

SUPPLEMENTARY INFORMATION TO

Manuscript IAI#00301-17

**H-IPSE is a pathogen-secreted host nucleus infiltrating protein (infiltrin)
expressed exclusively by the *Schistosoma haematobium* egg stage**

Luke F. Pennington^{a§}, Abdulaziz Alouffi^{b§}, Evaristus C. Mbanefo^c, Debalina Ray^d, David
M. Heery^e, Ted Jardetzky^a, Michael H. Hsieh^{c,f} and F. H. Falcone^b

^aStanford University School of Medicine, Stanford, CA, USA;

^bSchool of Pharmacy, Division of Molecular Therapeutics and Formulation, University of Nottingham, Nottingham, UK;

^cBiomedical Research Institute, Rockville, MD, USA

^dDepartment of Pathology, University of California, San Francisco, CA, USA

^eSchool of Pharmacy, Division of Biomolecular Science and Medicinal Chemistry, University of Nottingham, Nottingham, UK

^fDepartment of Urology, The George Washington University, Washington, D.C., USA

[§] Equal contributors

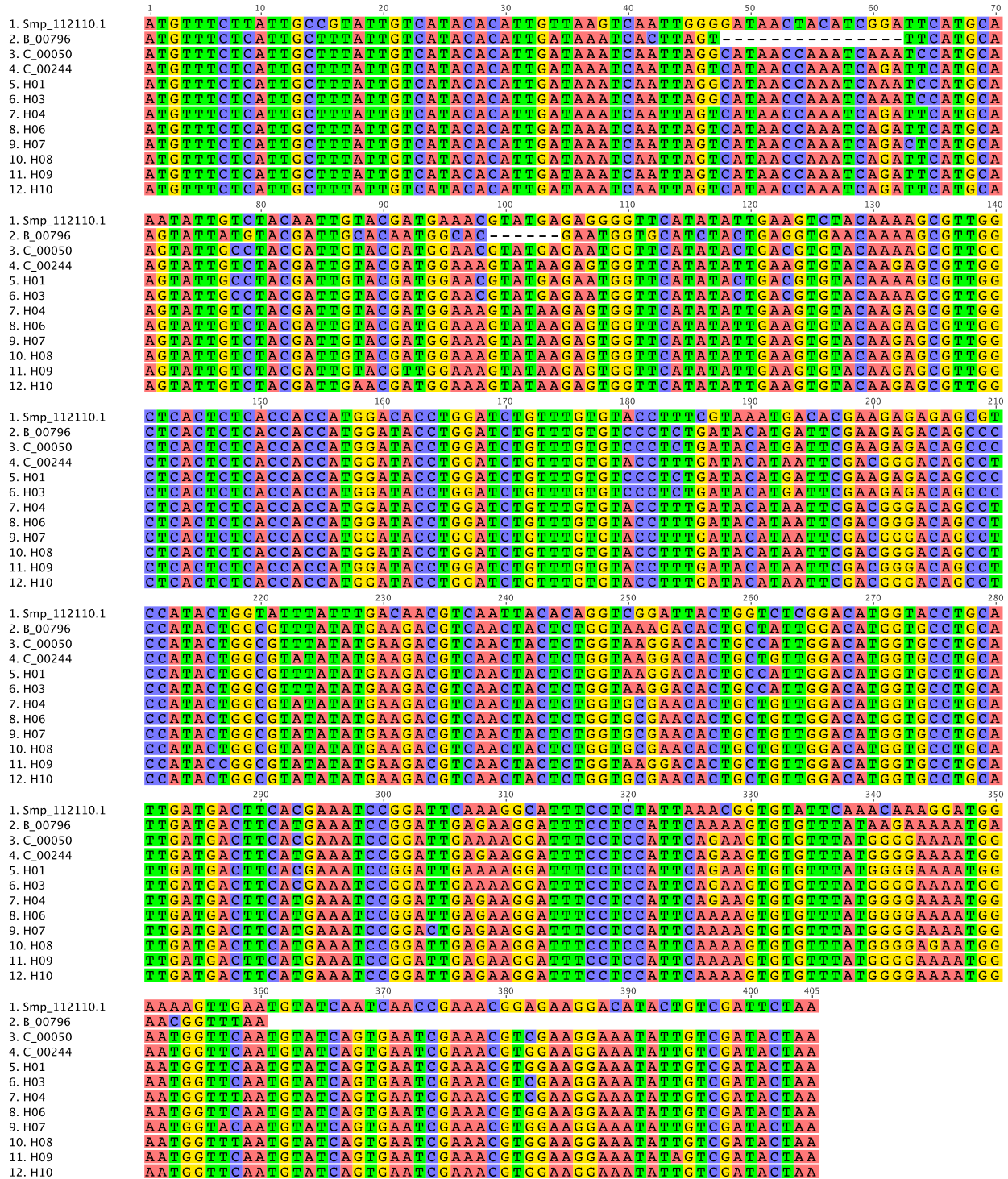


Figure S1: A Geneious global alignment with free end gaps and default alignment parameters (65% similarity (5.0/-4.0) cost matrix, 12 gap open penalty, 3 gap extension penalty, 2 refinement iterations) was performed using indicated sequences (Geneious v7.1.4).

