Supplementary Information

Table of contents

Supplementary Figures

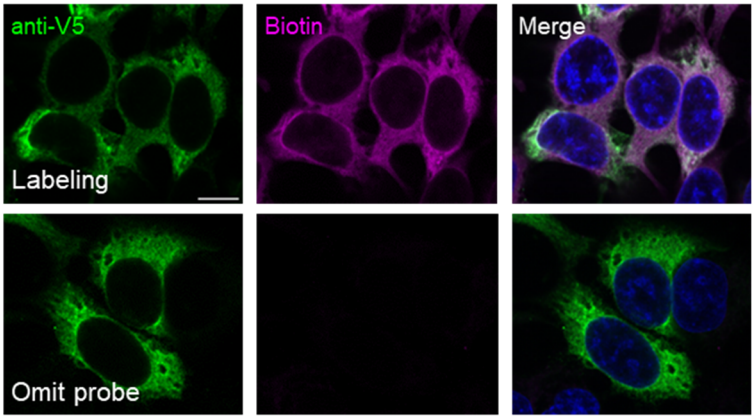

Figure S1. Representative immunofluorescence images of APEX2 expression (green), biotinylation (magenta), and DAPI (blue) in APEX2-NES cell line. Top: APEX2-NES cells labeled with BP. Bottom: Negative control omitting BP probe. Scale bar, 20 μm.

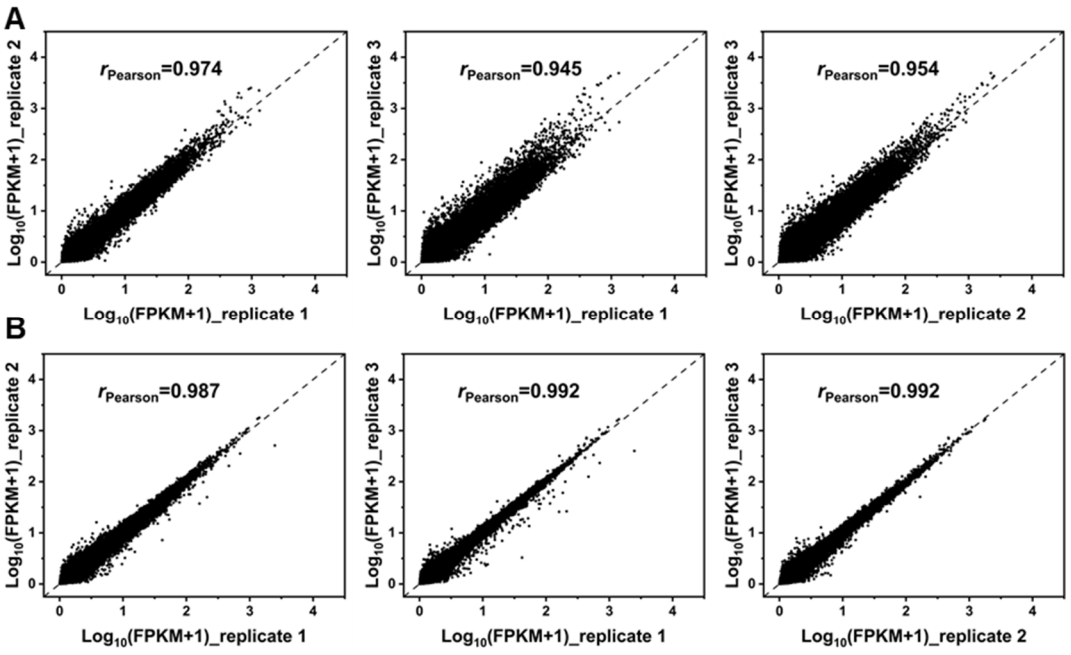

Figure S2. Comparisons of RNA abundance between labeled samples of three independent biological replicates in APEX2-PCNT-EGFP (A) and APEX2-NES (B) cell lines.

