Supplementary Information

Pairing of single mutations yields obligate Cre-type site-specific recombinases

Jenna Hoersten[1], Gloria Ruiz-Gómez[2], Felix Lansing[1], Teresa Rojo-Romanos[1], Lukas Theo Schmitt[1], Jan Sonntag[1], M. Teresa Pisabarro[2] and Frank Buchholz[1,*]

[1] Medical Faculty and University Hospital Carl Gustav Carus, UCC Section Medical Systems Biology, TU Dresden, 01307 Dresden, Germany

[2] Structural Bioinformatics, BIOTEC TU Dresden, Tatzberg 47-51, 01307, Dresden, Germany

[*]To whom correspondence should be addressed. Tel: +49 351 463 40277; Email: frank.buchholz@tu-dresden.de

Supplementary Figure S1: Plasmid map for assay to visualize recombination activity.

Supplementary Figure S2: Blue-white library screen to identify SSR clones with desired properties.

Supplementary Figure S3: Library construction, evolution and selection scheme.

Supplementary Figure S4: Gating strategy to evaluate recombination efficiencies of SSRs for the HEK293T reporter cell line.

Supplementary Figure S5: MD-based structural analysis of Cre/loxP and the investigated mutants.
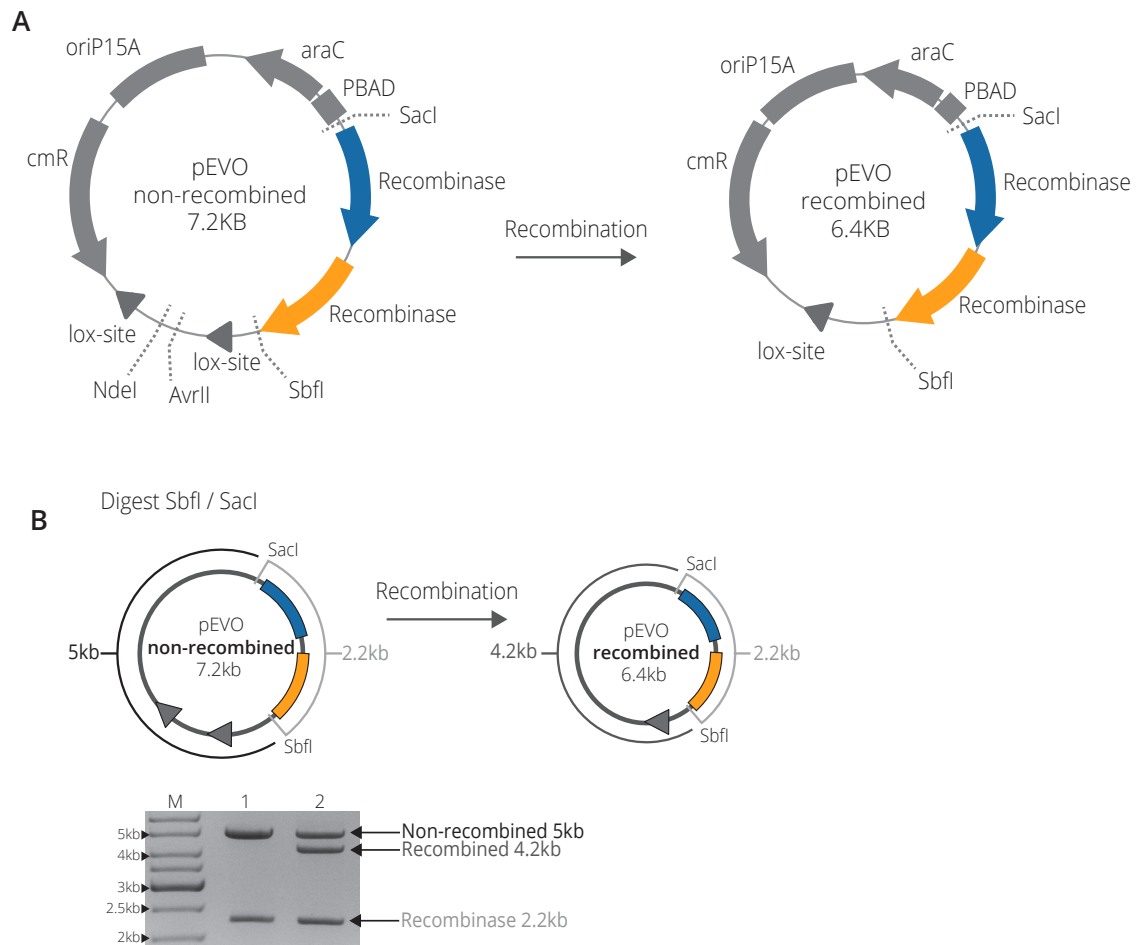
Supplementary Figure S6: Half-site combinations of the loxF8 target site.

**Supplementary Materials & Methods:** MD simulations protocol

Supplementary Table S1: Primers

Supplementary Table S2: Sanger sequencing results of colony screen.