	Smp_112110.1	B_00796	C_00050	C_00244	H01	H03	H04	H06	H07	H08	H09	H10
Smp_112110.1		73.9	80.7	80.7	80.7	80.7	81	81.2	80.5	81.5	80.2	81
B_00796	73.9		87.5	85.8	87.5	87.5	85.3	85.3	84.4	85.3	85.6	85.3
C_00050	80.7	87.5		94.6	99.8	100	93.8	93.6	92.8	93.1	93.6	93.3
C_00244	80.7	85.8	94.6		94.8	94.6	98.8	99	98.3	98.5	99	98.8
H01	80.7	87.5	99.8	94.8		99.8	93.6	93.8	93.1	93.3	93.8	93.6
H03	80.7	87.5	100	94.6	99.8		93.8	93.6	92.8	93.1	93.6	93.3
H04	81	85.3	93.8	98.8	93.6	93.8		99.3	98.5	99.3	97.8	99
H06	81.2	85.3	93.6	99	93.8	93.6	99.3		99.3	99.5	98.5	99.8
H07	80.5	84.4	92.8	98.3	93.1	92.8	98.5	99.3		98.8	97.8	99
H08	81.5	85.3	93.1	98.5	93.3	93.1	99.3	99.5	98.8		98	99.3
H09	80.2	85.6	93.6	99	93.8	93.6	97.8	98.5	97.8	98		98.3
H10	81	85.3	93.3	98.8	93.6	93.3	99	99.8	99	99.3	98.3	

Figure S2: A matrix showing the relative nucleotide identity of the indicated sequences.

>Smp_112110.1

```
ATTGCCGTATTGTCATACACATTGTTAAGTCAATTGGGGATAACTACATCGGATTCATGCAAATATTGTCTACAATTG
TACGATGAAACGTATGAGAGGGGTTTCATATATTGAAGTCTACAAAAGCGTTGGCTCACTCTCACCACCATGGACACCT
GGATCTGTTTGTGTACCTTTTCGTAATGACACGAAGAGAGAGCGTCCATACTGGTATTTATTTGACAACGTCAATTAC
ACAGGTCGGATTACTGGTCTCGGACATGGTACCTGCATTGATGACTTCACGAAATCCGGATTCAAAGGCATTTCTCT
ATTAACGGTGTATTCAAACAAAGGATGGAAAAGTTGAATGTATCAATCAACCGAAACGGAGAAGGACATACTGTCTGA
TTCTAA
```

>C_00244

```
ATTGCTTTTATTGTCATACACATTGATAAATCAATTAGTCATAACCAAATCAGATTCATGCAAGTATTGTCTACGATTG
TACGATGGAAGTATAAGAGTGGTTTCATATATTGAAGTGTACAAGAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTACCTTTGATACATAATTCGACGGGACAGCCTCCATACTGGCGTATATATGAAGACGTCAACTAC
TCTGGTAAGGACACTGCTGTTGGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGATTGAGAAGGATTTCTCTCC
ATTCAGAAGTGTGTTTATGGGGAAAATGGAATGGTTCAATGTATCAGTGAATCGAAACGTGGAAGGAAATATTGTCTGA
TACTAA
```

>C_00050

```
ATTGCTTTTATTGTCATACACATTGATAAATCAATTAGGCATAACCAAATCAAATCCATGCAAGTATTGCCTACGATTG
TACGATGGAACGTATGAGAATGGTTTCATATACTGACGTGTACAAAAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTCCCTCTGATACATGATTCGAAGAGACAGCCCCATACTGGCGTTTATATGAAGACGTCAACTAC
TCTGGTAAGGACACTGCCATTGGACATGGTGCCTGCATTGATGACTTCACGAAATCCGGATTGAAAAGGATTTCTCTCC
ATTCAGAAGTGTGTTTATGGGGAAAATGGAATGGTTCAATGTATCAGTGAATCGAAACGTGGAAGGAAATATTGTCTGA
TACTAA
```

