

Supplementary Figures and Tables

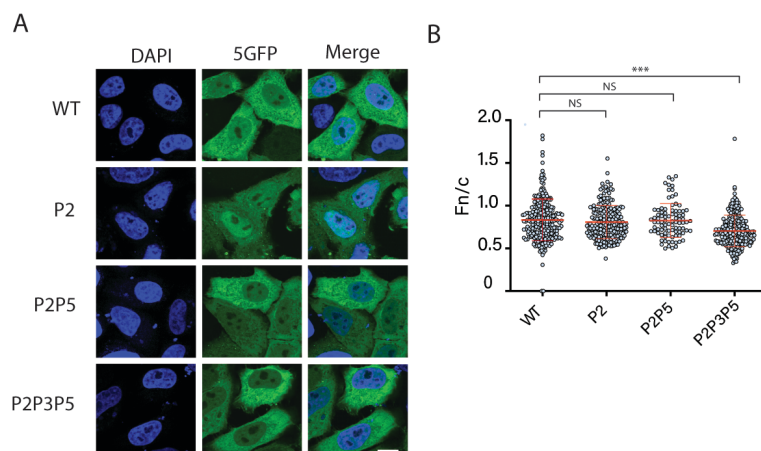


Figure S1. Functional analysis of the three NLS2 variants absent in influenza A virus NP. **(A)** Confocal images of HeLa cells transfected with plasmids expressing 5GFP or three remaining NLS2 variants fused to 5GFP 24 hours post-transfection. Nuclei were stained with DAPI. Scale bar, 10 μ m. **(B)** Quantification of the ratio of nuclear to cytoplasmic fluorescence (Fn/c) from the experimental conditions shown in B. Shown is the means \pm standard error of the means scored from 85-100 cells for each condition from three independent experiments. (NS, not significant; *** $p < 0.001$, one-way ANOVA followed by Tukey's tests).

Table S1. NLS2 variants located at NP amino acid residues different than 212-216.

NLS2 Variant	Location of NLS2 in the Protein	Accession Number
WT	218 to 222	BAL70445.1, BAL70455.1, BAK41897.1
P3	213 to 217	ADU16942.1, ADU53204.1
P3	218 to 222	ACU51033.1, ACU51030.1, BAM78353.1, ATW75159.1, ABY81789.2, ANZ03693.1, ANK78256.1, BAM78385.1, ATW75162.1, AAN46830.1, AAV30838.1, 2Q06_A, 2Q06_B, ATW75160.1, ABY70960.1, BAK41893.1, BAK41894.1, BAK41895.1, BAK41896.1, BAK41898.1, ATW75163.1, ATW75164.1, ATW75165.1, ATW75166.1, ABF48471.1, BAM65747.1, ATW75168.1, BAM78369.1, ACU51029.1, ACU51031.1, ATW75155.1, ATW75156.1, ATW75158.1, BAL70427.1, ADI52818.1, ADI52820.1, BAL70465.1, AEX30601.1, ATW75169.1, ATW75157.1, BAG49623.1, ALT19014.1, ACR01026.1, ATW75167.1, BAM78401.1, AAV30830.1, ACU51032.1, ACU51034.1, ACU51035.1, APP91838.1

Table S2. BLASTP E-values for influenza A NP and proteins from DNA virus families reported in Figures 2 and 4 ¹.

Variant	Influenza A NP	Adenoviridae	Herpesviridae	Circoviridae	Hepadnaviridae
WT	224.0 \pm 0.0	No hit	No hit	No hit	No hit
P2	No hit	13.0 \pm 0.0	71.0 \pm 0.0	No hit	No hit
P3	158.0 \pm 0.0	6.7 \pm 0.1	35.0 \pm 0.0	2.2	No hit
P5	318.0 \pm 0.0	No hit	71.0 \pm 0.0	No hit	No hit
P2P3	224.0 \pm 0.0	9.2 \pm 0.0	50.0 \pm 0.0	No hit	No hit
P2P5	No hit	No hit	Not hit	No hit	No hit
P3P5	224.1 \pm 0.0	9.2 \pm 0.0	No hit	No hit	No hit
P2P3P5	No hit	No hit	No hit	No hit	78.0 \pm 0.0
Random peptide (GLVTV)	No hit	No hit	No hit	6.9 \pm 0.0	No hit
SV40 Large T-antigen NLS	No hit	No hit	No hit	No hit	No hit

¹ Mean of E-value \pm standard error of the mean are listed.