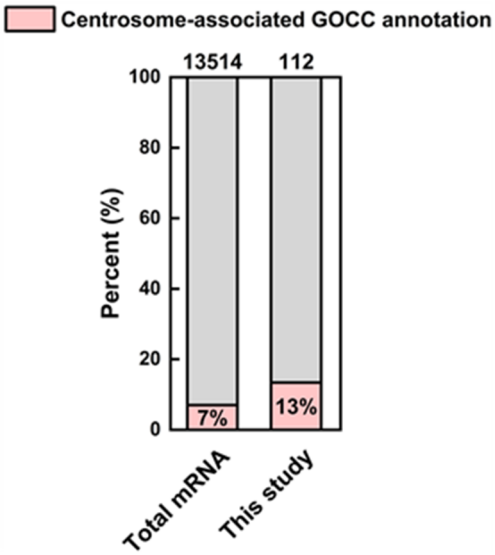

Figure S3. Stack bar graph indicating percentage of genes with centrosome-associated GOCC annotations in our dataset. Total mRNA denotes mRNAs with FPKM more than 0.1 in the APEX2-PCNT-EGFP cell line.

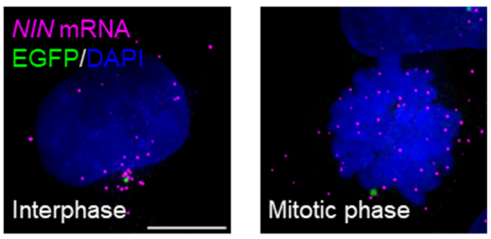

Figure S4. Representative smFISH images of *NIN* mRNA in HEK293T cell line in interphase and mitotic phase. Magenta: endogenous *NIN* mRNA; green: centrosome-localized EGFP fusion proteins; blue: DNA stained with DAPI. Scale bars, 10 μm.

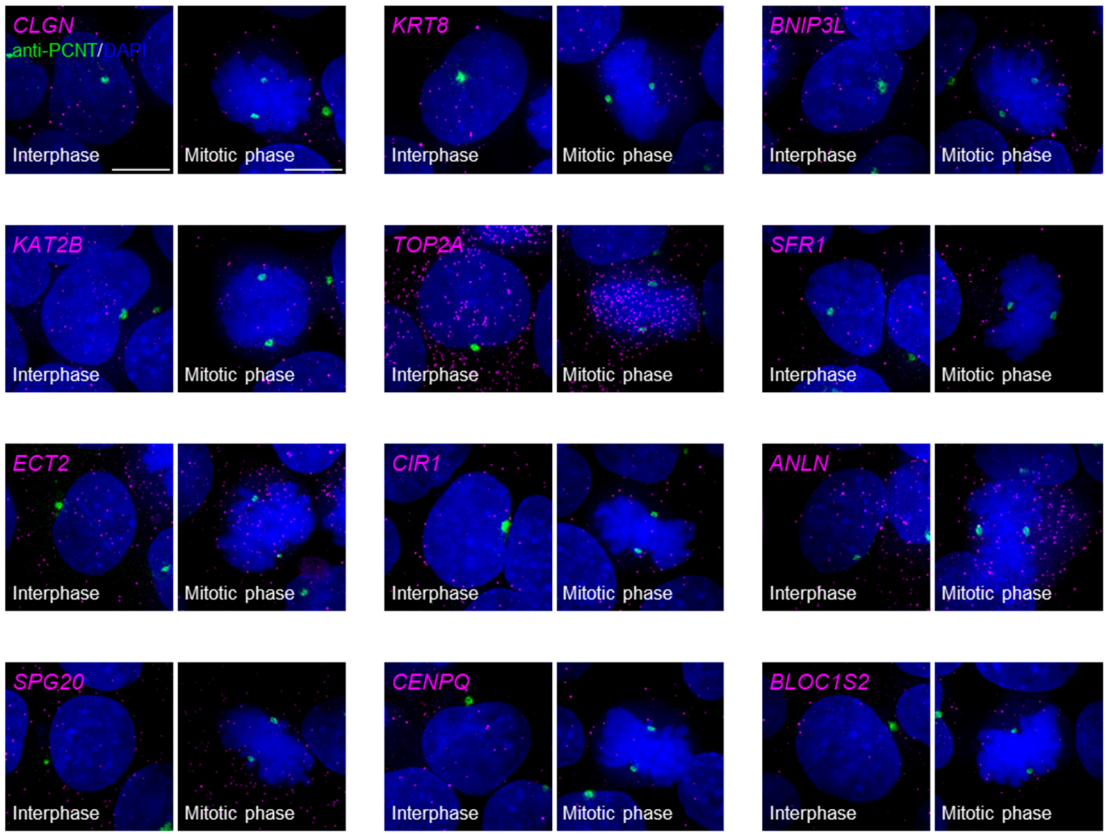

Figure S5. Representative smFISH images of enriched centrosome-associated mRNAs in HEK293T cell line in interphase and mitotic phase. Magenta: endogenous mRNA; green: anti-PCNT immunofluorescence to label the centrosome; blue: DNA stained with DAPI. Scale bars, 10 μm.