**Supplementary Table S3:** Complete sequencing results of colony screen.

Supplementary Figure S1: Plasmid map for assay to visualize recombination activity. A | The vectors are shown before and after SSR-mediated recombination of a 750bp fragment flanked by the lox-sites (gray triangles). Protein coding genes are depicted as arrows, the origin of replication (oriP15A) and the pBAD promoter are shown as rectangles. The protein coding genes include the chloramphenicol resistance gene (cmR), the arabinose regulatory protein (araC) and the genes encoding for the recombinase(s) or libraries of interest. Relevant restriction enzyme sites are shown as dotted lines. B | Bacterial recombination assay. Purified plasmids containing a mixture of recombined and non-recombined target sequences are digested to visualize the amount of recombined vs. non-recombined plasmids in the sample. By digesting the sample with SacI, located upstream of the first recombinase sequence (blue) and SbfI, located downstream of the second recombinase sequence (orange). If non-recombined, the plasmid is cut into fragments of size 2.2kb and 5kb, if recombined the plasmid is cut into fragments of size 2.2kb and 4.2kb. Gel indicates an example digestion where well 1 (1) contains a digested sample of only non-recombined plasmids and well 2 (2) contains a mixture of both recombined and non-recombined plasmids. Sizes are indicated with arrows. M = GeneRuler DNA Ladder Mix 10kb.
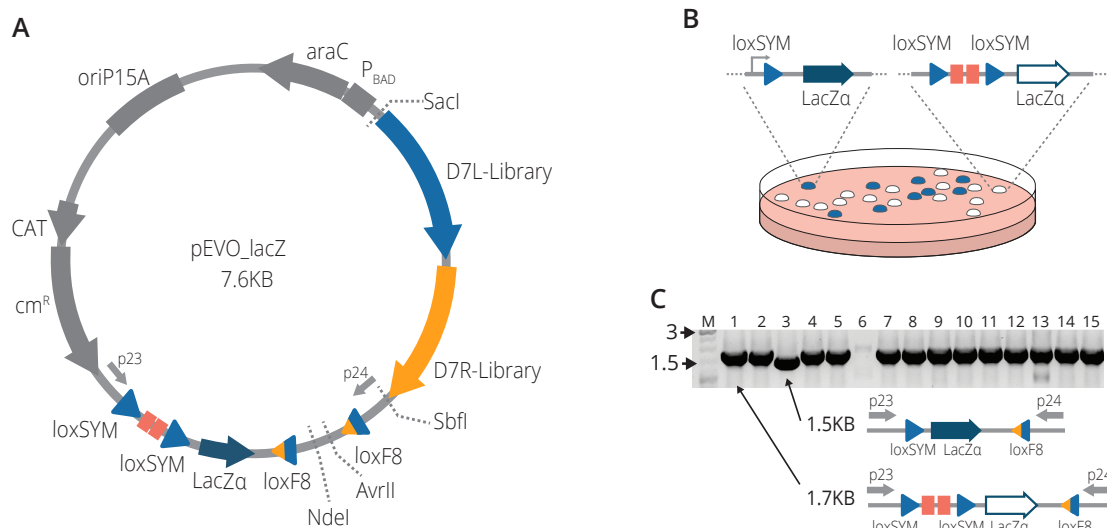
**Supplementary Figure S2 | Blue-white library screen to identify SSR clones with desired properties.**
A | Plasmid map of the employed vector. Relevant features of the vectors are shown (see Supplementary Figure 1). The selection plasmid contains transcriptional terminators (red squares) flanked by the symmetric sites (indicated as loxSYM) upstream of LacZα. Upon recombination of the symmetric sites, the transcriptional terminators are excised allowing for the expression of LacZα driven by the constitutive cat promoter (CAT). Inactive variants are removed via NdeI and AvrII digest. B | Scheme of blue-white screen. The blue colonies contain SSR mutants that did not recombine the symmetric sites. The white colonies are selected, containing mutants that did not recombine the symmetric sites, whereas they had recombined the loxF8 sites. C | Example displaying 15 of the 75 amplicon products from the colony PCR. Primers p23 and p24 were used to amplify over the target sites. The resulting amplicon contains a band of 1.5KB if the loxSYM site is recombined and a 1.7KB product if the site is non recombined. M = 10kb ladder (GeneRuler DNA Ladder Mix 10KB).
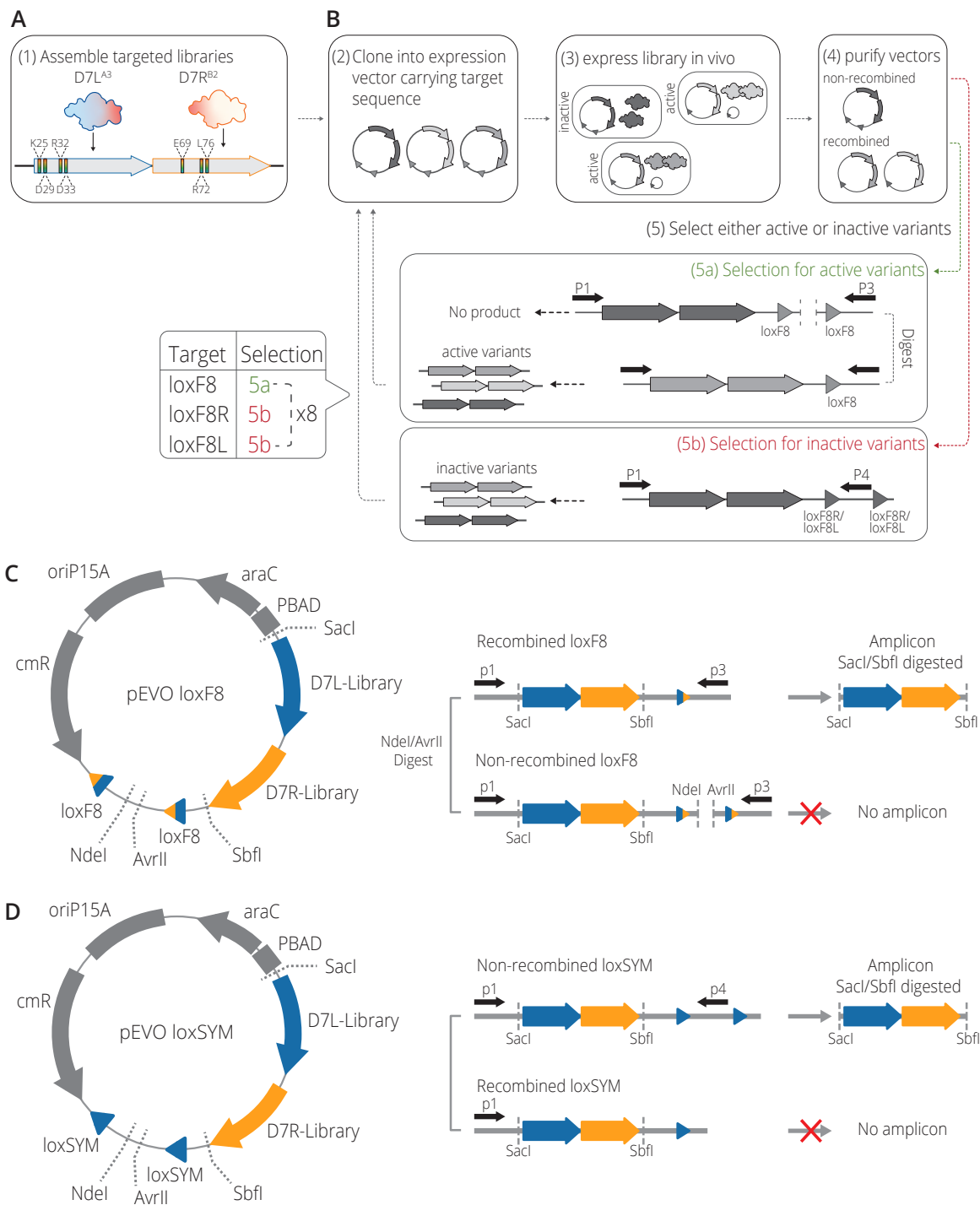
Supplementary Figure S3 | Library construction, evolution and selection scheme. A | (1) D7L recombinase (blue) and D7R recombinase (orange) with applied targeted positions indicating the original amino acid and position (rainbow). B | Evolution and selection scheme (2) the constructed libraries are cloned together into the target substrate vector. (3) The constructed vector is transformed into XL1-Blue E. coli and expression of both recombinase libraries is induced. (4) The resulting recombined and non-recombined vectors are purified from the bacteria. (5) Selection of variant pairs that have desired activity, either active on the loxF8 site (5a) or inactive on the loxF8L and loxF8R sites (5b). (5a) Variants active on loxF8 are selected by first digesting the sample with restriction enzymes of sites between the two full loxF8 sites to remove the non-active variants then selecting and evolving the active variants with an error-prone PCR. Primers bind upstream of the recombinase variants (p1) and downstream of the loxF8 site