Table S3. Proteins from DNA virus families containing variants of the sequence GRKTR.

GRKTR Variant	Protein	Accession Number
<i>Adenoviridae</i> family		
P2	Minor core protein pV	AQQ73631
	Viral RNA splicing factor L4-33 kDa	YP_009373251, ART33376
	Viral transcription factor L4-22 kDa	YP_002822221, ACJ14521, YP_002822220, ACJ14522
P3	Precursor terminal protein pTP	YP_007518314, AEK98448, QHR77492
	Encapsidation protein L1-52/55 kDa	QJC19249
	Late L2 mu core protein pX	AWB14650, YP_004123745, ADR77841, AP_000057, AAB05441, P68974, P68975
P2P3	Minor core protein pV	AJE59778, AGT76023, AJE59852, AKQ98460, AJE59724, AJE59783, AJE59788, AJE59808, AKQ98454, AFQ34581, AJE59813, AFQ34346, AGT76107, AGT77115, AFQ34307, AJE59734, AJE59798, AJE59719, AJE59754, AJE59759, AJE59769, AKQ98442, AJE59803, AGT77293, AJE59709, AJE59764, AGT76151, AGT75892, AGT76065, AGT76536, AJE59857, AJE59862, AKQ98376, AAZ13829, AFQ34385, AGT75547
P3P5	Hexon-associated structural protein (pVIII precursor)	AGT75675, AUN87399, AFQ34424, AJE59867, AJE59739, AJE59744, AAZ15253, AJE59881, AJE59895, AKQ98448, YP_002213776, AJE59876, AKQ98418, AET87309, AAW33441, AAW33258, ANQ44536, AAW33166, ABB1780, QEQ50111, QEQ50148, AMB61156, AP_000545, AAR89960, AAT97584, QBG39296, QEQ50074, AMB61089, QID75502, AAT97543, ACX32405, AEC11867, AAW33399, AMB61124, AYI99658, AET87268, AET87227, AET87145, AET87186, QFR07974
<i>Herpesviridae</i> family		
WT	Tegument protein UL37	YP_00917690, AIU39240, AKV40709, AIU39350, AKI81626, AKI81730, AAY59063
P2	GP84	CDI95420, YP_007417850, AGE11554, BAJ78542
P2	DNA polymerase processivity subunit	ACY41129, NP_044897, AAB66449, CAA70274
P3	Tegument protein VP22	BBM13221
P5	Tegument protein UL32	AEQ32128
P2P3	Assembly protein M80	CCE57080, AWW68505, ADD10452, CCE56751, CCE56914, CCE57244, YP_214083, AAA46000, AQQ81427, AWW68681, CCE56585, CCE57408, AAA46001, CAP08120, ACE95421, ACE95257, ACE95585
<i>Circoviridae</i> family		
P3	Capsid protein	QFR58250
<i>Hepadnaviridae</i> family		
P2P5	X protein	AGO96886

Table S4. Nuclear proteins containing variants of the sequence GRKTR found only in humans and their predicted NLSs.