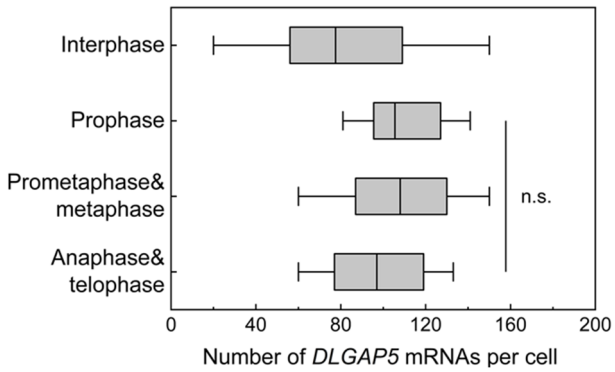

Figure S6. Box plots showing variations in the expression levels of *DLGAP5* mRNA in HEK293T cell line during different cell cycle phases, related to Figures 2A and 2B. Statistical significance was calculated by a two-sided Mann-Whitney test. n.s.: not significant.

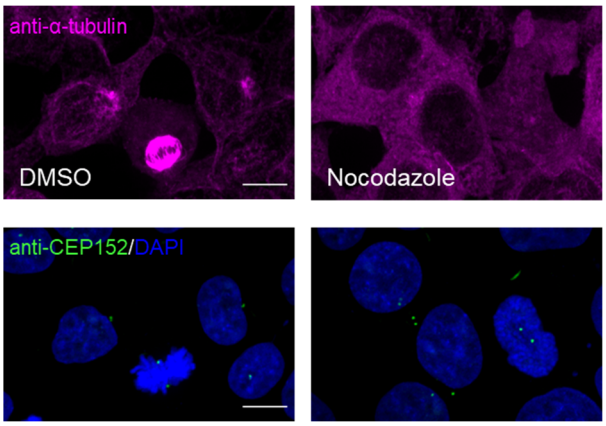

Figure S7. Representative immunofluorescent images of HEK293T cells treated with nocodazole. Magenta: anti-α-tubulin immunofluorescence to label microtubules; green: anti-CEP152 immunofluorescence to label centrosome; blue: DNA stained with DAPI. Scale bar, 10 μm.

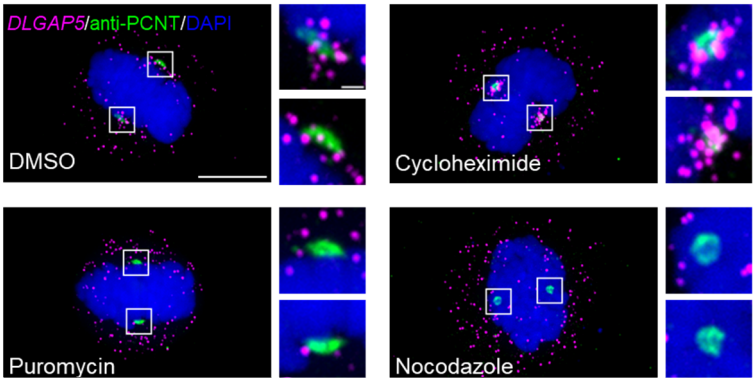

Figure S8. Representative smFISH images of *DLGAP5* mRNA in mitotic HeLa cells treated with DMSO, cycloheximide, puromycin, and nocodazole. Magenta: endogenous human *DLGAP5* mRNA; green: anti-PCNT immunofluorescence; blue: DNA stained with DAPI. Scale bars, 10 μm. Zoom-in images of boxed regions are shown in the right. Scale bar, 1 μm.

mRNA	CDS Length (nt)
ASPM	10434
PCNT	10011
CEP350	9354
NUMA1	6348
NIN	6273
CCDC88C	6087
BICD2	2475
HMMR	2175

Figure S9. CDS length of currently-discovered centrosomally-localized mRNAs.

