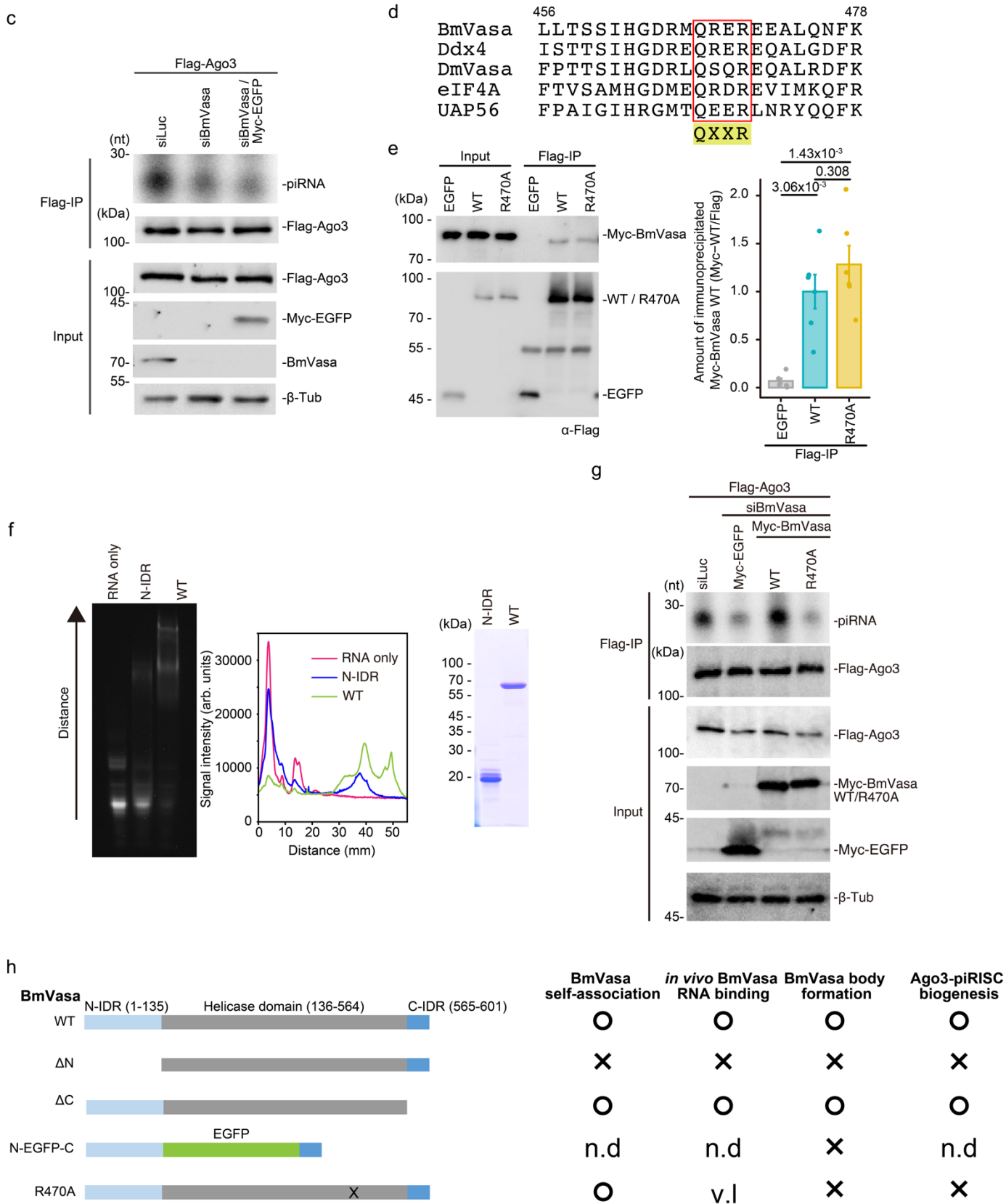
Supplementary Figure 2 LLPS of recombinant BmVasa proteins and behavior of BmVasa RK mutant in BmN4 cells.

Supplementary Figure 3 The droplets of BmVasa WT and R470A, and FRAP analysis of the droplets of BmVasa WT and E339Q.