Protein ¹	Organism	Putative NLS2 Variant	Predicted NLS ²
Ubl carboxyl-terminal hydrolase 18	<i>H. sapiens</i>	P2: 234GKKTR ²³⁸ 160GKKTR ¹⁶⁴	No NLS predicted
Transcriptional repressor CTCFL	<i>H. sapiens</i>	P3: 581GRRTR ⁵⁸⁵ 531GRRTR ⁵³⁵ 376 GRRTR ³⁸⁰ 319 GRRTR ³²³	576AASGK GRRTR KRKQTILKEATKGQKEA ⁶⁰²
Sororin	<i>H. sapiens</i>	P3: 3 GRRTR ⁷	12AQRSGPRAPSPTKPLRRSQRK ³² 212PEKQKRKKKKKMPEILSLRGRQRA ²³⁴
N-lysine methyltransferase SETD6	<i>H. sapiens</i>	P3: 46GRRTR ⁵⁰	No NLS predicted
Probable rRNA-processing protein EBP2	<i>H. sapiens</i>	P2P3: 349GKRTR ³⁵³ 294GKRTR ²⁹⁸	277QRKKAGAKGQQMRKGPSAKRRYKNQKFGF GGKKKGSKW ³¹⁴ 328AKTAHGRGLKRPKGKGSNKR PGRTR EKMK NRTH ³⁶¹
Torsin-1A-interacting protein 1	<i>H. sapiens</i>	P2P3: 82GKRTR ⁸⁶	No NLS predicted
Interferon-induced, double-stranded RNA-activated protein kinase]	<i>H. sapiens</i>	P2P3: 443GKRTR ⁴⁴⁷ 402GKRTR ⁴⁰⁶	No NLS predicted
Transforming growth factor beta regulator 1	<i>H. sapiens</i>	P2P5: 133GKKTK ¹³⁷	19KARMKKLPKKSQNEKYRLKYLRRLKA ⁴⁴ 133 GKKTK KEKKEKGKNNKLEVLKKTCKKKKM AGGARK ¹⁶⁸
Histone-lysine N-methyltransferase NSD2	<i>H. sapiens</i>	P2P5: 1218GKKTK ¹²²² 1149GKKTK ¹¹⁵³ 437GKKTK ⁴⁴¹	164KARRNRKRS ¹⁷² 1217 KGKKTK KKTRRRRAKGEGKRQ ¹²³⁷ 1339EKPPPEPGKPKGKRRRRRGWRRVTEGK ¹³⁶⁵
Chromobox protein homolog 5	<i>H. sapiens</i>	P2P5: 2GKKTK ⁶	69KYKKMKEGNNKPREKSEGNKRKS ⁹² 103SKKK ¹⁰⁶ 28VKKGKKDKKIKK ³⁹ 74KKKRDTRKGRRKK ⁸⁶ 119KPRGGKK ¹²⁵ 211KEPPKQGKEKAKK ²²³ 518KKLKELKAGGKSTK ⁵³¹
ATP-binding cassette sub-family F member 1	<i>H. sapiens</i>	P2P5: 123GKKTK ¹²⁷	
Cyclin-F	<i>H. sapiens</i>	P3P5: 567GRRTK ⁵⁷¹ 259GRRTK ²⁶³	21RRIRRR ²⁶

¹ All listed proteins are known to function in the cell nucleus.² NLSs were predicted using NLStradamus (Nguyen et al. 2009; BMC Bioinformatics. 10:202). Overlapping regions of the predicted NLSs containing GRKTR or its variants are highlighted in red.

Table S5. Nuclear proteins from different cellular organisms containing the sequence GRKTR or its variants and their predicted NLSs.