(p4). The active variants are then carried on through the next cycle. (5b) Selection for variants inactive on the loxF8L or loxF8R sites are isolated via PCR with a primer that binds upstream of the recombinase pair (p1) and between the two full lox sites. C | Plasmid maps for SLiDE strategy to evolve obligate D7 recombinases. Relevant features of the vectors are shown (see Supplementary Figure 1). Selection scheme for substrate-linked directed evolution to screen for recombinase pairs that recombine loxF8.The variants that are not active on the loxF8 target site are removed from the library by digesting the purified plasmid with NdeI/AvrII, linearizing any plasmid that has not undergone recombination. The active variants are then amplified with primers p1 and p3 By PCR to carry the variants to the next cycle of evolution. D | Selection scheme to remove variants capable of recombining the symmetric sites loxF8L or loxF8R (indicated as loxSYM with blue triangles). The variants are amplified with primer p1 and p4 that binds between the symmetric sites, amplifying only those SSRs that did not recombine the target site. SSRs without activity were preferentially amplified by PCR to bring them to the next cycle of evolution.
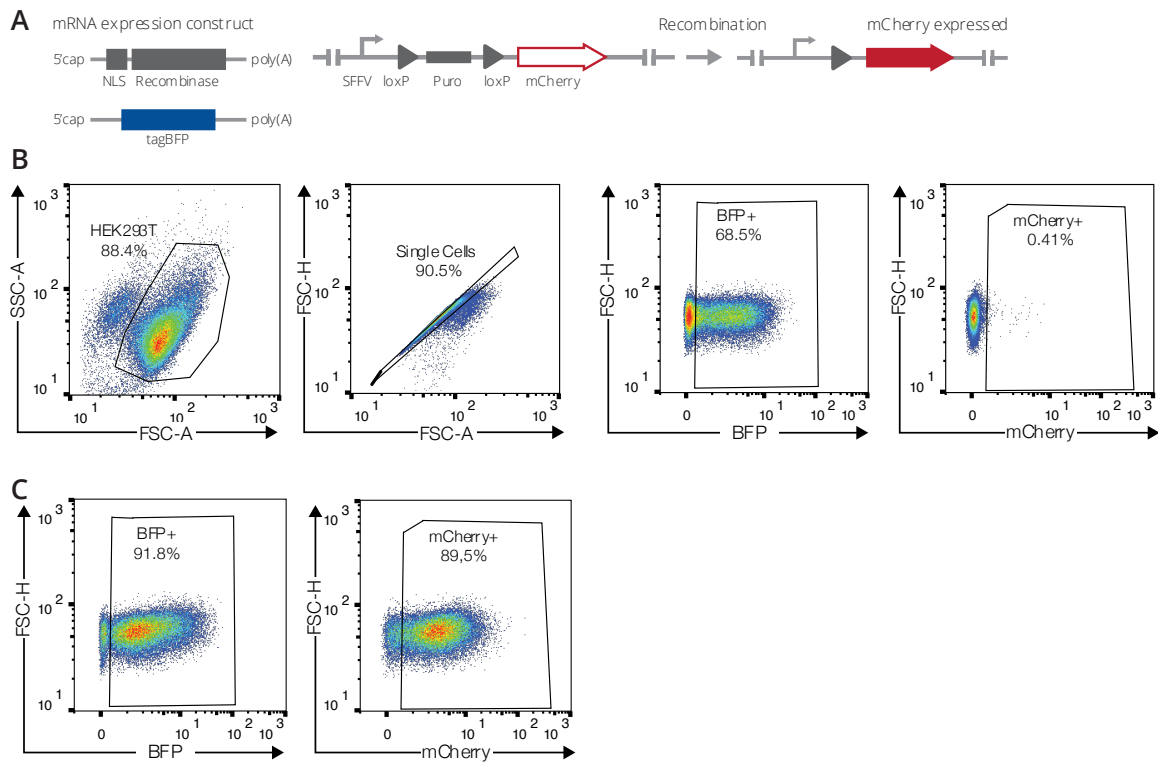
Supplementary Figure S4 | Gating strategy to evaluate recombination efficiencies of SSRs for the HEK293T reporter cell line. A | Schematic presentation of mRNA expression constructs and the reporter HEK293T cell line. Employed mRNAs with indicated features (5'cap and polyA tail) expressing a nuclear localization signal (NLS) fused to the recombinase and the tagBFP mRNA are shown. The stable reporter cell line harbours two loxP sites (gray triangles) that flank a puromycin selection gene (puro). Once successfully excised by recombination, mCherry is expressed from the SFFV promoter (arrow). B | Single cells were gated out for live and dead cells. The live cells were gated for BFP+ and mCherry+ cells. C | FACS plots of representative mRNA transfected HEK293T reporter cell line where BFP fluorescence indicates the cells transfected and mCherry fluorescence indicates the gated BFP+ cells with a recombination event.
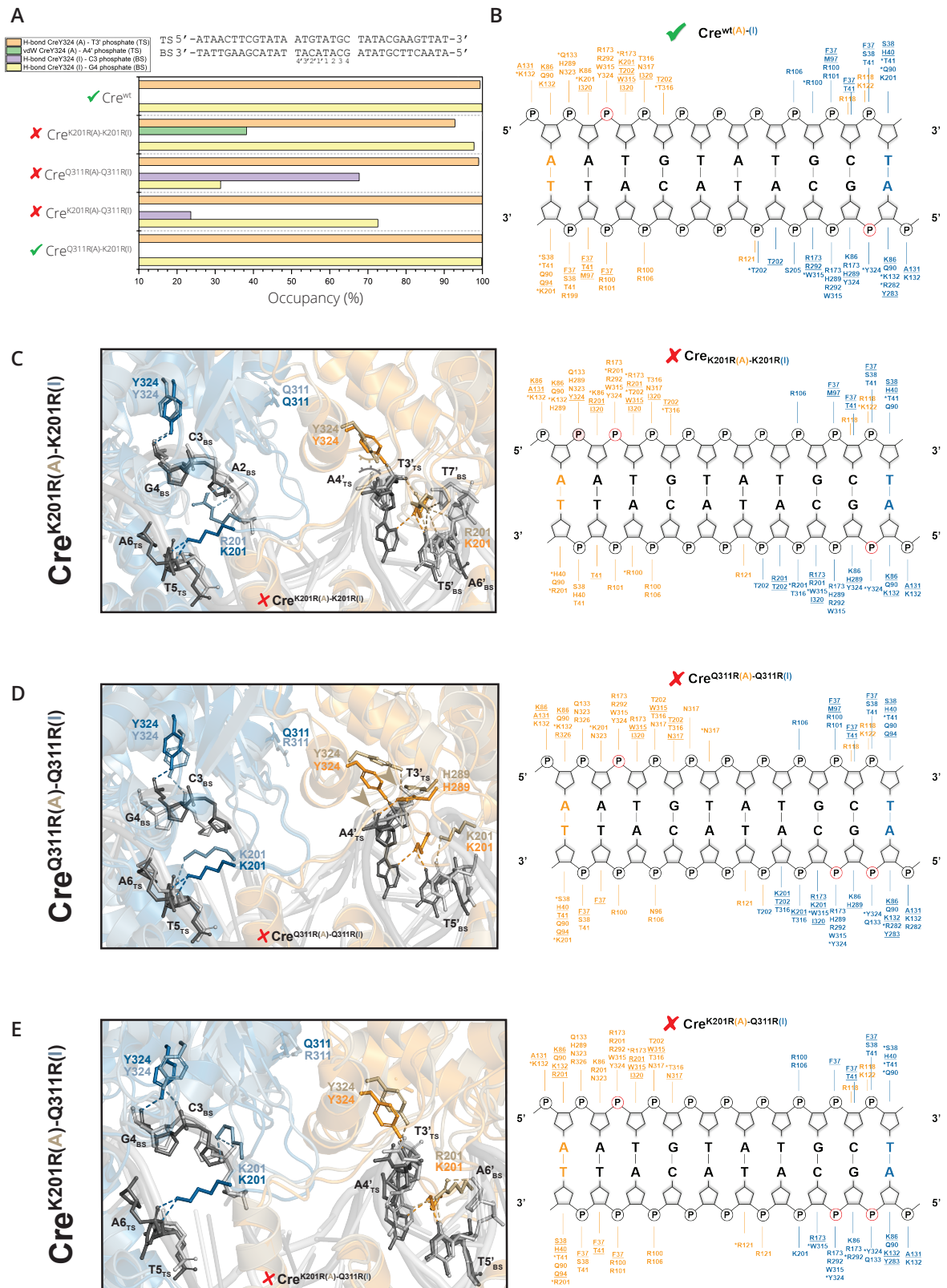
Supplementary Figure S5 | MD-based structural analysis of Cre/loxP and the investigated mutants.
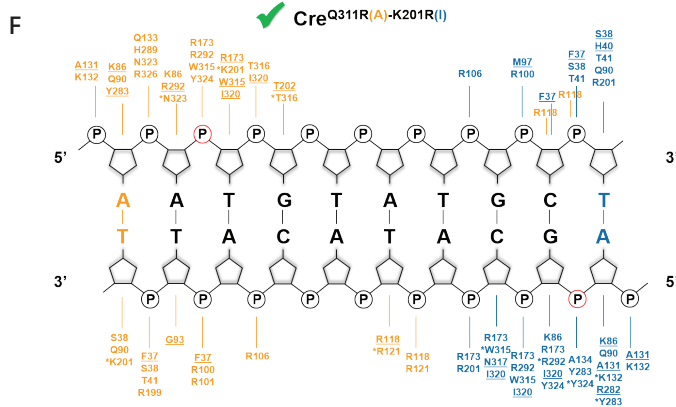
Supplementary Figure S5 | MD-based structural analysis of Cre/loxP and the investigated mutants. A | The top of the panel shows the DNA loxP sequence with top (TS) and bottom (BS) strands labelled, and the bases in the spacer region numbered at the bottom. The graph shows a summary of the intermolecular hydrogen bond (H-bond) and van der Waals (vdW) interactions established by the catalytic residue Y324 with the DNA in the last 100 ns of MD simulations for Cre wild type and the studied mutants. Interactions are shown in percentage of occupancy (*i.e.* appearance) with the following color code (summarized in the explanatory caption box at the top left): in orange the appearance (in % occupancy) of H-bond between Y324 and T3'phosphate from TS, in green the appearance of vdW interactions between Y324 and A4'phosphate from TS, in violet the appearance of H-bond between Y325 and C3 phosphate in BS, and in yellow the appearance of H-bond between Y325 and G4 phosphate in BS. The green check mark or red cross at the left of the enzyme names indicate if it is active or inactive, respectively. B | Intermolecular hydrogen bond and van der Waals contacts profile obtained from the analysis of the last 100 ns of MD simulations for Cre[wt]. The DNA is represented schematically with phosphates as circles and bases with the corresponding one-letter code. Enzyme residues establishing contacts with the DNA are listed in orange and blue for the active (A) and inactive (I) monomer, respectively. Bold is used to highlight residues forming H-bond, underlined and bold is used for van der Waals contacts, and residues involved in H-bond and van der Waals contacts simultaneously are highlighted in bold with an asterisk. Phosphate groups involved in Y324 recognition are highlighted with a red circle in the DNA scheme. C | Left panel: details of the superimposition of the MD-refined structures (taken from last 50 ns of simulation) of wild type Cre/loxP and mutant Cre[K201R(A)-K201R(I)]/loxP. In the wild type complex, Cre active (A) monomer is shown in orange, inactive (I) in blue and the DNA in dark gray, whereas for the mutant, the active monomer is shown in ochre, the inactive in gray-blue and the DNA in light gray. For clarity, cartoon representations are shown with transparency and side chains of relevant residues are in balls and sticks and labelled. Intermolecular H-bonds are depicted with dashed lines (ochre and light gray for the active and inactive monomers, respectively). Right panel: intermolecular hydrogen bond and van der Waals contacts profile from last 100 ns of MD simulations of Cre[K201R(A)-K201R(I)]. D | Left panel: details of the superimposition of the MD-refined structures (taken from last 50 ns of simulation) of wild type Cre/loxP and mutant Cre[Q311R(A)-Q311R(I)]/loxP. The lack of H-bond between H289 in the active monomer of Cre[Q311R(A)-Q311R(I)] and the phosphate group of A4'TS in loxP is highlighted with two ochre arrows pointing at those atoms. Right panel: intermolecular hydrogen bond and van der Waals contacts profile from last 100 ns of MD simulations of Cre[Q311R(A)-Q311R(I)]. E | Left panel: details of the superimposition of the MD-refined structures (taken from last 50 ns of simulation) of wild type Cre/loxP and mutant Cre[K201R(A)-Q311R(I)]/loxP. Right panel: intermolecular hydrogen bond and van der Waals contacts profile from last 100 ns of MD simulations of Cre[K201R(A)-Q311R(I)]. F | Intermolecular hydrogen bond and van der Waals contacts profile from last 100 ns of MD simulations of Cre[Q311R(A)-K201R(I)]. In each panel, active versus inactive mutants are highlighted with green tick and red cross symbols at the left of their names, respectively.