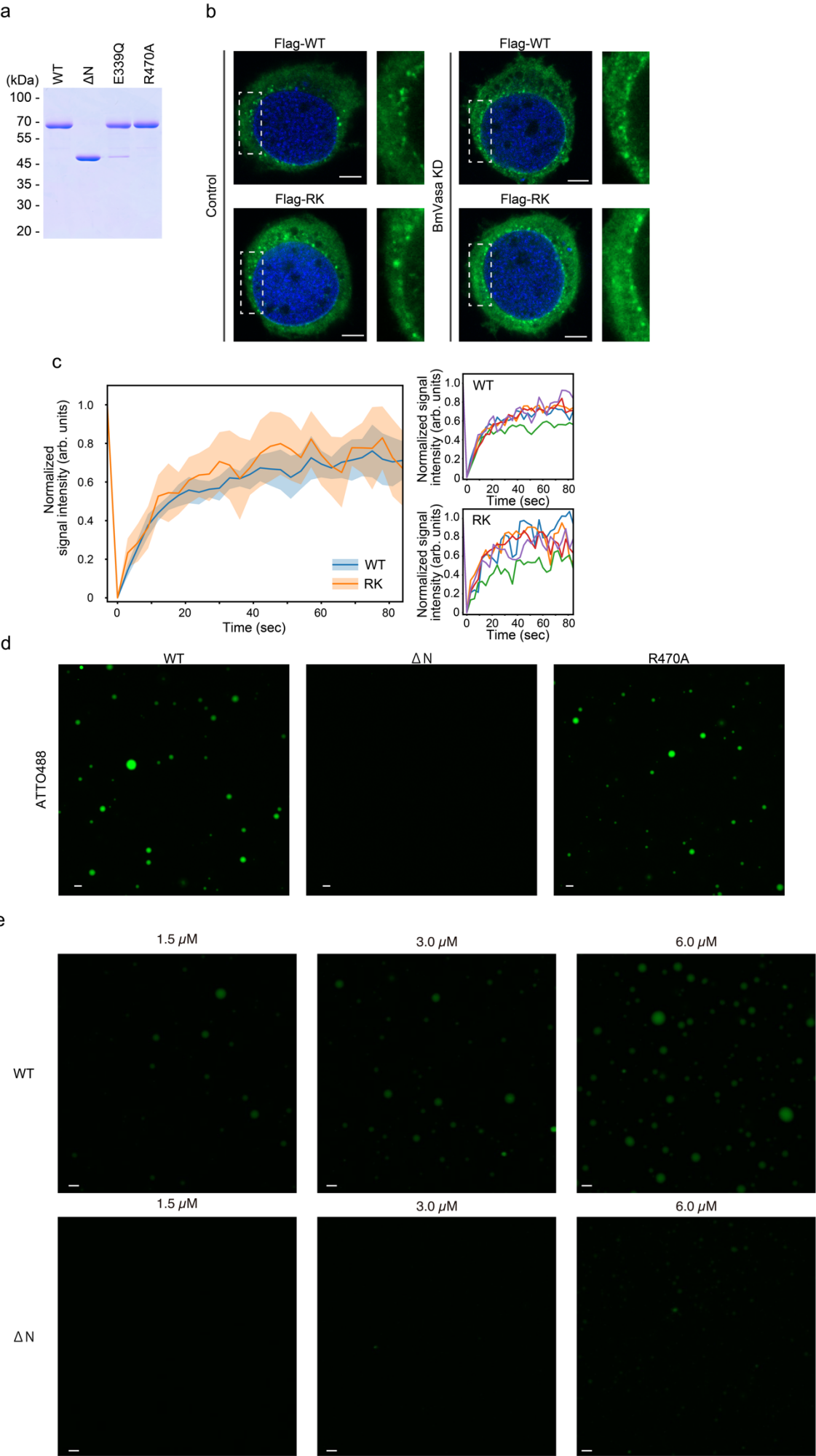
Supplementary Figure 4 N-IDR sequence alignment.

Supplementary Figure 5 Bioinformatic analysis of BmVasa FAST-iCLIP reads.