Protein. ¹	Species	Putative NLS2 Variant	Predicted NLS. ²
Suppressor of mec-8 and unc-52 protein	<i>G. max</i>	WT: 513GRKTR ⁵¹⁷ 515GRKTR ⁵¹⁷	6KKNPKEKPIRRKEEKPEEPEVPKYRDRAKER ³⁶ 215GKILKKKRKEKDAKGK ²³⁰
Spliceosomal protein Beag	<i>D. melanogaster</i>	WT: 504GRKTR ⁵⁰⁸	77LRRKKK ⁸² 314GGRNKKNKRDKDK ³²⁷
Methyl-CPG-binding domain protein 13	<i>A. thaliana</i>	WT: 54GRKTR ⁵⁸	No NLS predicted
U-box domain-containing protein	<i>G. max</i>	WT: 509GRKTR ⁵¹³ 501GRKTR ⁵⁰⁵	No NLS predicted
Regulator of apoptosis protein XIAP-associated factor	<i>M. musculus</i>	WT: 130GRKTR ¹³⁴	No NLS predicted
Atpase family AAA domain-containing protein 5	<i>M. musculus</i>	P2: 477GKKTR ⁴⁸¹	111KRKRK ¹¹⁵ 202RKRKR ²⁰⁶ 466PKEKSKKPNKKGKKTR ⁴⁸¹ 654PKSKSKSSKK ⁶⁶³ 1543LRKSQKRKQKK ¹⁵⁵³
Transcription factor Scarecrow-like protein 14	<i>G. max</i>	P2: 289GKKTR ²⁹³	290KKTRSKKGS ²⁹⁸
Origin of replication complex subunit 1	<i>G. max</i>	P2: 579GKKTR ⁵⁸³ 576GKKTR ⁵⁸⁰	106KKK ¹⁰⁸ 132RKREGEQGVVTRAKRRKSENREKS AKLP ¹⁵⁹
Lysine-specific demethylase 6B	<i>M. musculus</i>	P2: 1069GKKTR ¹⁰⁷³	972GSGKRRQKEHRRHRRACRDSVGRRPREGRAKAKAKAPKEKS RR ¹⁰¹⁴
Kinase regulator Mob2	<i>D. melanogaster</i>	P2: 394GKKTR ³⁹⁸ 232GKKTR ²³⁶ 203GKKTR ²⁰⁷ 97GKKTR ¹⁰¹	972GSGKRRQKEHRRHRRACRDSVGRRPREGRAKAKAKAPKEKS RR ¹⁰¹⁴
Ethylene-responsive transcription factor 9	<i>G. max</i> <i>A. thaliana</i>	P2: 46GKKTR ⁵⁰ 45GKKTR ⁴⁹ 43GKKTR ⁴⁷ 42GKKTR ⁴⁵	No NLS predicted
SMAD/FHA domain-containing transcription factor	<i>A. thaliana</i>	P2: 183GKKTR ¹⁸⁷	No NLS predicted
Smg-4/UPF3 family protein	<i>A. thaliana</i>	P3: 185GRRTR ¹⁸⁹	182RRGGRTRVVSANKPSRPSKRNSEKKKYVEK ²¹³
Lupus La protein homolog	<i>M. musculus</i>	P3: 374GRRTR ³⁷⁸ 373GRRTR ³⁷⁷	199SKVEAKLKAKQ ²⁰⁹ 329KSKGGHAGGRFGSHVFTAARRFKGKGKGNRPGYAGAPKG RGQFHGRTRFDGDDRRRGP MKRGRDGRDREEP ASKHKKREN GARDK ⁴¹⁵
Recq-mediated genome instability protein 1	<i>G. max</i>	P3: 263GRRTR ²⁶⁷	255LKKPPRPRGRTR ²⁶⁷
E3 ubiquitin-protein ligase RNF169	<i>D. rerio</i>	P3: 190GRRTR ¹⁹⁴	526RRGKKRSQK ⁵³⁴
Inflammatory response factor strawberry notch homolog	<i>C. elegans</i>	P2: 1789GRRTR ¹⁷⁹³	1253KDRKAKKRKRDEEEAERLREKVRKREERREKKRRRAIRRAER EKQRR ¹²⁹⁹