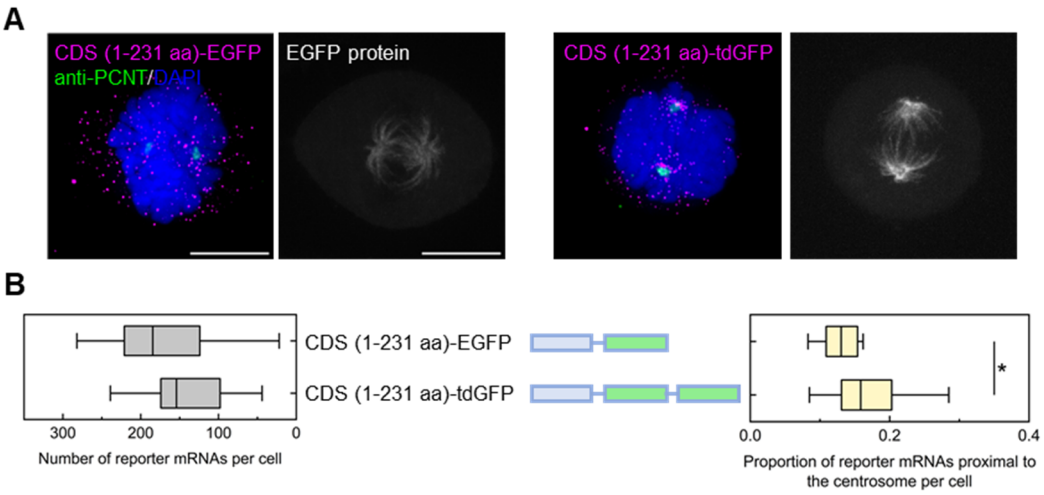

Figure S10. Evaluation of influence of CDS length on *DLGAP5* mRNA localization. (A) Representative smFISH images of mitotic HeLa cells expressing EGFP or tdGFP fusion proteins. Magenta: smFISH probes against EGFP; green: anti-PCNT immunofluorescence; blue: DNA stained with DAPI; grey: EGFP or tdGFP fusion proteins. Scale bars, 10 μm. **(B)** Left: Transcript abundance per cell for EGFP-containing reporter mRNAs as measured by smFISH. Right: box plots depicting the proportion of reporter mRNAs proximal to the centrosome. Data were collected from 18 cells in two biological replicates. Statistical significance was calculated by a two-sided Mann-Whitney test. *: $p < 0.05$.

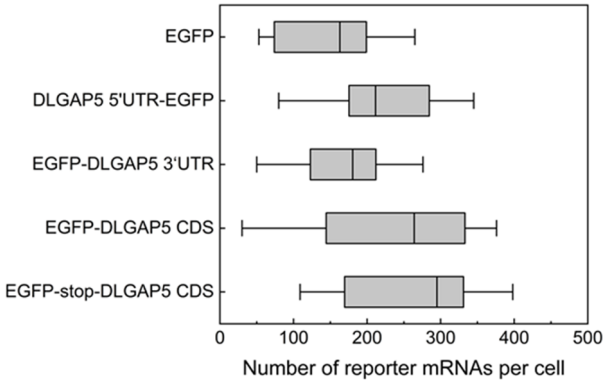

Figure S11. Transcript abundance per cell for EGFP-containing reporter mRNAs in HeLa cells as measured by smFISH, related to Figure 3C and 3D.

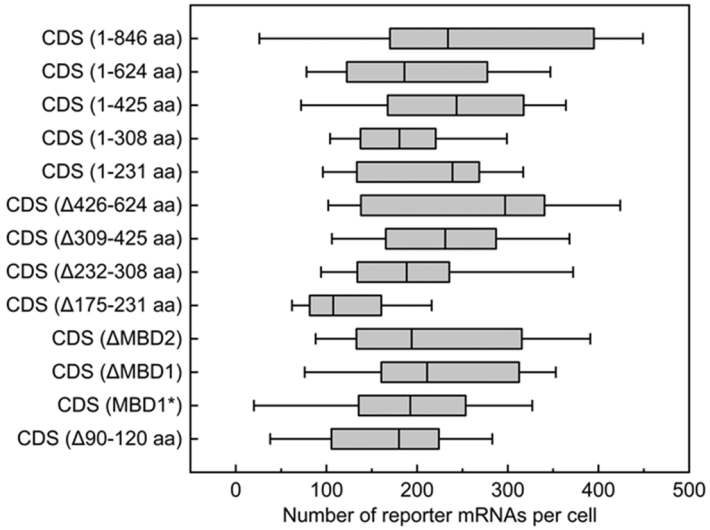

Figure S12. Transcript abundance per cell for EGFP-containing reporter mRNAs in HeLa cells as measured by smFISH, related to Figure 4.