A

Name	Half-site orientation	Full left site	Full right site
loxF8	LR-LR	ATAAATCTGTGGAAACGCTGCCACACAATCTTAG	ATAAATCTGTGGAAACGCTGCCACACAATCTTAG
loxF8L	LL-LL	ATAAATCTGTGGAAACGCTGCTCCACAGATTTAT	ATAAATCTGTGGAAACGCTGCTCCACAGATTTAT
loxF8R	RR-RR	CTAAGATTGTGTGAACGCTGCCACACAATCTTAG	CTAAGATTGTGTGAACGCTGCCACACAATCTTAG
IofF8:RL-RL	RL-RL	CTAAGATTGTGTGAACGCTGCTCCACAGATTTAT	CTAAGATTGTGTGAACGCTGCTCCACAGATTTAT
loxF8:LR-RL	LR-RL	ATAAATCTGTGGAAACGCTGCCACACAATCTTAG	CTAAGATTGTGTGAACGCTGCTCCACAGATTTAT
loxF8:LL-LR	LL-LR	ATAAATCTGTGGAAACGCTGCTCCACAGATTTAT	ATAAATCTGTGGAAACGCTGCCACACAATCTTAG
loxF8:RR-LR	RR-LR	CTAAGATTGTGTGAACGCTGCCACACAATCTTAG	ATAAATCTGTGGAAACGCTGCCACACAATCTTAG

B
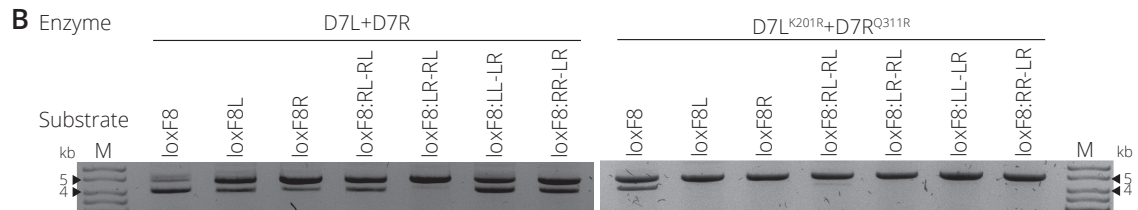
Supplementary Figure S6 | Half-site combinations of the loxF8 target site. A | Table showing the different half-site combinations. For example: The loxF8 target (LR-LR) site has: the half-site orientation of the full left loxF8 target sequence, left half-site (L) (blue) then right half-site (R) (orange); the half-site orientation of the full right loxF8 target sequence, left half-site (L) (blue) then right half-site (R) (orange). B | Bacterial assay of recombination activity for D7L+D7R compared to D7L^{K201R}+D7R^{Q311R} on the half-site combinations. Concentration of L-Arabinose for all samples 10 µg/ml. M = GeneRuler DNA Ladder Mix 10kb.

Supplementary Materials - MD

Materials & Methods- MD simulations protocol

The structures of Crewt and mutants in complex with loxP were energy refined by MD simulations in AMBER20 (1) (https://ambermd.org/) using the following protocol: Each complex structure was solvated using a truncated octahedral box of SPC water molecules and neutralized with K^+ ions. Simulations were performed with a concentration of 150 mM KCl using Joung and Cheatham parameters (2). Ions were randomly placed in the simulation cell at least 5 Å away from the DNA. MD simulations were preceded by two energy-minimization steps, first with position restraints for the solute (500 kcal/mol·$Å^2$) and then without restraints by applying in both cases 5000 cycles of steepest descendent and 5000 steps of conjugate gradient. The system was heated up from 200 K to 300 K in 50 ps with weak position restraints (10 kcal/mol·$Å^2$). Langevin temperature coupling with a collision frequency $\gamma = 1$ ps^{-1} was used at this step. Next, three equilibration steps of 500 ps each with consecutively decreased positions restraints (10, 5 and 2 kcal/mol·$Å^2$, respectively) were conducted under constant pressure of 1 atm using periodic boundary conditions (NPT conditions) at 300 K. The system was further equilibrated during 1 ns without restraints at 300 K under same conditions. A total of 200 ns MD simulation was carried out at 300 K NPT conditions for each complex. The SHAKE algorithm with a time step of 2 fs was used to constrain all bonds involving hydrogen atoms. A cutoff of 10 Å was applied to treat the non-bonded interactions, and the Particle Mesh Ewald (PME) method was used to treat long- range electrostatic interactions.

1. Case,D.A., Belfon,K., Ben-Shalom,I.Y., Brozell,S.R., Cerutti,D.S., Cheatham,T.E., III, Cruzeiro,V.W.D., Darden,T.A., Duke,R.E., *et al.* (2020) Amber 2020 University of California.

2. Joung,I.S. and Cheatham,T.E. (2008) Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J Phys Chem B*, **112**, 9020–9041.

Supplementary Table S1 | Oligonucleotides listed 5' to 3'

Primer (p) number	Description	Oligo Sequence (5'-3')
1	Sequence validation of recombinase in pEVO	TGCATCAGACATTGCCGTCA
2	Sequence validation of recombinase in pEVO	AGACCGCTTCTGCGTTCTGA
3	Selection of recombinases active on loxF8 (binds downstream of target sites)	AAGGGAATAAGGGCGACACG
4	Selection of recombinases inacvite on the loxF8L or loxF8R sites	CTAACTGACACACATTCCACA
5	degenerate VNS codon D7R Position 76 (SNB)	ACGCGTCTGAAGGTGSNBGAGGTAATC
6	degenerate VNS codon D7L Position 32 and 33 (SNB)	AGAAAACGCCCGGCGSNBSNBGAAGAC
7	Degenerate VNS codon D7L position 25 and 29	AGTGATGAGGCTCGCVNSAACCTGATGVNSGTCTTC
8	Degenerate VNS codon D7R position 69 and 72	GTAGAACCTVNSGATGTTVNSGATTACCTC
9	Forward nested PCR primer for D7L	ATGTCCAATCTACAGACCCTACACCAGAATTTG
10	Degenerate GHW codon D7L position 25 and 29	GAAGACWDCCATCAGGTTWDCGCGAGCCTCATCACT
11	Degenerate MDG codon D7L position 25 and 29	GAAGACCHKCATCAGGTTCHKGCGAGCCTCATCACT
12	Degenerate GHW codon D7L position 32, 33 and 35	GTCTTCGHWGHWCGCGHWGCGTTTTCTGAAGCT
13	Degenerate MDG codon D7L position 32, 33 and 35	GTCTTCMDGMDGCGCMDGGCGTTTTCTGAAGCT
14	Reverse nested PCR primer for D7L	ATTCAGCTTGCACCATGCCGCCCACGTCCGGCA
15	Forward nested PCR primer for D7R	CTGTCCGTTTGCCGGTCGTGGGCGGCATGGTGC
16	Degenerate GHW codon D7R position 69 and 72	GAGGTAATCWDCAACATCWDCAGGTTCTACGGG
17	Degenerate MDG codon D7R position 69 and 72	GAGGTAATCCHKAACATCCHKAGGTTCTACGGG
18	Degenerate GHW codon D7R position 76	GATTACCTCGHWCACCTTCAGACGCGTGGTCTG
19	Degenerate MDG codon D7R position 76	GATTACCTCMDGCACCTTCAGACGCGTGGTCTG
20	Reverse nested PCR primer for D7R	AGCCCGACGGTGAAGCATGTTTAGCGAGCCCAG
21	mRNA production adding T7 promoter	GCTAATACGACTCACTATAGGGAGAGCCGCCACCATGCCAAAAAAGAAGAGAAAGGTAATGTCCAATTTACTGACCGTACACCA
22	mRNA production for Cre adding poly(A) signal	TTTTTTTTTTTTTTTTTGGTTTATTCCTAATCGCCATCTTCCAGCAG
23	Colony selection	GCCCCTAACTCCGCCCATCC
24	Colony selection	ACAGTACTGCGATGAGTGGC

Supplementary Table S2 | Sanger sequencing results of colony screen. The table shows variants with mutations at positions D7L - 25, 29, 201, 282, 305 and D7R – 311 (60 out of the 75 sequenced variants). The first row is the original D7 sequence at these positions.