Protein timeless homolog	<i>C. elegans</i>	P2: 173GRRTR ¹⁷⁷	307RKAKKRAPNRR ³¹⁸ 520LKKVTKKVKVRKATKSK ⁵³⁷ 1011RSRKKVA ¹⁰¹⁷
Myb domain-containing protein		P2: 699GRRTR ⁷⁰³	283DRERDRDRERERERERDRERERERGRDR ³¹⁰ 556SRERSRSHSRSRSRSRSHSR ⁵⁷⁷ 621KRKSGGDDGKDVKRRKTGS ⁶³⁹
Nuclear inhibitor of protein phosphatase 1	<i>D. melanogaster</i>	P2: 186GRRTR ¹⁹⁰	No NLS predicted
Protein MAIN-LIKE 2-like	<i>G. max</i>	P2: 23GRRTR ²⁷	No NLS predicted
Nuclear receptor corepressor 2	<i>D. rerio</i>	P5: 2268GRKTK ²²⁷² 2258GRKTK ²²⁶² 2252GRKTK ²²⁵⁶ 2242GRKTK ²²⁴⁶ 2230GRKTK ²²³⁴ 2204GRKTK ²²⁰⁸ 1646GRKTK ¹⁶⁵⁰	653KIKSEKERKARRKGKA ⁶⁶⁸ 804KKEVKKKEH ⁸¹² 1674REREREREREREQRERERERERERDRERDRDREREKDRDRDRE R ¹⁷¹⁷ 2260KSKGGRSNGRKTKS ²²⁷³
Histone h1t	<i>M. musculus</i>	P5: 165GRKTK ¹⁶⁹	0KPSSKRRGKKPGLAPARKPRGF ⁴¹ 110LSKKAASGNDKGKGSASAKAKMGLPRASRSPKSSKTKAV KKPKATPTKASGSGRKTKGAKGVQQRK- SPAKARAANPNSGKAKMVMQKTDLRKAAGRK ²⁰⁹
Winged-helix DNA-binding transcription factor family protein	<i>A. thaliana</i>	P5: 35GRKTK ³⁹	29KKPAAKGRKTKNVKEVKEKKTVAAPKKRT ⁵⁸
AGAMOUS-like 57	<i>A. thaliana</i>	P5: 9GRKTK ¹³	7AKGRKTKGKQKI ¹⁸
MRX complex nuclease subunit	<i>S. cerevisiae</i>	P5: 650GRKTK ⁶⁵⁴	588KPKRVRTATKKKI ⁶⁰⁰ 651RKTKTKTRPAASTKTASRRGKGRASRTPKTDILGSLAKKRK ⁶⁹²
Transcription factor MIG1	<i>S. cerevisiae</i>	P5: 159GRKTK ¹⁶³	87RRIHTNSHPRGKRGRKKK ¹⁰⁴
Acyl-CoA N-acyltransferase with RING/FYVE/PHD-type zinc finger protein	<i>A. thaliana</i>	P5: 484GRKTK ⁴⁸⁸	64RRSGVDRKTLGAKEKFVRKR ⁸³ 266RKSKRKKSEYSKPKKE ²⁸¹ 302RRGRRGGGTDKQRL ³¹⁵ 396RKAKKTR ⁴⁰²
Topoisomerase 3alpha	<i>A. thaliana</i>	P5: 584GRKTK ⁵⁸⁸	862GRRGSRGRGRGGRGGQSSGGR ⁸⁸²
Lysine-specific demethylase 4C	<i>M. musculus</i>	P5: 694GRKTK ⁶⁹⁸ 677GRKTK ⁶⁸¹ 602GRKTK ⁶⁰⁶ 496GRKTK ⁵⁰⁰	379RKKKLRKPPKSLQGNK ³⁹⁴
Probable serine/threonine-protein kinase At1g54610	<i>G. max</i>	P5: 523GRKTK ⁵²⁷	423GAKKVRVRERERG ⁴³⁵
Caspase-14	<i>M. musculus</i>	P5: 151GRKTK ¹⁵⁵	No NLS predicted
ELM2 and Myb/SANT-like domain containing 1a	<i>D. rerio</i>	P5: 913GRKTK ⁹¹⁷	No NLS predicted
Mitotic deacetylase associated SANT domain protein a	<i>D. rerio</i>	P5: 873GRKTK ⁸⁷⁷	No NLS predicted
Trithorax	<i>D. melanogaster</i>	P5: 1224GRKTK ¹²²⁸ 856GRKTK ⁸⁶⁰	No NLS predicted
NAC domain-containing protein 90	<i>G. max</i> <i>A. thaliana</i>	P5: 123GRKTK ¹²⁷ 125GRKTK ¹²⁹	No NLS predicted