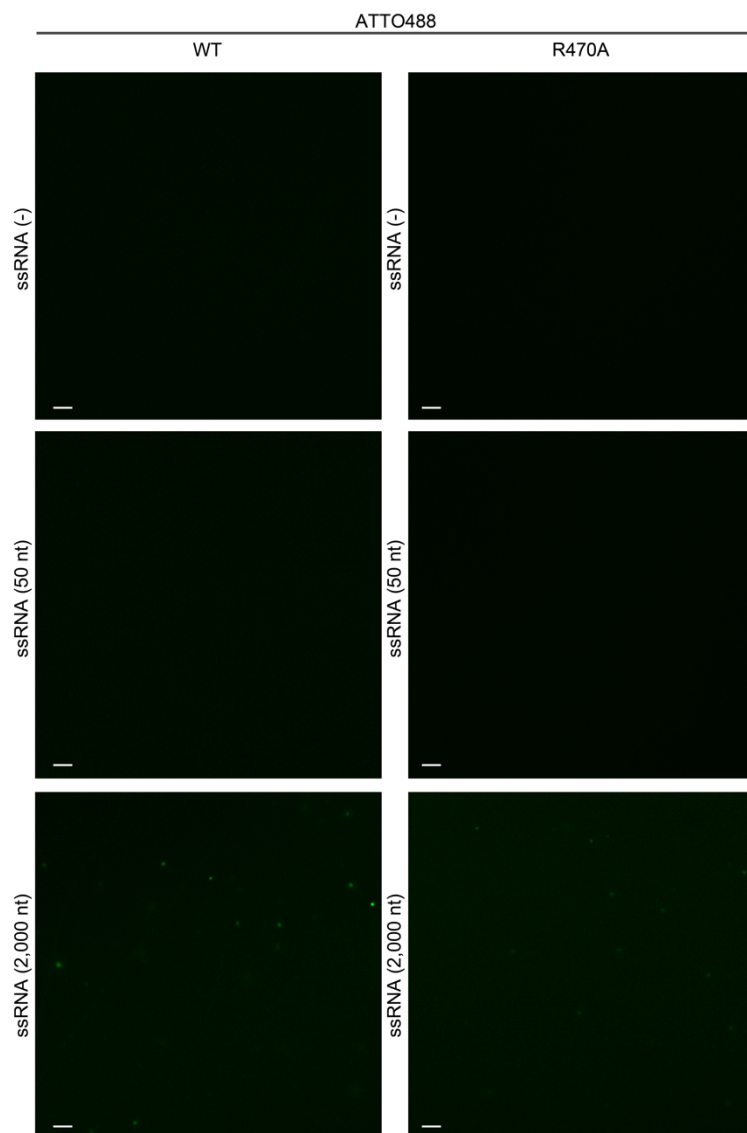


Supplementary Figure 1. Unique trait of each domain in BmVasa. **(a)** Disorder probability shows the N-IDR (residues 1–135) and the C-IDR (residues 565–601) in BmVasa. Red line on the disorder plots represents disorder probability = 0.5. The domain structure and amino acid sequence of BmVasa are also indicated. Light blue: N-IDR. Dark blue: C-IDR. Yellow: DEAD box. Red: Arg470. The RNA helicase domain: residues 136–564. **(b)** Subcellular localization of Flag-BmVasa WT (Flag-WT), Flag-BmVasa Δ C mutant (Flag- Δ C), Flag-BmVasa Δ N mutant (Flag- Δ N), Flag-BmVasa N-EGFP-C mutant (Flag-N-EGFP-C), and Flag-BmVasa R470A mutant (Flag-R470A) in BmN4 cells (gray) at various expression levels. Grayscale bar indicates signal intensity. Numerical value below the image is the mean signal intensity (a.u) in cytoplasm. Scale bar: 5 μ m. **(c)** piRNAs loaded onto Flag-Ago3 in BmVasa-depleted BmN4 cells (BmVasa KD). siLuc (Control) and Myc-EGFP were used as negative controls. Flag-IP, upper: 32 P-labeled piRNAs bound to Flag-Ago3. Flag-IP, lower: western blotting shows Flag-Ago3 in the immunoprecipitates. Input, from the top: Western blotting shows the levels of Flag-Ago3, Myc-EGFP, endogenous BmVasa, and β -Tubulin (β -Tub; a loading control) in input. **(d)** Conservation of Q-X-X-R and its surroundings in BmVasa (residues 456–478), Ddx4, DmVasa, and eIF4A of *Drosophila melanogaster* (UniProt: Q02748), and UAP56 of *Drosophila melanogaster* (UniProt: Q27268). The red box shows Q-X-X-R in each protein. **(e)** Self-association of WT BmVasa and the R470A mutant. Anti-Flag antibody was used for immunoprecipitation and western blotting (left). Anti-Myc antibody was used for western blotting (upper gel). Flag-EGFP was used as a negative control (lower gel). Myc-BmVasa signal was divided by Flag-tagged protein signal and plotted (right) (Error bar: standard error). Two-tailed Welch's t-test was performed and p values are indicated on the graph (n=6). **(f)** Electrophoretic mobility shift assay was performed with BmVasa WT, N-IDR and 50 nt ssRNA. Protein/RNA mixture was subjected to electrophoresis with non-denaturing polyacrylamide gel. The gel was stained with SYBR Gold to visualize RNA and is shown on the left. Quantified SYBR Gold signal along with the arrow was shown in the middle. BmVasa WT and N-IDR CBB-stained (SDS-PAGE) are shown on the right. **(g)** piRNAs loaded onto Flag-Ago3 in BmVasa-depleted BmN4 cells (siBmVasa). Prior to immunoprecipitation, Myc-BmVasa WT and Myc-BmVasa R470A were individually expressed. Myc-EGFP was used as a negative control. Flag-IP, upper: 32 P-labeled piRNAs bound to Flag-Ago3. Flag-IP, lower: western blotting shows Flag-Ago3 in the immunoprecipitates. Input, from the top: Western blotting shows Flag-Ago3, Myc-BmVasa WT, Myc-BmVasa R470A, Myc-EGFP, and β -Tubulin (β -Tub; a loading control) in input. **(h)** Summary of functional characteristics of each BmVasa mutant analyzed in this study and WT BmVasa. The domain structures of the proteins are also shown. v.l: very low. n.d: not determined. Source data are provided as a Source Data file.