	D7L					D7R
	25	29	201	282	305	311
D7	**K**	**D**	**K**	**G**	**S**	**Q**
H10				E		R
H09	Q	A				R
H07	V					
H05	Q	M	R			R
H03	Q		R			R
G12				E		
G10	M	M	R		P	R
G09	A	V		E		
G07	V			E		R
G06			R	E		
G05			R	E		
G02					P	R
F10	E	V	R	E		R
F08	Q	K	R			R
F07	Q	M		E		
F06	M	M	R	E	P	
F05			R			
F04	V	A	R			
F02	Q		R		P	
F01	Q	V	R			H
E12				E		R
E11			R	E		
E09	T	V	R			R
E07	M	M	R	E	P	
E06	L	K	R	E		R
E05	L	K	R	E	P	
E03			R	E	P	
E02			R			R
E01					P	
D11	R	M		E		
D10	E	A	R			R
D09			R			R
D07			R		P	
D06	Q			E		
D05	R	P		E	P	
D04	Q	M		E	P	
D03	W			E		
D02	Q					R
D01			R			
C12	Q	L				
C11			R	E		
C10				R		R
C09	E		R	E	P	
C08	R	R	R			
C07		K		E		R
C06	M		R	E		
C03			R	E		R
C02			R		P	
B12	V		R	E	P	
B09			R	E		R
B06				E	P	R
B03				E	P	R
B01	Q	A				R
A09		V		E	P	
A08	M	Q		E		R
A07	V	K	R			R
A06				R		R
A03				E		
A02				E		R
A01	Q		R	E		

Supplementary Table S3 | Complete sequencing results of colony screen. The table shows the sequencing results of the recombinase pairs selected from the final library. The recombinase pairs are indicated by the ID where the D7L recombinase variant is shown in blue and the D7R recombinase variant is shown in orange. The table first lists the full sequence of the D7L recombinase variant and then the full sequence of the paired D7R recombinase variant.

D7L ID	Sequence
A01	MFNLQTLHQNLAALLANATSDEARQNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVE RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYQTWSGHSARVGAA RDXARAGVSIAEIMQAGGWTTVESMSYXRNLDXEXGAMVRLLEDGD*
A02	MTNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCELNNRKWFPAEPEDVRGY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRSLDSETGAMVRLLEDGD*
A03	MSNLQTLHQNLSALLANATGDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCLDIRNLAFMGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESMSYLCNLXSETGAMVLLEDGD*
A04	MTNLQTLHQNLSALLANATSDEARKNLMDVLRNRRAFSEATWKTLQSVCRTWAAWCKLNNRKWFPAXPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDRVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIRDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTXLVE RWISVSGVAXDPNNCLFCRVRRYGVAXPSATSQLSTXALRGXFAAAHRLVYGXKDDSGQXYRTWSGHSARVGAA RDMARAGVXIPEIMQAGGWTTVKXVMHXXRNLDXEXGAMVXLLE*
A06	MTNLQTLHQNLSALLANXTSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSNAVSLVMRRIRRENVDAGERVRQALPFERTDFDKVRSL MGNSDRCLDIRNLAFLGVAYNTLLRISEIARIRIRDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVX RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQVSTDVLRGVFAAAHRLVYGTKDDSGQRYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVXLLEDGD*
A07	MTNLQTLHQNLSALLANATSDEARVNLMKVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSRLSTDVLRGVFAAAHRQVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
A08	MFNLQTLHQNLSALLANATSDEARMNLMQVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRTSEIARIRIRDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVTAPSATSQLSTDVLRGVFAAAHHLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRDLDSETGAMVRLLEDGD*
A09	MTNLQTLHQNLSTLLANATSDEARKNLMVVLRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKVNVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVX RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIXEIMQAGGWTTVESVMSYLRNLDSETGAMVXCWRXRL*
A10	MYNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWASWCKLSNRKWFPAEPEDVRDY LLHLQVRGLAMNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MENSDRCQDIRNLAFLGVAYNTLLRXSEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVE RWISVSGVAXDPNNYLFCRVRRYGVAXPSATSQLSTXVLXGXFXAAHRLVYGXKDDSGQEYLAWSGHSARVGAA RDMARAGVXIPEIMQAGGWTTVNSVMNYIRNLECETGAMVRLLE*
A11	MXXLQTLHQNLSALLANATSDEARXNLMXVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG

Supplementary Table S3

	RWISVSGVAGDPNNYLFCRVKNGVAAPSATSQLSTDVLRGVFAAAHRXVYGTKDDSGQXYLTWSGHSARVGAAR DMARAGVSIXEIMQAGGWTTVXXMSYLRNLDSETGAMVL*
A12	MTNLXTLHQXLSALLANATSDEARKNLMDVXRDRRAFSEATWKTLLSVXRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLXVNTIQRHLALLNMLHRRSGLPRPGDSNAVSLVMRRIRKENVDAGERXRQALAFERTDFDRVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVE RWISVSGVADDPNNYLFCRVRRYGVAXPSATSQLSTXXLXGIFAAAHRLVYGXKDDSGQXYLTWSGXSARVGPA RDMARAGVPIPEIMQAGGWTTVNSVMNXXRNLXXETGAMVXLL*
B01	MYSLQTLHQNLSALLANATSDEARQNLMAVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL IGNSDRCLDIRNLAFLGVAYNTLLRISEIARIRIKDISLTDGGRMLIRIGSTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTRDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
B02	MTNLQTLHQNLPALLANATSDEARKNLMDVSRDRRAFSEATWKTLLSVCRTWAAWCKLDNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVCSL MGNSDRCQDIRNLALLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYPTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
B03	MTNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCLDIRNLAFLGVAYNTLLRLSEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLAYGTKDDSGQEYLAWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
B04	MSNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKXNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSCAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL XGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVE RWISASGVADPNNYLFCRVRRYGAAXPSATSQLSTXVLRGIFAAAHRLVYGXKDDSSQXYXTWSGHSARVGAAR DMARAGVXIPEIMQAGGWTTVNSVMNXXRNLDXEXGAMVRLL*
B05	MXXLQTLHQNLSALLANXTSDEARQNLMXVFRDRRAFSEATWKTLLSVCRTWAXWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNSIQRHLALLNMLHRXSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVSRLVX RWISVSGVAXDPNNYLFCRVRKNGVAAPSATXQLSTDVLRGVFAAAHRLXYGTRDDSGQXYLTWSGHSARVGAA RDMARAGVXIXEIMQAGGWTTVEXMNXXRNLDSEXGAMVRLLEDGD*
B06	MSNLQTLHQNLSALLANXTSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLTVNTIQRHLALLNMLHRWSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGAEKALSLGVTRLVG RWISVSGVAXDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDCETGAMVRLLEDGD*
B07	MYNLQTLHQNLSALLANAASDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLXVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCLDIRNLAFLGVAYNTLLRIXEIARIRIKDISRTGGGRMLIQIGRTRTLVSTAGVEKALSLGVTRLVE RWISVSGVAXDPNNYLFCRVRRNGVAAPSAASQLSTDVLRGXFTAAHRLVYGTKDDSGQXYLTWSGHSARVGAA RDMARAGVXIPEIMQAGGWTTVKSVMXYLRNLDSETGAMVRLLEDGD*
B08	MTNLQTLHQNLXALLANATSDEARXNLMXVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVXTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRXSEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYXTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGX*
B09	MSNLQTQHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQARGLAVNTIQRHLALLNMLHRRSGLPRPGDSGAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGRSARVGAA RDMARAGVSIAEIMQAGGWTTVESVVSYLRNLDSETGAMVRLLEDGD*
B10	MSNLQTLYQNLSALLANATSDEARKNPMDVFRDRRAFSDATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSTVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL KGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGSTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPDNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQXCLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEXGD*

Supplementary Table S3

B11	MTNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQERGLAVNTILRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVHSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVE RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDXSGQXYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESMSYLRNLDSETGAMVRLLEXGD*
B12	MTNLQTLHQNLSALLANATSDEARVSLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL LGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDCETGAMVRLLEEGD*
C02	MTNLQTLHQNLSSLLANATSDEARKNLMDVFRDRRAFSKATWKTLLSVCRTWAAWCKLNNRKWFPAEPEGVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRTSEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEGGD*
C03	MTNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFATAHRLVYGTRDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
C06	MYYLQTLHQNLSALLANATSDEARMNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNREWFPAEPEGVRDY LLHLQVRGLAVNTIQRHLALLNMLHRWSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVE RWISVSGVAXDPNNYLFCRVRRNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGNSARVGAA RDMARAGVSIXEIMQAGGWTTVESVMSYLRDLDSEAGAMVRLLEDGD*
C07	MYNLQTLHQNLSALLANATSDEARKNLMKVFRDRRAFSEATWKTLLSVCRTWAAWCELNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVTAPSATSQLSTDVLRGVFAASHRLVYGTKDDSGQEYLTWSGNSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLCNLDSETGAMVRLLEDGD*
C08	MSNLQTLHQNLSALLANATSDEARRGLMRAFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVTTIQRHMALLNTLHRWSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDVSRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYRTWSGLSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLCNLDSETGAMVRLLEDGD*
C09	MTNLQTLYQNLSALLANATSDEARENLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSNRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDALRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESMSYLRNLDNETGAMVRLEDGD*
C10	MFNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQRYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRILDSETGAMVRLLEDGD*
C11	MTNLQTLHQNLSALLANATSGEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTVQRHLALLNMLHRRXGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRIPEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVGKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
C12	MYNLQTLHQNLSALLANATSDEARQNLMLVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRRALAFERTDFNKVRSL MGNSDRCQDIRSLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVE RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVXSVMSYLRNLDSETGAMVRLLEDGD*
D01	MSNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTGLVG

Supplementary Table S3

	RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYVTKDDSGQGYLTWSGLSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLCNLDSETGAMVRLLEDGD*
D02	MTNLQTLHQNLSALLANAISDEARQNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL KGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSSTSQLSTDVLRGVFAAAHRLVYGTRDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLHNLXSEXGAMVLLEDGD*
D03	MTNLQTLHQNLSALLANATSDEARWNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDRVRSL MGDSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTEDNSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLHNLDSETGAMVRLLEDGD*
D04	MSNLQTLHQNLSALLANATSDEACQNLLMVFRDRRAFSEATWKTLLSVCRTWAAWCKLSNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGGEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTAKSMSYLRNLXIEXGAMVLLEXSD*
D05	MYNLQTLHQNLSALLANATSDEARRNLMPVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGGEKALSLGVTRLVG RWISVSGVAGDPNNFLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
D06	MSNLQTLHQNLSALLANATSDEARQSLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQARGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVTRRIRKENVDAGERVWQALAFERTDLDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHCLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
D07	MSNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENIDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISHTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESMSYLRNLDSETGAMVXLLEDGD*
D09	MTNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHCRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDRVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEVARIRIKDISRTDGGRMLIHIGRTRTLASTAGVEKALSLGVTRLVG RWISVSGIAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSEXGAMVRLLEDGD*
D10	MTNLQTLHQNLSALLANATSDEARENLMAVFRDRRAFSEATWKTLLSVCRTWAAWCRLNNRKWFPAEPEDVRGY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGVRVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRIPEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSEXGAMVRLLEDGE*
D11	MTNLQTLHQNLSALLANATSDEARRNLMMVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQAPAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSVAEIMQAGGWTTLESVMSYLRNLDSETGAMVRLLEDGD*
D12	MXNLQTLHQNLSALLANATSDEXRXNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCXLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRXALAFERTDFDKXRSL MGNSDRCXDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVX RWISVSGVAXDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMXXLRNLDSETGAMVRLLEMAI*
E01	MYNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRRWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGXSDRCQDIRNLAFLGVAYNTLLRTSEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTRDDSGQGYLTWSGLSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*

Supplementary Table S3

E02	MSNLQTLHQNLPALLAYATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
E03	MTNLQTLHQNLSALLXNATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLXVNXIQRHLALLNMLHRXSGLPRPGDSSAVSLVMRRIRKENVDXGERVRQALAFERTDFDKVRSL MGNSDRXXDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSXAGVEKALSLGVTRLVE RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSEGAMVRLLEDGD*
E04	MXNLQTLHQNLTALLANATSDEARXNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCXLNNRKWFPAEPEDVRDY LLHLQIRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALXFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGLSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLED*
E05	MSNLQTLHQNLSAILANATSDEARLNLMKVFRDRRAYSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVX RWISVSGVAXDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIXEIMQAGGWTTVESVMSYXRNLDSEXGAMVRLLEDGD*
E06	MSNLQTLHQNLSALLANAASDEARLNLMKVFRDRQAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVVRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSIGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDXETGAMVRLLEXAI*
E07	MTNLQTLHQNLSALLANATSDEARMNLMMVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDIRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRTL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTSLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVXSVMNYLRNLDSETGAMVLLEDGD*
E08	MTNLQTLLQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNREWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRWSGLSRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKIRSL MGNSDRCQDIRNLAFLGVAYNTLLRIPEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWATVESMSYLRDLDSETGAMVRLLEDGD*
E09	MSNLQTLHQNLSALLANATSGEARTNLMVVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQARGLVVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSIGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKGDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESMSYLRNLDSETGAMVRLLEDGD*
E11	MTNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSETTWKTLLSVCRTWAAWCELNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVTRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLALLGVAYNTLLRISEIARIRIKDISHTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
E12	MTNPQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSACRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQTLAFERTDFDKVRTL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVADDSNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVVSYLRDLDSETGAMVRLLEDGD*
F01	MLNLQTLHQNLAALLANATSDEARQNLVVVFRDRRAYSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFATAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSCLRNLXSEXGAMVRLLEDGD*
F02	MTDLQTLHQNLSALLANATSDEARQDLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSAAGVEKALSLGVTRLVG

Supplementary Table S3

	RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVEGVMSYLRNLDSETGAMVRLLEDGD*
F03	MTNLQTLHQNLSALLANATSDEARMNLMXVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRXY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSXAVSLVMRRIRKENVDAGERVRQALAFERSDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVATPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
F04	MTNLQTLHQNLSALLANAASDEARVNLMAVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDLDKVRSL MGNSDRCQDIRNLALLGVAYNTLLRMPEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLXSEXGAMVRLLEDGD*
F05	MTNLQTLLQDLSALLANAASDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDRVRSL MGNSDRCQDIRNLAFLGVAYNTLLRTPEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVT RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
F06	MTNLQTLHQNLSALLANATSDEARMNLMMAFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVSTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSETGAMVLLEDGD*
F07	MTNLQTLHQNLSALLANATSDEARQNLMMVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDH LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRVSEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTNDDSGQEYXTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLSNLDSETGAMVRLLEDGD*
F08	MSNLQTLHQNLSALLANATSDEARQNLMKVFRDRRAFSEATWKTLLSVCRTWVAWCKLSNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTGLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESMSYLRNLDSETGAMVRLLEDGD*
F09	MXNLQTLHQNLXALLGNATSDEARENLMXVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQXRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISXTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVHGTKDDSGQEYLTWSGHSARVGPA RDMARAGVSIAEIMQAGGWTTVESAMNHLRNLDS*
F10	MSNLQTLHQNLSXLLANATSDEARENLMVVFRDRRAFSEATWKTLLSVCRTWAAWCELNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRXQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEXXTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSEAGAMVRLLEDGD*
G02	MFNLQTLHQNLSALLADATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGECVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTGLVG RWISVSGVAGDPNNYLFCRVRKNGVAVPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
G04	MXNLQTLHQXLSALLANATSDEARXNLMXVXRDRRAFSEATWKTLLSVCRTWAAWCKLNXRKWFPAEPEDVRDY LLHLQVRGLAVNTXQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERXDFDKVRXL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVE RWISVSGVAGDPNNYLFCRVRRYGVAAPSATSQLSTDVLRGXFAAAHRLVYGTKDDSGQXYLTWSGHSARVGAA RDMARAGVPIXEIMQAGGWTTVESVMSYXRNLDSETGAMVRLLEEGD*
G05	MTNLQTLHQNLSALLANATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLTVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERSDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYQTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMXYLCNLDSETGAMVRLLEDGD*

Supplementary Table S3

G06	MTNLQTLHQNLSALLANATSDEARKNLMDVIRDRRAFPEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRMSEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVX RWISVSGVAXDPNNYLFCRVRRYGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQECLTWSGHSARVGAA RDMARAGVSIXEIMQAGGWTTVKSMSYXRNLDSEXGAMVRLLEDG*
G07	MTNLQTLHQNLSALLANATSDVARVNLMDVFRDRRAFSEATWKTLLSVCRTWAVVCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRILIHIGRTKTLVCTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSAISQLSTDVLRGVFAAAHRLVYSTKDDSGQEYRTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
G09	MTNPQTLHQNLPALLANATSDEARANLMVVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLTVNTIQRHLALLNMLHRRSGLPRPGDSNAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLSWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESMSYLRDLDSETGAMXRLLEXGD*
G10	MYNLQTLHQNLSALLANATSGEARMNLMMVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLTVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNDVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESMSYLRNLDSETGAMVRLLEDGD*
G11	MTNLQTLHQNLSALLANAISDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRIPEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRQVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
G12	MTNLQTLHQNLSALLADATSDEARKNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDQVRSL MGNSDRCQDIRNLAFLGVAYNTLLRIPEIARIRIKDISRTDGGRMLIHISRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESMSYLRNLDSETGAMVRLLEDGD*
H02	MTNLQTPHQNLPALLANXTSDEARXNLMDVFRDRRAFSEATWKTXLSVCRTWAAWCKLNNRKWFPAXPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYQTWSGLSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVLLLEMAI*
H03	MSNLQTLHQNLSALRANATSDEARQNLMDVLRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLTVNTIQRHLALLNMLHRWSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRVSEIARIRIKDISRTDGGRMLIHIGRTRTLISTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSASSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTIESVMSYLRNLDSETGAMVRLLEDGD*
H04	MFNLQTLHQNLSALLGNATSDEARRNLMKVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGSTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQXYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMXYLRNLDSETGAMVRLLEDGD*
H05	MSNLQTLHQNLSALLANATSDEARQNLMMVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAKPEDVRDY LLHLQVRGLAMNTIQRHLALLNVLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRDLDSETGAMVRLL*
H06	MTNLQTLHQNLSALXANATSDEARQNLMDVXRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRXSGLPRPGDSSAVSLVMRRIRKENVDAGERXRQALPFERTDFDKVRSL MGNSDRCXDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVE RWISVSGVADDPNNYLFCRVRRYGVAAPSATSQLSTDVLRGIFAAAHRLVYGXKDDSGQXYLTWSGHSARVGAA RDMARAGVXIPEIMQAGGWTTVKGXNXXRNLDSETGAMVRLLEX*
H07	MTNLQTLHQNLSALLANATSDXARVNLMDVSRDHRAFSEATWKTLLSVCRTWAAWCKPNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRIXEIARIRIKDISRTDGGRMLIHISRTKTLVSTAGVEKALSLGVTRLVX

	RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRQVYGTMDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
H08	MXNLQTLHQNLSALLANATSDEARXNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTRTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVPIAEIMQAGGWTTVESVMSYLRNLDSEXGAMVRLLEDGD*
H09	MTNLQTLHQNLSALLANATSDEACQNLMAVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRIKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPDNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQGYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDXETGAMVRLLEDGD*
H10	MTNLQTLHQNLSALLANATSDEARKNLMDVLRDRRAFSEATWKTLLSVCRTWAAWCRLNNRKWFPAEPEDVRDY LLHLQVRGLAVNTIQRHLALLNMLHRRSGLPRPGDSSAVSLVMRRIRKENVDAGERVRQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRTPEIARIRIKDISRTDGGRMLIHISRTKTLVSAAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQEYLTWSGHSARVGAA RDMARAGVSIAEIMQAGGWTTVESVMSYLRNLDSETGAMVRLLEDGD*
H11	MTXLXTLHXNLXALLANATSDEARXNLMDVFRDRRAFSEATWKTLLSVCRTWAAWCKLNNRKWXPAXPEXVXXY LLHLQVXGLAVNTIQRHLALLNMLHRXSGLPXPGDXXAVSLVMRRIRKENXDAGERXRXALAFXRTDFXXVRSL XXNSDRCQDIRNLAFLGVAYNXLLXIXAIARIRIKDISRTDGGRMLIHIGRTXTLVSTAGVEKALSLGVTRLVG RWISVSGVAGDPNNYLFCRVRKNGVAAPSATSQLSTDVLRGVFAAAHRLV*

D7R ID	Sequence
A01	QFHIRTPHQSLSALLTDATSEVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTQVSSAGVEKALSLGVTKLVE RWISASGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGVMVRLLEDSD*
A02	MPNLRPPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCGLNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERALQALAFERTDFDQVRSL VGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDDGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
A03	MPNLQTPHQSLSALLTDATSDVTRKNLAYMFRDSQAFSEHTWKMLLSVCRSWAAWCKLNNRKWFPVEPEDVRDY LQHLQVRGLAVRTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGIAKPSATSQLSTYALQGIFGAAHRLVYSAKGVSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
A06	RSNLQTPPQSLCITDDATSDTCKNLANIFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYLL HLQTRGLAVKTIQHHLGSLNMLHRQAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFGRTDFDQVRSLVE NSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVERW ISVSGVSDDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAARD MARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
A07	SSPFPDPPPSLSALLTDAMSDVTRKNLADVFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDRVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHHLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
A08	ISNLQTPHQRLSALLTDATSDVTRKNLAVMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGTLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVVKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLTWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
A09	RSNLQTPHQSLSALLTDATSDVTRKDLAGMFRDGQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVRTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE

Supplementary Table S3

	RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQRIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B01	FSSIYSPHPKLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGEHAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B02	RSNLQTPHQSLSALLSDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDRVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVERALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B03	PVHIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPAEPEDVRDY LLHLQTRGLAVRTIQHHLGSLNMLHRRAGLPRPGGSNAVSLVMRRIRRENVDAGEHAQQALAFERTDFDQVRSL VENSDRCQDVRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISASGVAGDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B05	MSNIQTPHQLFALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPAEPEDVRDYL VHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVER WISVSGVADDPNNYLFCRVRRYGVAKPSAISQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMQAGGWSTVNSVMNYIRNLDSETGAMVRLLEDGG*
B06	MSNLQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLAMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGRKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B07	QVHIQTPHQKLSALLTGATSDVTRKNLADMFRDRQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSSAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRADGGRMLIHIGRTKTLVSTAGVERALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B08	VRRVQSTHTKLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTQVSTAGVEKTLSLGVTKLVE RWISVSGVSDDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHCLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B09	QFQFPTPHQSLPALLTDATSDVTRKNLANMFRDSQAFSEHTWKMLLSVCRSWVAWCELNNRKWFPAEPEDVRDY LRHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVSKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSEAGAMVRLLEDSD*
B10	RGHVQFPHPNFFALLTDATSDVTCKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VETSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
B11	RSNLQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDARDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPSDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVAGDPNNYLFCRVRRYGVAKPSATSRLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSS*
B12	MSNLQTPHQCLPALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHVGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*

Supplementary Table S3

C02	MSNLQTPHQSLSALLTDATSDVIRKNLADMFRDNQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LMHLQTRGLAVKTIQHHLGTLNMLHRGAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
C03	MSNIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRVY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVEAGERAQQALAFGRTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIQEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
C06	MSNLQTPLQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISHTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGIAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
C07	GSNLQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPAEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLISTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRFGVAKPSATSQLSTYALQGIFGVAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPLPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDGD*
C08	MSNVQTPHQSLSALLTGATSDVTRKNLAHMFRDSQAFSEHTWKMLLSVCRSWATWCELNNRKWLPAEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRSQDIRNLAFLGVAYNTLLRISEIARIRAKDISRTDSGRMLIHIGRTKTLVGTAGVEKVLSLGVTKLVE RWISVSGVADDPDNYLFCRVRRYGVAKPSATSQLSTYAMQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
C09	SSSNYRPHPNLSALLTDATSDVTRKNLVDMFRDSQAYSEHTWKMLLSVCRSWAAWCELNNRKRLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVVAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKTLSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSE*
C10	SSNLQTPPQLSALLTDAMSDVTRKNLANMFRDSQAFSEHTWKTLLSVCRSWAAWCELNDRKLLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSAAGVEKALSLGVTKLVER WISVSGVADDPNNYLFCRVRRYGVVKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
C11	MSNIQTPHQSLSALLTDAASDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKRLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGAEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGNSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
C12	SSNLQTPHQSLPALLTDATSDVIRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELSNRKWLPAEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNVVLLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VGNSDRCQDIRNLALLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGIEKALSLGVTKLVE RWISVSGVAEDPNNYLFCRVRRYGVAKPSATNQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGWLDPVNSVMNYFQNLDSETGAMCSLLQDTD*
D01	QVQLQTPHQSLSALLTDATSDVTRKNLADIFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFRRTDFDKVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTGGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGTKCDSGQKYLAWSGHSARVGAA RDMARAGVPIQGIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
D02	MSNLQPHQSLSALLTDAASDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYL LHLQARGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFGRTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTQVSTAGVEKALSLGVTKLVGR WISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGDKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
D03	MTNLQTPPQLFALLSDATSDARKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCEMNNRKWLPAEPEDVRDYLL HLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRESVDAGERAQQALAFERTDFDQVRSLVE NSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVERW

22

Supplementary Table S3

	ISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAARD MARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
D04	MSNLQTPHQLFALLTDATSDVTRKNLADMFRDSQAFSEHTWKTLLSVCRSWAAWCELNNRKWLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFGRTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVER WISVSGVADDPNNYLFCRVRRYDVAKPSATSQLSTYTLQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
D05	MSNIQTPHQSLSALLTDATSDVTCKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKRLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLASTAGVEKALSLGVTELVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYQAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
D06	SSSIHTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL MENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVRDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVAEDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVYSVMNYIRNLDSETGAMVRLLEDSD*
D07	MSSIQTPHQSLSALLTDATSGVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGGRAQQALAFERTDFDQVRSL VENSDRCRDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGLEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
D09	SSSIYRPHTKFVCILTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVRDISRTDGGRMLIHIGRTKTLASTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
D10	ISHFQAPNPSLFALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERADFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLVHIGRTKTRVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVCLLEDSD*
D11	MSNIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNREWFPVEPEDVRDY LLHLQARGLAVKTIQHHLGSLNMLHRRAGLPRPGDSIAVSLVMRRIRRENVDAGERAQQALAFERTDFDRVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYAMQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSEAGAMVRLLEDSD*
E01	RGNLTAPHQSLSALLTDATRDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTQVSTAGVERALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRMVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWATVNSVMNYIRNLDSETGAMVRLLEDSD*
E02	HFQFPTPRQSLSALLTDATSDVICKNLADMFRDSQAFSEHTWKTLLSVCRSWAAWCKLNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKVLSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
E03	SSPITDPTPKLFALLTDATSDLTRKNLADMFRDSQAFSEHTWKMLLSVCRSWTAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPILEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
E05	ISNLQTPHQGLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAVWCELNNRKWFPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNVVLLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHVGRTKTLVRTAGVEKALSLGVTKLVE RWITVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*

Supplementary Table S3

E06	RVQLQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSVHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVEAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDGETGAMVRLLEDSD*
E07	PVQLQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL AENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATRQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLGSETGAMVRLLEGSD*
E08	YVHIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARVRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYSLQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDGD*
E09	VISITTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LMHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENADAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHVGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDLNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGVKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
E11	MSNIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENIDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYPAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
E12	YVHLQTPHQSLSALLTGATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL AENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVVKALSLGVTKLVE RWISVSGVADDSNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F01	WTDIQTPHQSLSALLTDATSDVTRKNLVDMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKRLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMHAGGWSTVNSVMNYIRNLDSETGAMVRLLEDGD*
F02	WSNIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCWSWAAWCELNNRKWLPAEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERADFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTQVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSASSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F03	RVHVQLPHPNLSALLTVATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNVVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F04	HVQLQTPRQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRESVDAGERAQQALAFERTDFDQVRSL MENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLIE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVHGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F05	MSNLQTPRQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNTLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFGRTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVGKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F06	ASSLQTLHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWRMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNTVSLVMRRIRRENIDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRVSEIARIRVKDISRTDGGRMLIHIGRTKTLVSAAGVEKALSLGVTKLVE

Supplementary Table S3

	RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTANSVMNYIRNLDSETGAMVRLLEDSD*
F07	VQSIYSPTPNLSALLTDATSDVARKNLAVMFRDSQAFSEHTWRMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLSRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F08	MSNYRPHPKVVCLLTDATSDVTRKNLAVMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTELVER WISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGRKYLAWSGHSARVGAAR DMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F09	QSNIRTPHQLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLAMRRIRRENVDAGERAQQALAFERTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVER WISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
F10	QSNIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFTEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLSMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERTQQALAFERTDFNQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
G02	HVQLPTPTPSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLGSETGAMVRLLEDSD*
G05	MSNLQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNLKRLPAEPEDVRDY LRHLQARGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAAPSATSQLSTDVLRGVFAAAHRLVYGTKDDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
G06	MSNLQTPPNLSALLTDATSDARKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYLL HLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSLVE NSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVERW ISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAARD MARAGVPIPEIMQAGGWTTVNSVMNYIRNMDSETGAMVRLLEDSD*
G07	HVQLQTPHQSLSALLTDATSDVTRKNLADMFRDSPAFSEHTWKMLLSVCRSWAAWCEQNNRKWLPVEPEDVRDY LLHLQMRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVGKTLSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
G08	MSNLQTPHQSLSALLTDATRDVTRKNLVDMFRDSQAFSEHTWRMLLSVCRSWTAWCELNNRKRLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVRDISRTDGGRMLIHIGRTKTLISTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAT RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
G09	MSNLQTPPESFALLTDATSDVTRKNLADVFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDKVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTRLVER WISVSGVADDPSNYLFCRVRRYGIAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMQAGGWTTVNSVVNYIRNLDSETGAMVRLLEDSD*
G10	RSNLRTPHQSLSALLTDATSDVTRRNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATGQLSTYALQGIFGAAHRLIYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*

Supplementary Table S3

G11	MSNIQTPHQSLFALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDRVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSAAGVEKALSLGVSKLVE RWISVSGVADDPDNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
G12	MSNLRTPHQRLSALLADATSDVTRMNLADIFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRRWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDKVRSL MGNSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTENSVMNYIRNLGSETGAMVRLLEDSD*
H02	MPNLRPPHQSLSALLTDATSDVTRKNLADVFRDSQAFSEHTWKTLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL MGNSDRCQDVRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLASTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWATVNSVMNYIRNLDSETGAMVRLLEDSD*
H03	YSSITDPPPNFFALLTDATSDVTRKNLADMLRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VETSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGNSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
H04	RSRITDPPPNLFALLGDATSDACQNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFGRTDFDQVRSLV ENSNRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVER WMSVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHHLVYGAKGVSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSEAGAMVRLLEDSD*
H05	MSNLQTPHQSSSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPAEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSKLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
H07	MSNIQTPHQGLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQARGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*
H08	MSNIRPHPQIVALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGAEKALSLGVTKLVER WISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEDGD*
H09	ISNLQTPTKVVCILTDATSDATRKNLADMFRDSQAFSEHTWKMLLSVFRSWPAWCELNNRKWLPVEPEDVRDYL LHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSLV ENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTQVSTAGIEKALSLGVTKLVER WISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTFALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAAR DMARAGVPIPEIMRAGGWTTVNSVMNYIRNMDSETGAMVRLLEDSD*
H10	SSNLRPPPQILSIIDEATMMDSQNLADMFRDSQAFSGHTWKMLLSVCRSWAAWCELNNRKWPPVEPEDVRDYLL HLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSLVE NSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLVHIGRTKTLVSTAGVEKALSLGVTKLVERW ISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAARD MARAGVPIPEIMRAGGWTTVNSVMNYIRNLDSETGAMVRLLEGSD*
H11	HVHIQTPHQSLSALLTDATSDVTRKNLADMFRDSQAFSEHTWKMLLSVCRSWAAWCELNNRKWLPVEPEDVRDY LLHLQTRGLAVKTIQHHLGSLNMLHRRAGLPRPGDSNAVSLVMRRIRRENVDAGERAQQALAFERTDFDQVRSL VENSDRCQDIRNLAFLGVAYNTLLRISEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVE RWISVSGVADDPNNYLFCRVRRYGVAKPSATSQLSTYALQGIFGAAHRLVYGAKGDSGQKYLAWSGHSARVGAA RDMARAGVPIPEIMQAGGWTTVNSVMNYIRNLDSETGAMVRLLEDSD*