Protein Wiz	<i>D. rerio</i>	P2P3: 275GKRTR ²⁷⁹ 269GKRTR ²⁷³	252RRKKR ²⁵⁶ 276KRTRGRPPGKGRNKT ²⁹⁰ 598RRKDVAFKSIGNKRGRGGRGRRRG ⁶²¹
Tetratricopeptide repeat protein 27	<i>M. musculus</i>	P2P3: 276GKRTR ²⁸⁰	No NLS predicted
Kinesin-like protein KIF22	<i>D. rerio</i>	P2P3: 363GKRTR ³⁶⁷	No NLS predicted
Meiotic recombination protein SPO11-1	<i>G. max</i>	P2P3: 3GKRTR ⁷	No NLS predicted
Transcription factor DIVARICA-TA	<i>G. max</i>	P2P3: 74GKRTR ⁷⁸	No NLS predicted
RNA polymerase III subunit C160	<i>D. melanogaster</i>	P2P3: 1236GKRTR ¹²⁴⁰	No NLS predicted
Recq-mediated instability protein (DUF1767)	<i>A. thaliana</i>	P2P3: 282GKRTR ²⁸⁶	No NLS predicted
DNA binding protein	<i>A. thaliana</i>	P2P3: 23GKRTR ²⁷	No NLS predicted
Small subunit processome component 20 homolog	<i>M. musculus</i>	P2P5: 2506GKGTK ²⁵¹⁰	1321KKKKN ¹³²⁵ 2505K GKKT KKK ²⁵¹² 2727KRALRKRRKAL ²⁷³⁷ 2744DIAAKKKLKKHKNKSEAKKRKIEFLRPGYKAKRQK ²⁷⁷⁸
Methyl-CPG-binding domain 9	<i>A. thaliana</i>	P2P5: 1779GKGTK ¹⁷⁸³	734RKAKKPKL ⁷⁴¹ 916RGRK ⁹¹⁹ 1353RRK ¹³⁵⁵ 1776KKR GKKT K ¹⁷⁸³ 2162RVRRSKGKKRKEP ²¹⁷³
High mobility group B4	<i>A. thaliana</i>	P2P5: 25GKGTK ²⁹	17LKTRGRKA GKKT KDPNQPKRP ³⁸
FACT complex subunit ssrp1-B	<i>C. elegans</i>	P2P5: 552GKGTK ⁵⁵⁶	525PKKKESKEKKNKREKKEKPVKEKAVKK GKKT KDPNEPK ⁵⁶²
Galactose oxidase/kelch repeat superfamily protein	<i>A. thaliana</i>	P2P5: 2GKGTK ⁶	3 KKT KPGKGKEKTERKTAKADEKKARREGKKL ³⁴ 353SKDKAK ³⁵⁸ 554KSLRRKEKRARI ⁵⁶⁵ 657RKGGIAKKKR ⁶⁶⁶
ATPase family AAA domain-containing protein 5	<i>D. rerio</i>	P2P5: 911GKGTK ⁹¹⁵	93KPVRGRGQKRTRTKDKKKSK ¹¹³ 182KKDGGNKKKSALRRNRKAK ¹⁹⁹ 443KKMKVVRKSGKKAQAKKA ⁴⁵⁹ 595AKKRKQAKKL VQAKALQQNKD TKKEKETVRRS ⁶²⁹ 897KRK ⁸⁹⁹ 941GRGRRGRSLRQK ⁹⁵² 1151PRRVVSSPRKPPQSPR ¹¹⁶⁶
Cat eye syndrome critical region protein 2 homolog	<i>M. musculus</i>	P2P5: 479GKGTK ³⁸³ 407GKGTK ⁴¹¹ 266GKGTK ²⁷⁰	195GKTGKRRGRPPKRKK ²⁰⁹ 358KREMEEKVKAVEDRAKRRKL ³⁷⁸ 537KREKRRSRSGR ⁵⁴⁷
Selenocysteine insertion sequence-binding protein 2-like	<i>M. musculus</i>	P2P5: 480GKGTK ⁴⁸⁴ 369GKGTK ³⁷³ 157GKGTK ¹⁶¹	133KGRRRR ¹³⁸ 333KSQKK ³³⁷ 444KAKKGKEKEIAKLKRPTALKKVILKEREKKGR ⁴⁷⁶
E3 ubiquitin-protein ligase UHRF2	<i>M. musculus</i>	P2P5: 648GKGTK ⁶⁵² 612GKGTK ⁶¹⁶ 424GKGTK ⁴²⁸ 388GKGTK ³⁹²	408KLSKKKAK ⁴¹⁵