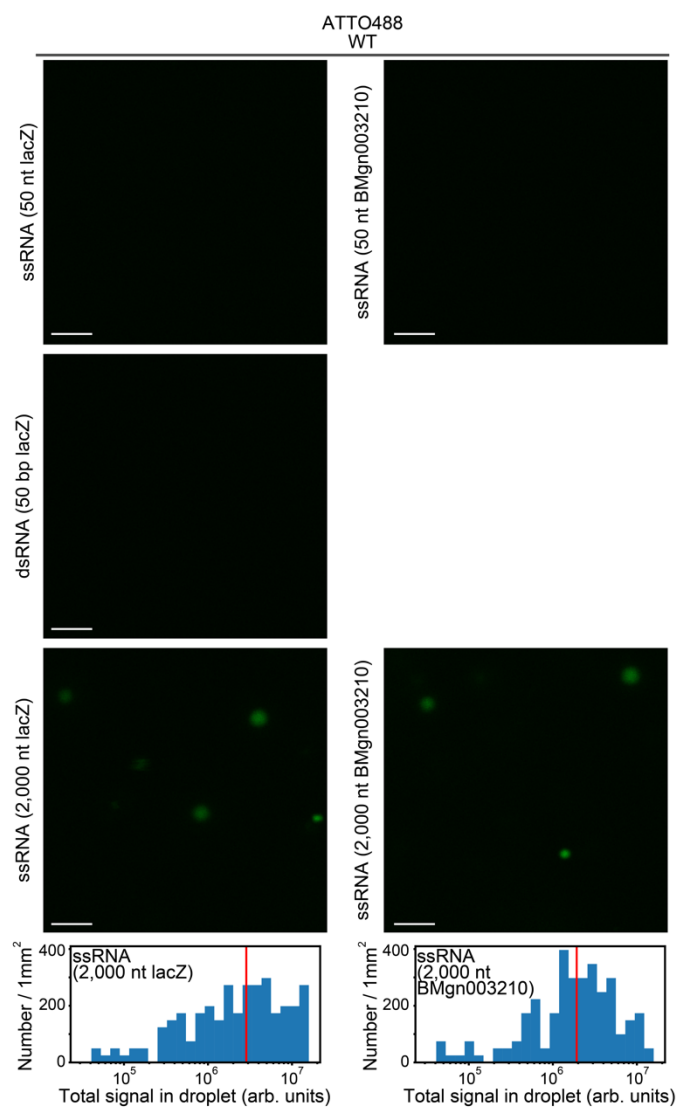


Supplementary Figure 2. LLPS of recombinant BmVasa proteins and behavior of BmVasa RK mutant in BmN4 cells. **(a)** CBB-stained protein gel shows the purity of BmVasa WT and its mutants, Δ N, E339Q, and R470A, employed in the *in vitro* droplet formation assay. **(b)** Subcellular localization of Flag-BmVasa WT and RK mutant (Flag-RK) (green) in control BmN4 cells (Control) and cells lacking endogenous BmVasa (BmVasa KD). DAPI (blue): nuclei. Scale bar: 5 μ m. Enlarged images of the insets indicated with white dotted lines are shown on the right. **(c)** FRAP analysis of the EGFP-BmVasa WT and RK expressed in BmN4 and localized in Vasa bodies was performed. Mean (line) and standard deviation (shade) are shown (left) (n=5). Individual data are plotted on the right. **(d)** Uncropped images of Fig. 2g. Scale bar: 5 μ m. **(e)** *In vitro* droplet formation assay in the presence of 15% PEG 6000 was performed with different concentrations of BmVasa WT and Δ N. Scale bar: 5 μ m. Source data are provided as a Source Data file.

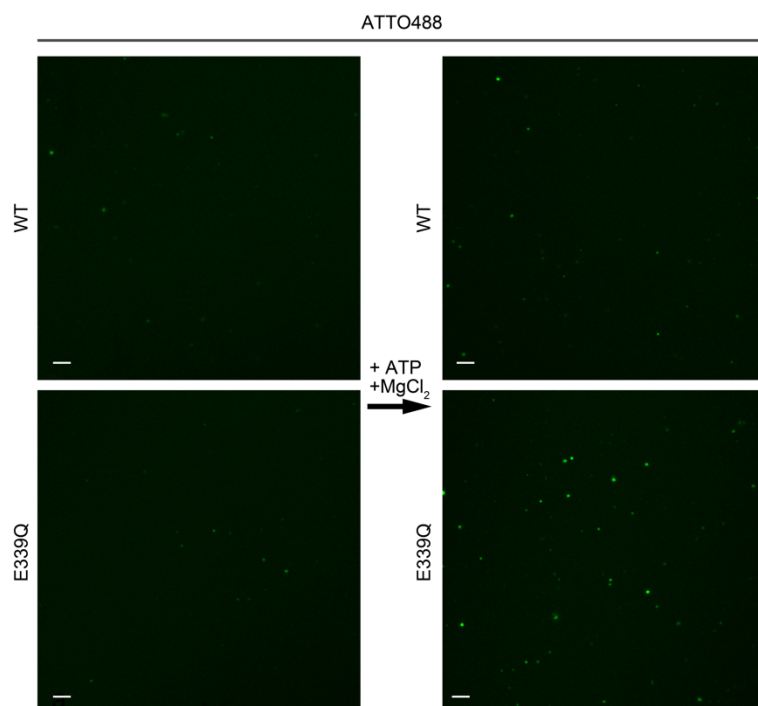
a



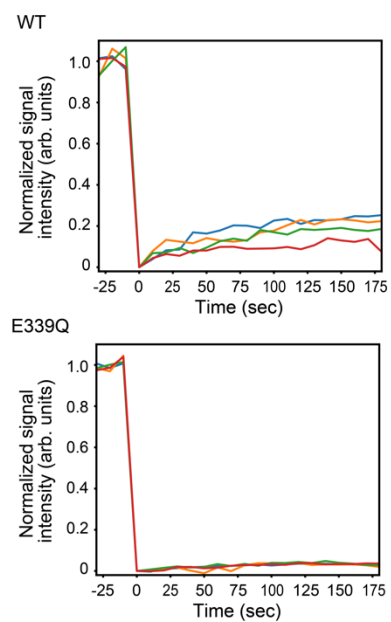
b



c



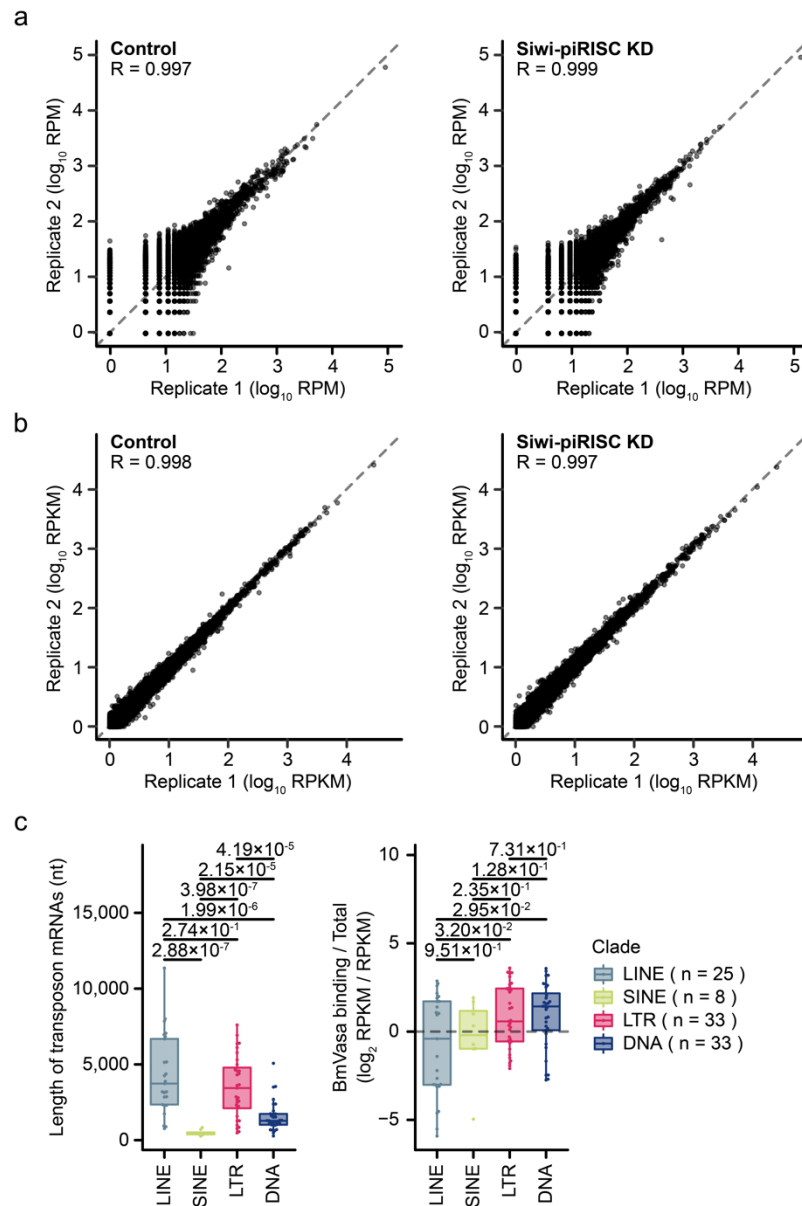
d



Supplementary Figure 3. The droplets of BmVasa WT and R470A, and FRAP analysis of the droplets of BmVasa WT and E339Q. (a) Uncropped images of Fig. 3a. Scale bar: 5 μm . (b) *In vitro* droplet formation assay of ATTO488-labeled recombinant BmVasa WT (1.5 μM) in the presence of 5% PEG 6000 and 50 nt ssRNA (derived from lacZ or BMgn003210), 50 bp dsRNA (derived from lacZ) or 2,000 nt ssRNA (derived from lacZ or BMgn003210). Only the 2,000 nt ssRNA promoted droplet formation. Histograms of total ATTO488 signal intensity of a droplet in the presence of 2,000 nt ssRNAs are shown. Red line indicates the median; the difference between 2,000 nt ssRNA derived from lacZ and BMgn003210 was small for clarity. ($p = 0.0406$ by two-tailed Mann–Whitney U test). Scale bar: 5 μm . (c) Uncropped images of Fig. 3b. Scale bar: 5 μm . (d) FRAP analysis of the droplets of BmVasa WT and E339Q in the presence of 2,000-nt-long ssRNAs, ATP, and MgCl_2 . Individual data ($n = 4$) of Fig. 3c are shown. Source data are provided as a Source Data file.

BmVasa	1	MDDDWDD	SCEAVVPPP	PLQNHDS	VDEGHSL	SRG-RGF	PSFNED	DEKENG
DmVasa	1	MSDDWDDE	-----	PI-----	VD-----	TRGAR	GGDWSDD	EDTAKSF
BmVasa	50	YGERRGRG	-----	-----	-----	GR-GGGR	GRGGRG	SGSREQHS
DmVasa	32	SGEAE	GDGVGG	SGGEGGY	QGGNRDVF	GRI	GGGRGG	GAGGYRGGNR-----
BmVasa	80	DYETNGD	HDDRGRGER	-----	-----	GRGRGR	GGGRGGGG	-----
DmVasa	78	--DGGGF	HGGRRER	GERDFR	GGEGGFRG	QGGSRGG	QGGSRGG	QGGFRGGE
BmVasa	114	-----	-----	-----	-----	GGDRND	YEDNEI	--GENGE-----
DmVasa	126	GGFRGR	LYENED	GDERRGR	LDREER	GGERRGR	LDREER	GGGERGERGDGGF
BmVasa	131	-----	-----	-----	-----	TKKPVT	YVPPEPT	NDETEIFSS
DmVasa	176	ARRRRN	EDDINN	NNNNIV	EDVERK	REFYIP	PEPSNDA	IEIFSSGIASGIHF

Supplementary Figure 4. N-IDR sequence alignment. The sequences of BmVasa and DmVasa were subjected to pairwise sequence alignment. Amino acids 1–160 of BmVasa and 1–225 of DmVasa are shown. Blue and orange highlights indicate N-IDR and conserved amino acid, respectively.



Supplementary Figure 5. Bioinformatic analysis of BmVasa FAST-iCLIP reads. **(a)** Scatter plots show RPM values of the BmVasa FAST-iCLIP library (BmVasa binding) for mRNAs in two replicates produced in control BmN4 cells (left) and cells lacking Spn-E (Siwi-piRISC KD; right). R is Pearson's correlation coefficient. **(b)** Scatter plots show the expression levels of mRNAs in two replicates produced in control BmN4 cells (left) and cells lacking Siwi (Siwi-piRISC KD; right). R is Pearson's correlation coefficient. **(c)** Box plots show the length distribution (left) and the enrichment value (BmVasa binding relative to Total; right) of transposon mRNAs with the annotation of clade. Two-tailed Mann–Whitney U test was performed and p values are indicated in the box plots. Center line: median. Box limits: upper and lower quartiles. Whiskers: 1.5 × interquartile range. Source data are provided as a Source Data file.

Supplementary Table 1. Sequences of oligonucleotides

Experiment	Name	Sequence*
Plasmid Construction	3xFLAG-BmVasa ΔN Forward (vector)	ACGCGTACCGGTCATCATCACCATCAC
	3xFLAG-BmVasa ΔN Reverse (vector)	GAATTCGATATCCTTGTATCGTCATC
	3xFLAG-BmVasa ΔN Forward (insert)	AAGGATATCGAATTCACCTATGTGCCTCCAGAACCA
	3xFLAG-BmVasa ΔN Reverse (insert)	ATGACCGGTACGCGTCTACCATCTTCTTCAGGTTT
	3xFLAG-BmVasa ΔC Forward (vector)	ACGCGTACCGGTCATCATCACCATCAC
	3xFLAG-BmVasa ΔC Reverse (vector)	GAATTCGATATCCTTGTATCGTCATC
	3xFLAG-BmVasa ΔC Forward (insert)	AAGGATATCGAATTCATGGATGATGACTGGGATGAT
	3xFLAG-BmVasa ΔC Reverse (insert)	ATGACCGGTACGCGTCTACTTTAAGAAAGTCTGGGAC
	3xFLAG-BmVasa N-EGFP-C Forward (vector)	GGAGGTGGCACTGCTACATTTCAAG
	3xFLAG-BmVasa N-EGFP-C Reverse (vector)	TACCGGCTTCTTCTGTTTCCCGTT
	3xFLAG-BmVasa N-EGFP-C Forward (insert)	ACGAAGAAGCCGGTAATGGTGAGCAAGGGCGAGGAG
	3xFLAG-BmVasa N-EGFP-C Reverse (insert)	AGCAGTGCCACCTCCCTGTACAGCTCGTCCATGCC
	3xFLAG-BmVasa RK Forward (vector)	ACCTATGTGCCTCCAGAACCAAC
	3xFLAG-BmVasa RK Reverse (vector)	CATCCATGAATTCGATATCCTTGTCTATCG
	3xFLAG-BmVasa RK Forward (insert)	TCGAATTCATGGATGATGACTGGG
	3xFLAG-BmVasa RK Reverse (insert)	TGGAGGCACATAGGTTACCGG
	3xFLAG-BmVasa N(Ddx4)-EGFP-C Forward (vector)	ATGGTGAGCAAGGGCGAGG
	3xFLAG-BmVasa N(Ddx4)-EGFP-C Reverse (vector)	GAATTCGATATCCTTGTATCGTCATCCTTG
	3xFLAG-BmVasa N(Ddx4)-EGFP-C Forward (insert)	AGGATATCGAATTCATGGGAGATGAGGATTGGGAAGC
	3xFLAG-BmVasa N(Ddx4)-EGFP-C Reverse (insert)	GCCCTTGCTCACCATACTTTCGCCGCTTCAGC
	3xFLAG-BmVasa N(DmVasa)-EGFP-C Forward (vector)	ATGGTGAGCAAGGGCGAGG
	3xFLAG-BmVasa N(DmVasa)-EGFP-C Reverse (vector)	GAATTCGATATCCTTGTATCGTCATCC
	3xFLAG-BmVasa N(DmVasa)-EGFP-C Forward (insert)	AAGGATATCGAATTCATGTCTGACGACTGGGATGATGAG
	3xFLAG-BmVasa N(DmVasa)-EGFP-C Reverse (insert)	GCCCTTGCTCACCATCTCTGCTTCTTTCGACATCTTCG
	3xFLAG-BmVasa N(Ddx4 FA)-EGFP-C Forward (vector)	GAAGGCGGCGAAAGTATGGTGAGCAAGGGCGAGG
	3xFLAG-BmVasa N(Ddx4 FA)-EGFP-C Reverse (vector)	ATCCTCATCTCCCATGAATTCGATATCCTTGTATCGTCATCC
	3xFLAG-EGFP-BmVasa WT Forward (vector)	GGTTCATCGAATTCATGGATGATGACTGGG
	3xFLAG-EGFP-BmVasa WT Reverse (vector)	GCTCACCATGGTGAATCCTTGTATCGTCATCCTTGTATC
	3xFLAG-EGFP-BmVasa WT Forward (insert)	TTCAACATGGTGAGCAAGGGC
	3xFLAG-EGFP-BmVasa WT Reverse (insert)	GAATTCGATAGAACCTTGTACAGCTCGTCCATGC
	3xFLAG-EGFP-BmVasa RK Forward (vector)	ACCTATGTGCCTCCAGAACCAAC
	3xFLAG-EGFP-BmVasa RK Reverse (vector)	GAATTCGATAGAACCTTGTACAGCTC
	3xFLAG-EGFP-BmVasa RK Forward (insert)	GGTTCATCGAATTCATGGATGATGACTGGGATGATTCTG
	3xFLAG-EGFP-BmVasa RK Reverse (insert)	TGGAGGCACATAGGTTACCGG
	3xFLAG-BmVasa R470A Forward (mutagenesis PCR)	GCTGAGGAAGCGTTACAGAATTTT
	3xFLAG-BmVasa R470A Reverse (mutagenesis PCR)	CTCACGTTGCATGCGATCGCCGTG
	3xMyc-BmVasa WT Forward (vector)	TAGACGCGTACCGGTCATCATC
	3xMyc-BmVasa WT Reverse (vector)	GAATTCATAATCCTCTCACTAAT
	3xMyc-BmVasa WT Forward (insert)	GAGGATTAGAAATTCATGGATGATGACTGGGATGATTCC
	3xMyc-BmVasa WT Reverse (insert)	ACCGGTACGCGTCTACCATCTTCTTCAGGTTT
	3xMyc-BmVasa ΔN Forward (vector)	TAGACGCGTACCGGTCATCATC
	3xMyc-BmVasa ΔN Reverse (vector)	GAATTCATAATCCTCTCACTAAT
	3xMyc-BmVasa ΔN Forward (insert)	GAGGATTAGAAATTCACCTATGTGCCTCCAGAACCAACA
	3xMyc-BmVasa ΔN Reverse (insert)	ACCGGTACGCGTCTACCATCTTCTTCAGGTTT
	3xMyc-BmVasa ΔC Forward (vector)	TAGACGCGTACCGGTCATCATC
	3xMyc-BmVasa ΔC Reverse (vector)	GAATTCATAATCCTCTCACTAAT
	3xMyc-BmVasa ΔC Forward (insert)	GAGGATTAGAAATTCATGGATGATGACTGGGATGATTCC
	3xMyc-BmVasa ΔC Reverse (insert)	ACCGGTACGCGTCTACTTTAAGAAAGTCTGGGACAGATTG
	3xMyc-BmVasa R470A Forward (mutagenesis PCR)	CGTGAGGCAGAGGAAGCGTTACAGAATTTAAGAG
	3xMyc-BmVasa R470A Reverse (mutagenesis PCR)	TTCTCTGCCTCAGTTCATGCGGATC
	pET47b-BmVasa WT Forward (vector)	TAGGGGTACCAGGATCCGAATTC
	pET47b-BmVasa WT Reverse (vector)	TCCAGGGGTCCCTGAAAG
	pET47b-BmVasa WT Forward (insert)	CAGGGACCCCTGGGAATGGATGATGACTGGGATGATTCC
	pET47b-BmVasa WT Reverse (insert)	ATCCTGGTACCCCTACCATCTTCTTCAGGTTT
	pET47b-BmVasa ΔN Forward (vector)	TAGGGGTACCAGGATCCGAATTC
	pET47b-BmVasa ΔN Reverse (vector)	TCCAGGGGTCCCTGAAAG
	pET47b-BmVasa ΔN Forward (insert)	CAGGGACCCCTGGGAACCTATGTGCCTCCAGAACCAACA
	pET47b-BmVasa ΔN Reverse (insert)	ATCCTGGTACCCCTACCATCTTCTTCAGGTTT
	pET47b-BmVasa R470A Forward (mutagenesis PCR)	CGTGAGGCAGAGGAAGCGTTACAGAATTTAAGAG
	pET47b-BmVasa R470A Reverse (mutagenesis PCR)	TTCTCTGCCTCAGTTCATGCGGATC
	pET47b-BmVasa E339Q Forward (mutagenesis PCR)	CTGGATCAGGCTGACCGATGTAGATATGG
	pET47b-BmVasa E339Q Reverse (mutagenesis PCR)	GTCAGCCTGATCCAGTACGACAAATCTTACACTGC
	pET47b-BmVasa N-IDR Forward (mutagenesis PCR)	AGCCGGTATAGGGGTACCAGGATCCGAATTTCTG
	pET47b-BmVasa N-IDR Reverse (mutagenesis PCR)	ACCCCTATACCGGCTTCTGTTTCCCGTTTTC
in vitro ssRNA transcription	50 nt RNA (lacZ) Forward	AAAAATTAATACGACTCACTATAGGGATGATAGATCCCGTCGTTTTA
	50 nt RNA (lacZ) Reverse	CCAGGGTTTTCCAGTCACG
	50 nt RNA (lacZ antisense) Forward	ATGATAGATCCCGTCGTTTTACAAC
	50 nt RNA (lacZ antisense) Reverse	AAAAATTAATACGACTCACTATAGGGCCAGGGTTTTCCAGTCACG
	50 nt RNA (CD) Forward	AAAAATTAATACGACTCACTATAGGGATGGCTAAACAAGCAAAGATAG
	50 nt RNA (CD) Reverse	TCAAAATGGGCATCATTATGTC
	2000 nt RNA (lacZ) Forward	AAAAATTAATACGACTCACTATAGGGATGATAGATCCCGTCGTTTTA
	2000 nt RNA (lacZ) Reverse	AGAGGCACTTCACCGCTTGCC
	2000 nt RNA (CD) Forward	AAAAATTAATACGACTCACTATAGGGATGGCTAAACAAGCAAAGATAG
	2000 nt RNA (CD) Reverse	ATACCAAAAAATATTTCATAAGGAGTTTATTCTC
RNAi	siLuc Sense	CGUACGCGGAUACUUCGATT
	siLuc antisense	UCGAAGUUAUCCGCGUACGTT
	siBmVasa Sense	CAGCACGCGACUAUGAGATT
	siBmVasa antisense	UCUCAUAGUCGUGUGUGTT
	siSiwi Sense	CACUCGGAGGAUAUCUUUUTT
	siSiwi antisense	AAGAGUAUCCUCCGAGUGTT
	siSpn-E Sense	CAGACACUCGUUAUACAUAUATT
	siSpn-E antisense	UAAUGUAUACGAGUGUCUGTT

*: RNA is shown in italics

Description of Additional Supplementary Files

Supplementary Movie1: Live imaging of EGFP-BmVasa in BmN4 cells. A longer time-frame movie of Fig. 2a. Frame rate: 7 fps. 1 frame: 3 sec. Scale bar: 2 μm .