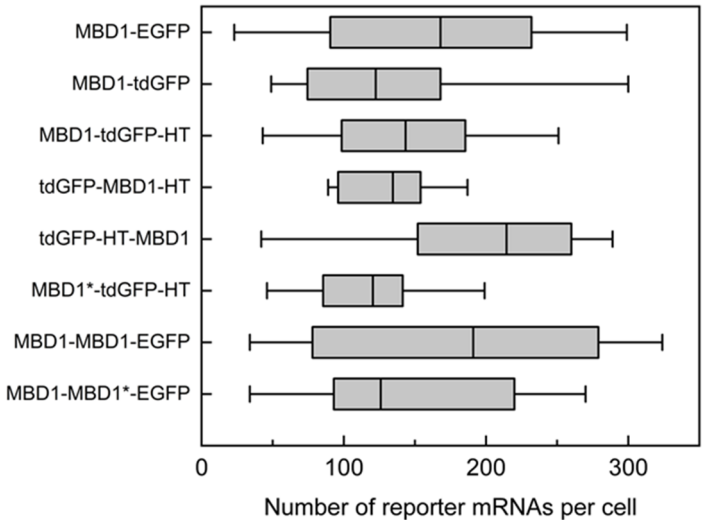

Figure S13. Transcript abundance per cell for EGFP-containing reporter mRNAs in HeLa cells as measured by smFISH, related to Figure 5.

Supplementary Tables

Table S1. Reagents & resources used in this study

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
V5-Tag Monoclonal Antibody (3C8)	Biodragon	Cat# B1005
Anti-CEP152 antibody produced in rabbit	Sigma-Aldrich	Cat# HPA039408
Monoclonal Anti-γ-Tubulin antibody produced in mouse	Sigma-Aldrich	Cat# T5326
Anti-Pericentrin antibody	abcam	Cat# ab4448
Anti-CDK5RAP2 antibody produced in rabbit	Sigma-Aldrich	Cat# HPA046529
alpha Tubulin Monoclonal Antibody (YL1/2)	Invitrogen	Cat# MA180017
Streptavidin, Alexa Fluor 568 conjugate	Invitrogen	Cat# S11226
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	Invitrogen	Cat# A-21236
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 568	Invitrogen	Cat# A-11036
Goat anti-Rat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	Invitrogen	Cat# A-11006
Chemicals, peptides, and recombinant proteins		
Dulbecco’s Modified Eagle’s Medium	Gibco	Cat# C11995500BT
McCoy's 5A Medium	BI	Cat# 01-075-1ACS
0.25% Trypsin-EDTA	Gibco	Cat# 25200056
0.05% Trypsin-EDTA	Gibco	Cat# 25300120

Fetal bovine serum	Gibco	Cat# 10099141C
Opti-MEM I Reduced Serum Medium	Gibco	Cat# 31985062
Lipofectamine 3000 Transfection Reagent	Invitrogen	Cat# L3000015
Matrigel Matrix	Corning	Cat# 356234
Blasticidin S HCl	Selleck	Cat# S7419
TRIzol reagent	Invitrogen	Cat# 15596018
DNase I (RNase-free)	NEB	Cat# M0303
Yeast tRNA	Invitrogen	Cat# 15401011
Bovine serum albumin	Sangon Biotech	Cat# A500023-0100
Dynabeads MyOne Streptavidin C1 beads	Invitrogen	Cat# 65002
RiboLock RNase Inhibitor	Thermo Scientific	Cat# EO0384
NaCl (5 M), RNase-free	Invitrogen	Cat# AM9759
EDTA (0.5 M), pH 8.0, RNase-free	Invitrogen	Cat# AM9260G
UltraPure 1 M Tris-HCl Buffer, pH 7.5	Invitrogen	Cat# 15567027
PBS (10×) pH 7.4, RNase-free	Invitrogen	Cat# AM9625
D-Biotin Solution (50 mM)	Invitrogen	Cat# B20656
DMSO	Aladdin	Cat# D103273
Formamide	Sigma-Aldrich	Cat# F9037
Urea	Sigma-Aldrich	Cat# U5378
Sodium hydroxide solution	Sigma-Aldrich	Cat# S2770
SDS, 10% Solution, RNase-free	Invitrogen	Cat# AM9822
Tween-20	Sigma-Aldrich	Cat# P1379
Triton X-100	Aladdin	Cat# T109026

BeyoPure Ultrapure Water	Beyotime	Cat# ST876
Glycogen, RNA grade	Thermo Scientific	Cat# R0551
SSC (20×), RNase-free	Invitrogen	Cat# AM9770
tRNA from *E. coli* MRE 600	Roche	Cat# 10109541001
UltraPure Salmon Sperm DNA Solution	Invitrogen	Cat# 15632011
Ribonucleoside Vanadyl Complex (RVC)	NEB	Cat# S1402S
Dextran sulfate sodium salt	Sigma-Aldrich	Cat# D6001
Paraformaldehyde	Sigma-Aldrich	Cat# V900894
$MgCl_2$ (1 M)	Invitrogen	Cat# AM9530G
Glycine	Solarbio	Cat# G8200
Nocodazole	TargetMol	Cat# T2802
Puromycin	J&K	Cat# 168086
Cycloheximide	Fluorochem	Cat# 375034
Sodium ascorbate	Aladdin	Cat# S105024
Trolox	Sigma-Aldrich	Cat# 238813
Biotin-phenol (BP)	Sigma-Aldrich	Cat# SML2135
2-Amino-6-mercaptopurine-9-D-riboside Hydrate (s^6G)	Sigma-Aldrich	Cat# 858412
Fluoromount-G Anti-Fade	SouthernBiotech	Cat# 0100-35
PowerUp SYBR Green Master Mix	Applied Biosystems	Cat# A25742
Lightening cloning kit	Biodragon	Cat# BDIT0014
Phanta Max Super-Fidelity DNA Polymerase	Vazyme	Cat# P505
VAHTS DNA Clean Beads	Vazyme	Cat# N411
RNA Clean & Concentrator-25	Zymo Research	Cat# R1018

Critical commercial assays		
NEBNext Ultra II RNA Library Prep Kit for Illumina	NEB	Cat# E7770
NEBNext Poly(A) mRNA Magnetic Isolation Module	NEB	Cat# E7490L
Fragment Analyzer RNA Kits	AATI	Cat# DNF-471-0500
Fragment Analyzer DNA/NGS Kits	AATI	Cat# DNF-474-0500
ProtoScript II First Strand cDNA Synthesis Kit	NEB	Cat# E6560
Experimental models: Cell lines		
HEK293T	NICLR	1101HUM-PUMC000212
HeLa S3	NICLR	1101HUM-PUMC000188
SH-SY5Y	NICLR	1101HUM-PUMC000026
U-2 OS	NICLR	1101HUM-PUMC000028
Neuro-2a	NICLR	1101MOU-PUMC000291
MCF-7	Prof. Mo Li (PKU)	N/A
MDA-MB-231	Prof. Mo Li (PKU)	N/A
Oligonucleotides		
smFISH oligos	This study	Data S2
Recombined DNA		
EGFP reporter plasmids	This study	Table S2
pLX304-V5-APEX2-NES	This study	BamHI-V5-APEX2-NES-NheI; NES: LQLPPLERLTLD
pLX304-V5-APEX2-PCNT-EGFP	This study	BstBI-V5-APEX2-PCNT_PACT domain-EGFP-NheI
Software and algorithms		
HISAT2	Reference [1]	RRID: SCR_015530

HTSeq	Reference [2]	RRID: SCR_005514
DESeq2	Reference [3]	RRID: SCR_015687
Fiji	Reference [4]	RRID: SCR_002285
FISH-quant	Reference [5]	PMID: 23538861

* National Infrastructure of Cell Line Resource

Table S2. EGFP reporter plasmids used in this study

Reporter name	Description	CDS length (nt)
EGFP	pUbC-EGFP	753
EGFP-DLGAP5 3'UTR	pUbC-EGFP-DLGAP5 3'UTR	753
DLGAP5 5'UTR-EGFP	pUbC-DLGAP5 5'UTR-EGFP	753
EGFP-DLGAP5 CDS	pUbC-EGFP-DLGAP5 CDS	3288
EGFP-stop-DLGAP5 CDS	pUbC-EGFP-*UAG*-DLGAP5 CDS	753
CDS (1-846 aa)	pUbC-DLGAP5 CDS-EGFP	3273
CDS (1-624 aa)	pUbC-DLGAP5 CDS (1-624 aa)-EGFP	2607
CDS (1-425 aa)	pUbC-DLGAP5 CDS (1-425 aa)-EGFP	2010
CDS (1-308 aa)	pUbC-DLGAP5 CDS (1-308 aa)-EGFP	1659
CDS (1-231 aa)-EGFP	pUbC-DLGAP5 CDS (1-231 aa)-EGFP	1428
CDS (Δ175-231 aa)	pUbC-DLGAP5 CDS (Δ175-231 aa)-EGFP	3102
CDS (Δ232-308 aa)	pUbC-DLGAP5 CDS (Δ232-308 aa)-EGFP	3042
CDS (Δ309-425 aa)	pUbC-DLGAP5 CDS (Δ309-425 aa)-EGFP	2922
CDS (Δ426-624 aa)	pUbC-DLGAP5 CDS (Δ426-624 aa)-EGFP	2676
CDS (ΔMBD2)	pUbC-DLGAP5 CDS (Δ2-69 aa)-EGFP	3069
CDS (ΔMBD1)	pUbC-DLGAP5 CDS (Δ65-174 aa)-EGFP	2943
CDS (MBD1*)	pUbC-DLGAP5 CDS (K105A, K107A, R110A, K112A, K114A, R115A)-EGFP	3273
CDS (Δ90-120 aa)	pUbC-DLGAP5 CDS (Δ90-120 aa)-EGFP	3180
CDS (1-231 aa)-tdGFP	pUbC-DLGAP5 CDS (1-231 aa)-tdGFP	2172
MBD1-EGFP	pUbC-DLGAP5 CDS (65-174 aa)-EGFP	1068

MBD1-tdGFP	pUbC-DLGAP5 CDS (65-174 aa)-tdGFP	1812
MBD1-tdGFP-HT	pUbC-DLGAP5 CDS (65-174 aa)-tdGFP-Halotag	2724
MBD1*-tdGFP-HT	pUbC-DLGAP5 CDS (65-174 aa, K105A, K107A, R110A, K112A, K114A, R115A)-tdGFP-Halotag	2724
tdGFP-MBD1-HT	pUbC-tdGFP-DLGAP5 CDS (65-174 aa)-Halotag	2724
tdGFP-HT-MBD1	pUbC-tdGFP-Halotag-DLGAP5 CDS (65-174 aa)	2724
MBD1-MBD1-EGFP	pUbC-2×DLGAP5 CDS (65-174 aa)-EGFP	1422
MBD1-MBD1*-EGFP	pUbC-DLGAP5 CDS (65-174 aa)-DLGAP5 CDS (65-174 aa, K105A, K107A, R110A, K112A, K114A, R115A)-EGFP	1422

Table S3. Analysis of previously-reported centrosome-localized mRNAs

Gene ID	Gene name	PCNT_En vs NES_En		PCNT_In vs NES_In		PCNT_En vs PCNT_Ctrl		Significantly enriched?
		log_2FC	padj	log_2FC	padj	log_2FC	padj	
ENSG00000072571	HMMR	2.78	7.80E-31	1.47	7.49E-17	0.50	2.52E-01	No
ENSG00000066279	ASPM	0.94	1.85E-04	1.71	2.03E-27	1.41	1.10E-03	No
ENSG00000135837	CEP350	0.84	3.10E-03	1.97	4.39E-22	1.48	5.32E-05	No
ENSG00000185963	BICD2	-1.20	8.47E-07	-1.11	7.91E-16	3.51	1.01E-21	No
ENSG00000100503	NIN	1.65	1.52E-11	0.98	6.55E-08	2.92	4.58E-16	Yes
ENSG00000160299	PCNT	4.34	1.47E-71	1.33	1.83E-32	3.69	6.49E-27	No
ENSG00000015133	CCDC88C	0.48	6.12E-09	1.40	6.12E-09	0.59	3.79E-01	No
ENSG00000137497	NUMA1	-0.26	1.52E-07	0.63	1.52E-07	-0.07	9.10E-01	No

References

1. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357-60. doi: 10.1038/nmeth.3317. PMID: 25751142.
2. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9. doi: 10.1093/bioinformatics/btu638. PMID: 25260700.
3. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PMID: 25516281.
4. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82. doi: 10.1038/nmeth.2019. PMID: 22743772.
5. Mueller F, Senecal A, Tantale K, Marie-Nelly H, Ly N, Collin O, et al. FISH-quant: automatic counting of transcripts in 3D FISH images. Nat Methods. 2013;10(4):277-8. doi: 10.1038/nmeth.2406. PMID: 23538861.