Nucleolar protein 56	<i>G. max</i>	P2P5: 448GKKTK ⁴⁵²	449KKTKKKKQK ⁴⁵⁷ 479KSEKKKKKEK ⁴⁹⁰ 518KKKKNKKDSNGEAL EAGGESK ⁵⁴⁷
Protein DEK	<i>G. max</i>	P2P5: 183GKKTK ¹⁸⁷ 167GKKTK ¹⁷¹	91KEKEKAKGKGKRRRSK ¹²³ 272KGKKRTRKSFP ²⁸⁵
SAGA histone acetyltransferase complex subunit SPT7	<i>S. cerevisiae</i>	P2P5: 1064GKKTK ¹⁰⁶⁸	333KGKKKRS ³³⁹
E3 ubiquitin-protein ligase TRIM69	<i>M. musculus</i>	P2P5: 386GKKTK ³⁹⁰ 368GKKTK ³⁷²	No NLS predicted
PAZ domain-containing protein / piwi domain-containing protein	<i>A. thaliana</i>	P2P5: 166GKKTK ¹⁷⁰	No NLS predicted
AGAMOUS-like-56	<i>A. thaliana</i>	P2P5: 3GKKTK ⁷	No NLS predicted
NAC (No Apical Meristem) domain transcriptional regulator superfamily protein	<i>A. thaliana</i>	P2P5: 157GKKTK ¹⁶¹	No NLS predicted
CDKN2A-interacting protein	<i>D. rerio</i>	P2P5: 130GKKTK ¹³⁴ 129GKKTK ¹³³ 43GKKTK ⁴⁷	No NLS predicted
Fap1p	<i>S. cerevisiae</i>	P2P5: 559GKKTK ⁵⁶³	No NLS predicted
Serum response factor homolog	<i>C. elegans</i>	P2P5: 41GKKTK ⁴⁵	No NLS predicted
IQ domain-containing protein IQM2	<i>G. max</i>	P2P5: 325GKKTK ³²⁹	No NLS predicted
Probable WRKY transcription factor 32	<i>G. max</i>	P2P5: 315GKKTK ³¹⁹	No NLS predicted
PR domain zinc finger protein 10	<i>D. rerio</i>	P3P5: 425GRRTK ⁴²⁹	405RPKSRGRGRGRKRFGGARRPGRRT ⁴²⁸ 711RKDKLREHMQRMHNP EREAKKADRIHRTKA ⁷⁴⁰
Bromodomain testis-specific protein	<i>D. rerio</i>	P3P5: 264GRRTK ²⁶⁸	455KPKKKEKS ⁴⁹⁹ 581RKKTETNKNKPPKSKIKEKD ⁶⁰² 666TKRKQSKGPGHANKIKKAL ⁶⁸⁵
Xeroderma pigmentosum, complementation group C	<i>D. melanogaster</i>	P3P5: 1043GRRTK ¹⁰⁴⁷ 1042GRRTK ¹⁰⁴⁶	61KRGSDHKAPSGIKGSSVKKRKPTGQS ⁸⁶ 269KKAGRG ²⁷⁴ 295RIK ²⁹⁷ 542KPKKDKKAGKPAEKESKSTISKEAEKNNAKKAEAKPLSK ⁵⁸²
Transcription factor unc-3	<i>C. elegans</i>	P3P5: 244GRRTK ²⁴⁸	No NLS predicted
Pre-mRNA processing factor 31	<i>D. discoideum</i>	P2P3P5: 330GKRTK ³³⁴	327PEEGKRTKRGGKKARL ³⁴²
Crossover junction endonuclease MUS81	<i>G. max</i>	P2P3P5: 93GKRTK ⁹³	92KGKRTK ⁹⁷
Cyclin-D3-1-like	<i>G. max</i>	P3P5: 311GRRTK ³¹⁵	313RTKLKKR ³¹⁹
Toutatis	<i>D. melanogaster</i>	P2P3P5: 2611GKRTK ²⁶¹⁵ 2574GKRTK ²⁵⁷⁸ 2538GKRTK ²⁵⁴²	1088LARNKEKARQEKNSKLEQQRKDKE ¹¹¹¹ 2511KNKSSKKSSAKKQATPSKKQQQKNKKE ²⁵³⁹ 2612KRTKKKSGGKRRR ²⁶²⁴ 2746PPPKKRSAGGTSGSSSKRRDRDRESGGS AKRR ²⁷⁷⁸ 3011LKQEKKKEKHATK ³⁰²²
Restriction endonuclease, type	<i>A. thaliana</i>	P2P3P5:	339KSKEHVKNKSGKKRNS ³⁵⁴