>B_00796

```
ATTGCTTTTATTGTCATACACATTGATAAATCACTTAGTTTTATGCAAGTATTATGTACGATTGCACAATGGCACGAAT
GGTGCATCTACTGAGGTGAACAAAAGCGTTGGCTCACTCTCACCACCATGGATACCTGGATCTGTTTGTGTCCCTCTG
ATACATGATTCGAAGAGACAGCCCCATACTGGCGTTTATATGAAGACGTCAACTACTCTGGTAAAGACACTGCTATT
GGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGATTGAGAAGGATTTCTCTCATTCAAAGTGTGTTTATAAG
AAAATGAAACGGTTTTAA
```

>H1

```
ATTGCTTTTATTGTCATACACATTGATAAATCAATTAGGCATAACCAAATCAAATCCATGCAAGTATTGCCTACGATTG
TACGATGGAACGTATGAGAATGGTTTCATATACTGACGTGTACAAAAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTCCCTCTGATACATGATTCGAAGAGACAGCCCCATACTGGCGTTTATATGAAGACGTCAACTAC
TCTGGTAAGGACACTGCCATTGGACATGGTGCCTGCATTGATGACTTCACGAAATCCGGATTGAAAAGGATTTCTCTCC
ATTCAGAAGTGTGTTTATGGGGAAAATGGAATGGTTCAATGTATCAGTGAATCGAAACGTGGAAGGAAATATTGTCTGA
TACTAA
```

>H3

ATTGCTTTATTGTCATACACATTGATAAATCAATTAGGCATAACCAAATCAAATCCATGCAAGTATTGCCTACGATTG
TACGATGGAACGTATGAGAATGGTTCATATACTGACGTGTACAAAAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTCCCTCTGATACATGATTGGAAGAGACAGCCCCATACTGGCGTTTATATGAAGACGTCAACTAC
TCTGGTAAGGACACTGCCATTGGACATGGTGCCTGCATTGATGACTTCACGAAATCCGGATTGAAAAGGATTTCTCTCC
ATTCAGAAGTGTGTTTATGGGGAAAATGGAATGGTTC AATGTATCAGTGAATCGAAACGTGCAAGGAAATATTGTCTCGA
TACTAA

>H4

ATTGCTTTATTGTCATACACATTGATAAATCAATTAGTCATAACCAAATCAGATTTCATGCAAGTATTGTCTACGATTG
TACGATGGAAGTATAAGAGTGGTTCATATATTGAAGTGTACAAGAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTACCTTTGATACATAAATTCGACGGGACAGCCTCCATACTGGCGTATATATGAAGACGTCAACTAC
TCTGGTGCGAACACTGCTGTTGGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGATTGAGAAGGATTTCTCTCC
ATTCAGAAGTGTGTTTATGGGGAAAATGGAATGGTTC AATGTATCAGTGAATCGAAACGTGCAAGGAAATATTGTCTCGA
TACTAA

>H6

ATTGCTTTATTGTCATACACATTGATAAATCAATTAGTCATAACCAAATCAGATTTCATGCAAGTATTGTCTACGATTG
TACGATGGAAGTATAAGAGTGGTTCATATATTGAAGTGTACAAGAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTACCTTTGATACATAAATTCGACGGGACAGCCTCCATACTGGCGTATATATGAAGACGTCAACTAC
TCTGGTGCGAACACTGCTGTTGGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGATTGAGAAGGATTTCTCTCC
ATTCAAAAGTGTGTTTATGGGGAAAATGGAATGGTTC AATGTATCAGTGAATCGAAACGTGGAAGGAAATATTGTCTCGA
TACTAA

>H7

ATTGCTTTATTGTCATACACATTGATAAATCAATTAGTCATAACCAAATCAGACTCATGCAAGTATTGTCTACGATTG
TACGATGGAAGTATAAGAGTGGTTCATATATTGAAGTGTACAAGAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTACCTTTGATACATAAATTCGACGGGACAGCCTCCATACTGGCGTATATATGAAGACGTCAACTAC
TCTGGTGCGAACACTGCTGTTGGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGACTGAGAAGGATTTCTCTCC
ATTCAAAAGTGTGTTTATGGGGAAAATGGAATGGTTC AATGTATCAGTGAATCGAAACGTGGAAGGAAATATTGTCTCGA
TACTAA

>H8

ATTGCTTTATTGTCATACACATTGATAAATCAATTAGTCATAACCAAATCAGATTTCATGCAAGTATTGTCTACGATTG
TACGATGGAAGTATAAGAGTGGTTCATATATTGAAGTGTACAAGAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTACCTTTGATACATAAATTCGACGGGACAGCCTCCATACTGGCGTATATATGAAGACGTCAACTAC
TCTGGTGCGAACACTGCTGTTGGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGATTGAGAAGGATTTCTCTCC
ATTCAAAAGTGTGTTTATGGGGAGAATGGAATGGTTC AATGTATCAGTGAATCGAAACGTGGAAGGAAATATTGTCTCGA
TACTAA

>H9

ATTGCTTTATTGTCATACACATTGATAAATCAATTAGTCATAACCAAATCAGATTTCATGCAAGTATTGTCTACGATTG
TACGTTGGAAAGTATAAGAGTGGTTCATATATTGAAGTGTACAAGAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTACCTTTGATACATAAATTCGACGGGACAGCCTCCATACCGGCGTATATATGAAGACGTCAACTAC
TCTGGTAAGGACACTGCTGTTGGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGATTGAGAAGGATTTCTCTCC
ATTCAAAAGTGTGTTTATGGGGAAAATGGAATGGTTC AATGTATCAGTGAATCGAAACGTGGAAGGAAATATAGTCTCGA
TACTAA

>H10

ATTGCTTTATTGTCATACACATTGATAAATCAATTAGTCATAACCAAATCAGATTTCATGCAAGTATTGTCTACGATTG
AACGATGGAAGTATAAGAGTGGTTCATATATTGAAGTGTACAAGAGCGTTGGCTCACTCTCACCACCATGGATACCT
GGATCTGTTTGTGTACCTTTGATACATAAATTCGACGGGACAGCCTCCATACTGGCGTATATATGAAGACGTCAACTAC
TCTGGTGCGAACACTGCTGTTGGACATGGTGCCTGCATTGATGACTTCATGAAATCCGGATTGAGAAGGATTTCTCTCC
ATTCAAAAGTGTGTTTATGGGGAAAATGGAATGGTTC AATGTATCAGTGAATCGAAACGTGGAAGGAAATATTGTCTCGA
TACTAA

Figure S3: FASTA-formatted sequences used for alignments

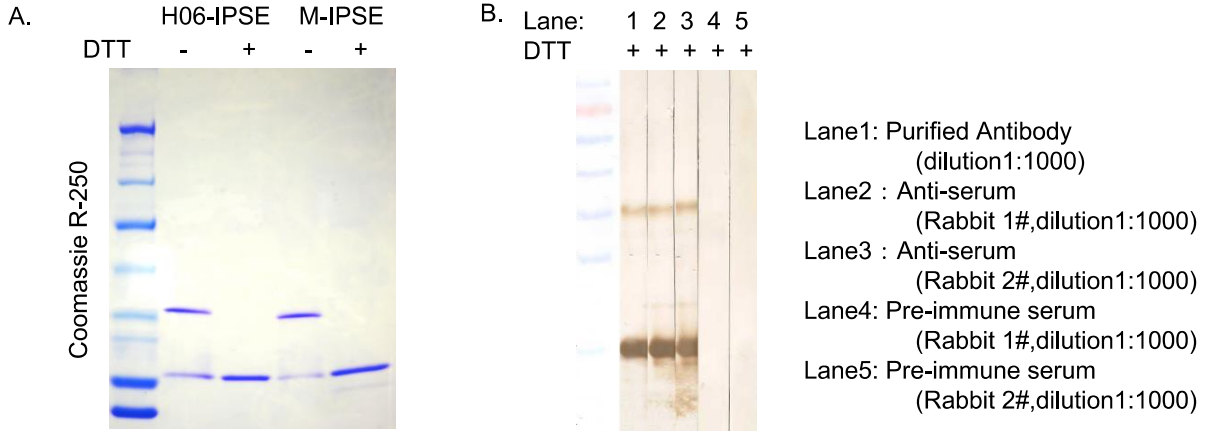


Figure S4: Generation of polyclonal anti-H06-IPSE antibodies: A.) 4-20% gradient SDS-PAGE electrophoresis of refolded H06-IPSE and M-IPSE stained with Coomassie R250. H06-IPSE protein was supplied as immunogen to ProMab Biotechnologies for the generation of anti-IPSE polyclonal rabbit antibodies. B.) A western blot H06-IPSE under reducing conditions with anti-serum and purified antibody from immunized rabbits before and after IPSE injection.

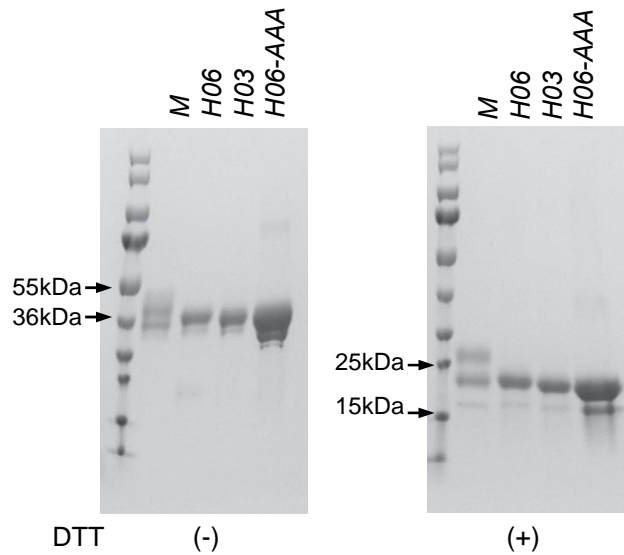


Figure S5: Mammalian derived IPSE: 4-20% gradient SDS-PAGE electrophoresis of IPSE variants expressed in HEK-6E cells and purified with Ni-NTA or TALON affinity resins stained with Coomassie R250. Proteins exhibited the expected homodimeric (DTT negative) and monomeric (DTT positive) molecular weights corresponding to differential glycosylation at the multiple potential N-linked glycosylation sites.

Table T1

Mutant/wt	Purpose Sequence (5'→3')
	<i>Construction pTT5 expression constructs</i>
M-IPSE for	GGCGGAGAGAACCTGTATTTTCAGGGCGGATCCGAC AGCTGCAAGTACTGCCTGCAG
M-IPSE rev	AGCGGCCGCGGTACCAAGCTAGTCATTACTIONAGAAC GGCAGTAGGTCTCCGCCGCTTGG
H03-IPSE for	GGCGGAGAGAACCTGTATTTTCAGGGCGGATCCAAC CCCTGCAAGTACTGCCTGCG
H03-IPSE rev	GCGGCCGCGGTACCAAGCTAGTCATTACTIONAGTACCG GCAGTACTTTCTCCGCCGCTTGC
H06-IPSE for	GGCGGAGAGAACCTGTATTTTCAGGGCGGATCCAAG AGCGACAGCTGCAAGTACTG
H06-IPSE rev	AGCGGCCGCGGTACCAAGCTAGTCATTACTIONAGTACC GGCAGTACTTTCTGCCCGCTTGC
VEGF-SS-8xHIS-TEV for VEGF-SS-8xHIS-TEV rev	CCAGGTCCAAGTTTAAACGGATCTCTAGCG GGATCCGCCCTGAAAATACAGGTTCTCTCCGCC
H06-SK AA AKY for	GAGAGCAAGGCCGCTGCCAAGTACTGCCGGTACTAG TAATGACTAGC
H06-SK AA AKY rev	GCAGTACTTGGCAGCGGCCTTGCTCTCGCTGATGCA CTGCACCATGCC
	<i>Cloning IPSE transcripts</i>
H-IPSE for H-IPSE rev H-IPSE RACE for.1 H-IPSE RACE (nested) for.2 Anchor Insert 3' RACE rev.1 Anchor Insert 3' RACE (nested) rev.2	ATGTTTCTCATTGCTTTATTGTCATACAC ACACAAACAGATCCAGGTATCC CGATGTTTCTCATTGCTTTATTGTC GTTTCTCATTGCTTTATTGTCATAC GTCATCGGCTTAGCATTG CTTAGCATTGATGTCTCGC
	<i>Tetra-EGFP- NLS mutagenic oligonucleotides</i>
PKKKRKV-For (SV40 pos control) PKKKRKV -Rev	* <u>GATCT</u> CCGAAGAAGAAGAGGAAGGTAA * <u>GATCT</u> TACCTTCCTCTTCTTCTCGGA
SKRGRKY-For (H06-IPSE WT) SKRGRKY-Rev	* <u>GATCT</u> AGCAAGAGGGGAAGGAAGTACA * <u>GATCT</u> GTACTIONCTTCCCCTCTTGCTA
SKR RR KY-For (H03-IPSE WT) SKR RR KY-Rev	* <u>GATCT</u> AGCAAGAGGAGG AGG AAGTACA * <u>GATCT</u> GTACTIONCTT CCT CTTGCTA
S AA GAAY-For (H06 NLS mut) S AA GAAY-Rev	* <u>GATCT</u> AGCG CAGCA GAG AGCAGCA TACA * <u>GATCT</u> GTA TGCTGCTCCTGCTGCGCTA

Table S1. Oligonucleotide sequences used in this study. Key: Bold, underlined: restriction sites; Bold, blue: mutation introduced. * denotes 5'-phosphorylation.