II-like superfamily protein		267GKRTK271	
Fanconi anemia group E protein	M. musculus	P2P3P5: 193GKRTK197	No NLS predicted
Protein POLLENLESS 3-LIKE 2	G. max	P2P3P5: 155GKRTK159	No NLS predicted
Transcription factor bhlh91	G. max	P2P3P5: 193GKRTK197 180GKRTK184	No NLS predicted
Proliferation marker protein Ki-67		P5:	558GRK560
		1807GRKTK1811	627KRQRRP632
		1636GRKTK1640	1311RRPR1314
			1526RKPAKR1531
		P2P5:	1900RRPR1903
		2215GKKTK2219	2137KRRPR2141
		2071GKKTK2075	
		1895GKKTK1899	
		1388GKKTK1392	
		1244GKKTK1248	
		1147GKKTK1151	
		1124GKKTK1128	
		1068GKKTK1072	2725RPRRV2729
			2965PVPEKKRAASSKRH2978
		P2P3P5:	
		2215GKRTK2219	
		1388GKRTK1392	
		1147GKRTK1151	
		1003GKRTK1007	
		827GKRTK831	
		P2P5:	
		265GKKTK269	
C2H2-like zinc finger protein	A. thaliana G. max	P2P3P5: 24GKRTK28 20GKRTK24 18GKRTK22 17GKRTK21	No NLS predicted

¹ All listed proteins are known to function in the cell nucleus

² NLSs were predicted using NLStradamus (Nguyen et al. 2009; BMC Bioinformatics. 10:202). Overlapping regions of the predicted NLSs containing GRKTR or its variants are highlighted in red.

Table S6. BLASTP E-values for nuclear proteins from different cellular organisms reported in Figure 5.¹

Variant	<i>M. musculus</i>	<i>H. sapiens</i>	<i>D. rerio</i>	<i>G. max</i>	<i>A. thaliana</i>	<i>D. melano-</i> <i>gaster</i>	<i>C. elegans</i>	<i>S. cerevisiae</i>
WT	341.2 ± 0.1 (p < 0.001)	342.6 ± 0.6 (p < 0.001)	343.0	343.0 ± 0.0 (p < 0.001)	343.0 ± 0.0 (p < 0.001)	343.0	No hit	No hit
P2	485.5 ± 0.2 (p < 0.001)	489.5 ± 0.3 (p < 0.001)	488.25 ± 0.8 (p < 0.001)	490.0 ± 1.3 (p < 0.001)	491.7 ± 0.9 (p < 0.001)	487.6 ± 0.43	489.0	489.0
P3	242.2 ± 0.5 (p < 0.001)	242.5 ± 0.2 (p < 0.001)	242.0 (p < 0.001)	242.5 ± 0.5 (p < 0.001)	242.0 (p < 0.001)	243.0	241.0 ± 0.0 (p < 0.001)	No hit
P5	487.2 ± 0.6 (p < 0.001)	487.0 ± 0.0 (p < 0.001)	484.3 ± 0.1 (p < 0.001)	489.7 ± 0.8 (p < 0.001)	488.1 ± 1.4 (p < 0.001)	484.0 ± 0.0	No hit	486.5 ± 0.50
P2P3	341.3 ± 0.1 (p < 0.001)	342.2 ± 0.3 (p < 0.001)	341.3 ± 0.2 (p < 0.001)	345.8 ± 1.6 (p < 0.001)	344.0 ± 0.3 (p < 0.001)	342.0	No hit	No hit
P2P5	691.4 ± 0.9 (p < 0.001)	689.9 ± 0.57 (p < 0.001)	691.4 ± 1.0 (p < 0.001)	690.6 ± 0.2 (p < 0.001)	698.3 ± 2.3 (p < 0.001)	No hit	692.0 ± 2.0 (p < 0.001)	688.5 ± 0.50
P3P5	No hit	343.5 ± 0.5 (p < 0.001)	342.0 ± 0.0 (p < 0.001)	No hit	No hit	342.0 ± 0.0	344.0 (p < 0.001)	No hit
P2P3P5	489.3 ± 1.5 (p < 0.001)	No hit	No hit	488.2 ± 0.42 (p < 0.001)	487.5 ± 0.7 (p < 0.001)	484.0 ± 0.0	No hit	No hit
Random peptide (GLVTV)	751.9 ± 1.8	755.7 ± 2.0	753.2 ± 0.5	752.0 ± 0.0	753.0	No hit	757.4 ± 3.1	No hit
SV40 Large T-antigen NLS	No hit	No hit	No hit	No hit	No hit	No hit	No hit	No hit

¹ Mean of E-values ± standard error of the mean are listed. E-values for each cellular organism were compared with that of the corresponding random peptide by One-way ANOVA followed by Tukey's test using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA).