

Supplementary Figures

The repeats in SpTransformer sequences result in multiple but limited alternative alignments

Megan A. Barela Hudgell, L. Courtney Smith

Figure S1 | The repeat-based alignment illustrates the elements for a representative set of deduced SpTrf protein sequences.

Figure S2 | The NJGT-PRANK alignment of representative deduced SpTrf protein sequences increases the number of gaps in the alignment.

Figure S3 | The MLGT-PRANK alignment of representative SpTrf protein sequences with a provided maximum likelihood guide tree results in regions of misaligned sequence.

Figure S4 | The ClustalW alignment without manual correction includes a non-matching region of the deduced SpTrf proteins.

	10	20	30	40	50	60	70	80	90	100							
A1--2	AHAQSD	FNERRG	KENGRER	GQDRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPRFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
A2--G	AHAQSD	FNERRG	KENGRER	GQDRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPRFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
G2--10	AHARRD	FNELRG	KENGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPRFG	GAPPMGGP	RQDGGP	MGGRRF	DG
G3--10	AHARRD	FNELRG	KENGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPRFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B2--10	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B3--2	AHARRD	YNERRG	NENGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B3--4	AHARRD	FNERRG	NENGRER	GQDRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B3--10	AHARRD	YNERRG	NENGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B5--10	AHARRD	FNERRG	NENGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B6--10	AHAQRD	FNELRG	KEENGR	KRGQ	GRFR	CGRP	PGM	LMGGP	RQDGG	SMGGRR	FDDG	PG	LGAE	-----	-----	-----	-----
B7--10	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B8--10	AHARRD	FNERRG	NENGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
B8--G	AHARRD	FNERRG	NENGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
C2--4	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
C3--4	AHAQRD	YNERRG	NENGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPGFG	AEQMGGPR	QNGGPM	GGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
C4--G	AHARRD	FNERRG	RENGR	KRGQ	GGF	GGRRPD	GMGG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
C5--4	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D1--2	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D1--2	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D1--4	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D1--10	AHAQRD	YNELRG	NENGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D1d-G	AHAQRD	YNERRG	NENGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D1f-G	AHAQRD	YNERRG	NENGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D1y-G	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D5--10	AHAQRD	FNELRG	KEENGR	KRGQ	GRFR	CGRP	PGM	LMGGP	RQDGG	SMGGRR	FDDG	PG	LGAE	-----	-----	-----	-----
D6--4	AHAQRD	YNELRG	NKNGRER	GQGRF	GGRRPD	GMGGSR	QDGGP	MGGRRF	DGPD	SGAEMD	GRRQDGGP	MGGRRF	DGPGFG	GAPPMGGP	RQDGGP	MGGRRF	DG
D6--10	AHARRD	YNERRG	NENGRER	GQGRF	GGRRPD	GMGGPR	QDGGP	MGGRRF	DGPD	SGAEMD							

110 120 130 140 150 160 170 180 190 200

A1--2 PGFGAQMGGPRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGEGRRDGEETDAAQQIGDGLGGPGRCDGPGHGHYGHQAGARPFPGNPPPFNPEQ

A2--G PGFGAQMGGPRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGEGRRDGEETDAAARQIGPGR---FDGPGHGHYGHQAGARPFPGNPPPFNPEQ

G2--10 PGFGAQMGGPRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGEGRRDGEETDAAARQIGDGLGGPGRFDGPGHGHYGHQAGARPFPGNPPPFNPEQ

G3--10 PGFGAQMGGPRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGEGRRDGEETDAAARQIGDGLGGPGRFDGPGHGHYGHQAGARPFPGNPPPFNPEQ

B2--10 -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGHRHHG

B3--2 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

B3--4 -----HMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

B3--10 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

B5--10 -----HMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

B6--10 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGHRHHG

B7--10 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGSDRFDGLGRHHG

B8--10 -----HMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

B8--G -----HMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDASQQIGDGLGGRGQFDGHRHHG

C2--4 -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGSDRFDGPRRHHG

C3--4 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGSDRFDGPRRHHG

C4--G -----GRRFDGPGFGGSRPDAG-RPFFGQGKRGDGEETDAAQQIGDGLGGSDRFDGPRRHHG

C5--4 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGSDRFDGPRRHHG

D1--2 -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D1--2 -----EMDGRRQNGGPMGRRFDGPGFGGSRPVGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D1--4 -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D1--10 -----EMDGTENGGPMGGRELDPGFGGSRPDGGGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D1d-G -----EMDGRRQNGGPMGRRFDGPGFGGSRPVGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D1f-G -----EMDGRRQNGGPMGRRFDGPGFGGSRPVGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D1y-G -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D5--10 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D6--4 -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D6--10 -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

D7--2 -----EMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

E2--2 -----QMEGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRDQFDGHRHHG

E2--4 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQLIGDGLGGRGQFDGHRHHG

E2--10 -----QMEGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDDDDQKMGDGPGRGQFDGHRHHG

E2b-G -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGSGQFDGPRRHHG

E2--G -----QMEGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

E3--2 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQMGDGLGGRGQFDGPRRHHG

E3--2 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGSGQFDGPRRHHG

E3--4 -----QMEGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQLIGDGLGGRGQFDGPRRHHG

E6--2 -----EMDGRRQNGGPMGRRFDGPGFGGSRPVGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

E7--2 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

E8--2 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGPGRGQFDGHRHHG

E9--2 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGPGRGQFDGHRHHG

E10--4 -----QMGGRQNGGPMGRRFDGPGFGGSRPDAGGRPFEGGRRDGEETDAAARQIGDGLGGPGQFDGPRRHHG

F1--2 -----HMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGRGQFDGHRHHG

01--2 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

01--4 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

01--10 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

01--10 -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

01--10 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

01--G -----QMDGRRQNGGPMGRRFDGPGFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

08--2 -----QMDGRRQNGGPMGRRFDGPRFGGSRPDGAGGRPFEGGRRDGEETDAAQQIGDGLGGPGQFDGPRRHHG

	310	320	330	340	350	360	370	380	390	400
A1--2	HTEGHQGHNETGDBPHRHHSKTV	DGDQDT	---	GHHGHHGHHEHHHHQH	DHREGHQDHD	DRPMFGMR	PPFRFNPFGRKPF	GDHPFG	RRNHTEGHQGHNETG	---
A2--G	HTEGHQGHNETGDBPHRHHSKTV	DGDQDT	---	GHHGHHGHHEHHHHQH	DHREGHQDHD	DRPMFGMR	PPFRFNPLGRKPF	GDHPFG	RRNHTEGHQGHNETG	---
G2--10	HTEGHQGHNETGDBPHRHHSKTV	DGDQDT	---	GHHGHHGHHEHHHHQH	DHREGHQDHD	DRPMFGMR	PPFRFNPFGRKPF	GDHPFG	RRNHTEGHQGHNETG	---
G3--10	---	---	---	---	---	---	---	---	---	---
B2--10	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
B3--2	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
B3--4	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
B3--10	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
B5--10	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
B6--10	---	GHNETGDH	---	---	---	---	---	---	---	---
B7--10	---	GHNETGDH	---	---	---	---	---	---	---	---
B8--10	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
B8--G	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
C2--4	---	GHSETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
C3--4	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
C4--G	---	GHNDTGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
C5--4	---	GHNETGDQ	---	---	---	---	---	---	---	---
D1--2	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D1--2	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D1--4	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D1--10	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D1d-G	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D1f-G	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D1y-G	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D5--10	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D6--4	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D6--10	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
D7--2	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
E2--2	---	GHNETGDH	---	---	---	---	---	---	---	---
E2--4	---	GHNETGDH	---	---	---	---	---	---	---	---
E2--10	---	GHNETGDH	---	---	---	---	---	---	---	---
E2b-G	---	GHNETGDH	---	---	---	---	---	---	---	---
E2--G	---	GHNETGDH	---	---	---	---	---	---	---	---
E3--2	---	GHNETGDH	---	---	---	---	---	---	---	---
E3--2	---	GHNETGDH	---	---	---	---	---	---	---	---
E3--4	---	GHNETGDH	---	---	---	---	---	---	---	---
E6--2	---	GHNETGDH	---	---	---	---	---	---	---	---
E7--2	---	GHNETGDH	---	---	---	---	---	---	---	---
E8--2	---	GHNETGDH	---	---	---	---	---	---	---	---
E9--2	---	GHNETGDH	---	---	---	---	---	---	---	---
E10-4	---	GHNETGDH	---	---	---	---	---	---	---	---
F1--2	---	GHNETGDQ	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
O1--2	---	---	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
O1--4	---	---	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
O1--10	---	---	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
O1--10	---	---	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
O1--10	---	---	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
O1--G	---	---	---	---	---	---	---	---	RRNHTEGHQGHNETG	---
O8--2	---	---	---	---	---	---	---	---	RRNHTEGHQGHNETG	---

110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500

A1--2 DH⁺PH⁺RH⁺HS⁺SKT⁺GD⁺GQ⁺QDR⁺PMF⁺ETR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺PH⁺PH⁺GN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PL⁺KVIE⁺IAINE⁺VD⁺T

A2--G DH⁺PH⁺RH⁺HS⁺SKT⁺GD⁺GQ⁺QDR⁺PMF⁺ETR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺PH⁺PH⁺GN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PL⁺KVIE⁺IAINE⁺VD⁺T

G3--10 -----OH⁺DR⁺PMF⁺EMR⁺PF⁺RNF⁺PF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GQRR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVIE⁺IAV⁺NEED⁺V

G2--10 DH⁺PH⁺RH⁺HS⁺SKT⁺GD⁺GQ⁺QDR⁺PMF⁺ETR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺PH⁺PH⁺GN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PL⁺KVIE⁺IAINE⁺VD⁺T

B2--10 DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF⁺EMR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LV--AINEED⁺I

B3--2 DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF⁺ETR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAI⁺-----

B3--4 DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF⁺ETR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAV⁺NEED⁺V

B3--10 DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF⁺ETR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAV⁺NEED⁺V

B5--10 DH⁺PH⁺RH⁺HN⁺KTRD⁺GQ⁺QDR⁺PMF⁺EMR⁺PF⁺RNF⁺PF⁺GR⁺KPF⁺GGR⁺PF⁺DR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LV--AINEED⁺I

B6--10 --PH⁺RH⁺HN⁺KTRD⁺GQ⁺QDR⁺PMF⁺EMR⁺PF⁺RNF⁺PF⁺GR⁺KPF⁺GGR⁺PF⁺DR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAF⁺NEED⁺V

B7--10 --PH⁺RH⁺HN⁺KTRD⁺GQ⁺QDR⁺PMF⁺EMR⁺PF⁺RNF⁺PF⁺GR⁺KPF⁺GGR⁺PF⁺DR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAF⁺NEED⁺V

B8--10 DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF⁺EMR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAV⁺NEED⁺V

B8--G DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF⁺EMR⁺PF⁺WV⁺NPF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAV⁺NEED⁺V

C2--4 DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF⁺ESR⁺PF⁺RNF⁺PF⁺GR⁺KPF⁺GDR⁺PF⁺FR⁺NGT⁺EEG⁺SP⁺RRD⁺GH⁺RR⁺PYGN⁺RGR⁺WG⁺ENE⁺SEE⁺KEH⁺PTTES⁺VTT⁺SSP⁺PE⁺LVVE⁺IAV⁺NEED⁺V

C3--4 DH⁺PH⁺RH⁺HN⁺KT⁺GD⁺GQ⁺QDR⁺PMF

```

      . . . . | . .
A1--2  NVVAEV*
A2--G  NVVAEV*
G3--10 NVVAEV*
G2--10 NVVAEV*
B2--10 NVVAEV*
B3--2  NDVAEV*
B3--4  NVVAEV*
B3--10 NVVAEV*
B5--10 NVVAEV*
B6--10 HVVAEV*
B7--10 NVVAEV*
B8--10 NVVAEVL
B8--G  NVVAEVY
C2--4  NVVAEV*
C3--4  NVVAEV*
C4--G  NVVAEV*
C5--4  NVVAEV*
D1--2  NVVAEV*
D1--2  NVVAEV*
D1--4  NVVAEV*
D1--10 NVVAEV*
D1d-G  VAEV*--
D1f-G  -----
D1y-G  VAEV*--
D5--10 NVVAEV*
D6--4  NVVAEV*
D6--10 NVVAEV*
D7--2  NVVAEVK
E2--2  -----
E2--4  -----
E2--10 -----
E2b-G  -----
E2--G  -----
E3--2  -----
E3--2  -----
E3--4  -----
E6--2  --VAEV*
E7--2  NVVAEV*
E8--2  NVVAEV*
E9--2  -----
E10-4  -----
F1--2  NVVAEV*
O1--2  NGVAEV*
O1--4  --VAEV*
O1--10 --VAEV*
O1--10 --VAEV*
O1--10 --VAEV*
O1--G  NDVAEV*
O8--2  -----

```

Figure S1. The repeat-based alignment illustrates the elements for a representative set of deduced SpTrf protein sequences. The deduced amino acid sequences encoded in the second exon of the *SpTrf* genes are shown. The alignment was done manually in BioEdit (ver 7.2.5) based on a set of pre-aligned deduced sequences according to Buckley and Smith (2007). The mature SpTrf proteins were deduced from the second exon of genes sequenced from animals 2, 4, and 10 to which was added the SpTrf protein sequences from the sea urchin genome (animal G) (1–3). The protein names are located to the left and indicate the element pattern (A-G) and the sea urchin number or letter from which the sequence was obtained (2, 4, 10, G). The ruler above each alignment indicates the aa position in the alignment. Dashes (-) indicate the insertion of artificial gaps in the alignment that optimize matching sequences. Asterisks (*) indicate a stop codon. Element borders are identified by vertical black lines. The sequences are arranged in the alignment with the L proteins at the top, the O proteins at the bottom, and the remaining S proteins in alphabetical order.



	110	120	130	140	150	160	170	180	190	200	
A1--2		MGRRFDGPGFGAP	PMGGPRQDGG	PMGGRRFDGPGFGAP	PMGGPRQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FEGGRRGD			
A2--G		MGRRFDGPGFGAP	PMGGPRQDGG	PMGGRRFDGPGFGAP	PMGGPRQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FEGGRRGD			
G2--10		MGRRFDGPGFGAP	PMGGPRQDGG	PMGGRRFDGPGFGAP	PMGGPRQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FEGSRRGD			
G3--10		MGRRFDGPGFGAP	PMGGPRQDGG	PMGGRRFDGPGFGAP	PMGGPRQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FEGSRRGD			
B2--10	PDSGAP	MDGRRQDGG	PMG		GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B3--2		MG			GRRFDGPGFGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B3--4		MG			GRRFDGPGFGAP	HMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B3--10		MG			GRRFDGPGFGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B5--10		MG			GRRFDGPGFGAP	HMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B6--10		MG			GRRFDGPGFGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B7--10						MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B8--10		MG			GRRFDGPGFGAP	HMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
B8--G		MG			GRRFDGPGFGAP	HMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
C2--4		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
C3--4		MG			GRRFDGPGFGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
C4--G								MGRRFDGPGFGGS	RPDGAG	RPFFQGGKRGD	
C5--4								NGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD
D1--2		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
D1--4		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
D1--10		MG			GRGFDGPGFGAP	EMDGR	TEKNGGPM	MGRELDPGFGGS	RPDGGGRPFF	FQGGRRGD	
D1--10		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPVGAGGRPFF	FQGGRRGD	
D1d-G		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPVGAGGRPFF	FQGGRRGD	
D1f-G		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPVGAGGRPFF	FQGGRRGD	
D1y-G		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
D5--10		MG			GRRFDGPGFGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
D6--4		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
D6--10		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
D7--2		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E2--2		MG			GRRFDGPESGAP	MEGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E2--4		MG			GMRFDPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E2--10		MG			GRRFDGPESGAP	MEGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E2b-G		G			GRRFDGPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E2--G		MG			GRRFDGPESGAP	MEGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E3--2		MG			GRRFD			GPFGGS	RPDGAGGRPFF	FQGGRRGD	
E3--2		MG			GRRFDGPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E3--4		MG			GRRFDGPESGAP	MEGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E6--2		MG			GRRFDGPGFGAP	EMDGR	RQNGGPM	MGRRFDGPGFGGS	RPVGAGGRPFF	FQGGRRGD	
E7--2		MG			GMRFDPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E8--2		MG			GRRFDGPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E9--2		MG			GRRFDGPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
E10-4		MG			GRRFDGPGFGAP	PMGGPRQNGGPM	MGRRFDGPGFGGS	RPDCAGGRPFF	FEGGRRGD		
F1--2		VV			GRRFDGPGFGAP	HMDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
O1--2		MG			GMRFDPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
O1--4		MG			GMRFDPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
O1--10		MG			GRRFDGPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
O1--10		MG			GRRFDGPGFGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
O1--10		MG			GRRFDGPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
O1--G		MG			GMRFDPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	
O8--2		MG			GMRFDPESGAP	MDGR	RQNGGPM	MGRRFDGPGFGGS	RPDGAGGRPFF	FQGGRRGD	

[illegible]

[illegible]

Here Er16 and Er17 are misaligned with Er19 and Er20										
	Er15	Er16	Er18	Er19	Er20	Er21	Er22	Er23	Er24	Er25
	410	420	430	440	450	460	470	480	490	500
A1--2	DHPRHHSKTVLDGDQDTSHHGHHGHHEHHHHQHHRREGHQDHD				RPMFGMRPFREFNPFGRKPFGDHPFGRRNHTEGHQGHNETGDHPRHHSKTVLD					
A2--G	DHPRHHSKTVLDGDQDTSHHGHHGHHEHHHHQHHRREGHQDHD				RPMFEMRPFREFNPFGRKPFGDHPFGRRNHTEGHQGHNETGDHPRHHSKTVLD					
G2--10	DHPNRHHSKTVLDGDQDTSHHGHHGHHEHHHHQHHRREGHQDHD				RPMFGMRPFREFNPFGRKPFGDHPFGRRNHTEGHQGHNETGDHPRHHSKTVLD					
G3--10										
B2--10	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTGL		
B3--2	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTGL		
B3--4	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTGL		
B3--10	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTGL		
B5--10	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTRD		
B6--10	D								HPRHHNKTRD	
B7--10	D								HPRHHNKTRD	
B8--10	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTGL		
B8--G	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTGL		
C2--4	D								HPRHHNKTGL	
C3--4	D								HPRHHNKTGL	
C4--G	D			QDQDKPNDT		RPFRFNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNQTTGL			
C5--4										
D1--2	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D1--4	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D1--10	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D1--10	D			QDQDKLNDT		RPFRYNYFGRKPFGDRTF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D1d-G	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D1f-G	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D1y-G	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D5--10	D			QDQDKLNDT		RPFRYNYFGRKPFGDRTF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D6--4	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
D6--10	D			QDQDKLHGT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHSKTVLD			
D7--2	D			QDQDKLHDT		RPFRYNHFGRKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
E2--2								RNHTEGHQGHNETGDHPRHHNKTGL		
E2--4								RNHTEGHQGHNETGDHPRHHNKTGL		
E2--10								RNRTEGHQGHNETGDHPRHHNKTGL		
E2b-G								RNHTEGHQGHNETGDHPRHHNKTGL		
E2--G								RNHTEGHQGHNETGDHPRHHNKTGL		
E3--2								RNHTEGHQGHNETGDHPRHHNKTGL		
E3--2	D								HPRHHNKTGL	
E3--4								RNHTEGHQGHNETGDHPRHHNKTGL		
E6--2	D								HPRHHNKTGL	
E7--2								RNHTEGHQGHNETGDHPRHHNKTGL		
E8--2								RNRTEGHQGHNETGDHPRHHSKTVLD		
E9--2								RNRTEGHQGHNETGDHPRHHNKTGL		
E10-4			HGRHHQH HH					RNHTEGHQGHNETGDHPRHHNKTGL		
F1--2	D			QDQDKPNDT		RPFRFNH		FGRRNHTEGHQGHNETGDHPRHHNKTGL		
O1--4							RKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD		
O1--2							RKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD		
O1--10							RKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD		
O1--10						GRKPFGDRPF	FGRRNHTEGNQGHNETGDHPRHHNKTRD			
O1--10						RKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD			
O1--G							RKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD		
O8--2							RKPFGDRPF	FGRRNHTEGHQGHNETGDHPRHHNKTRD		

	510	520	530	540	550	560	570	580	590	600	
A1--2		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	DRRNGTEEGSPRRDGHPH	PHGNRGRWGENE	SEEKEHPTTESVTTSSPLK	VIEI--AINE	----	VDTNVVAEV*	
A2--G		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	DRRNGTEEGSPRRDGHPH	PHGNRGRWGENE	SEEKEHPTTESVTTSSPLK	VIEI--AINE	----	VDTNVVAEV*	
G2--10		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	DRRNGTEEGSPRRDGHPH	PHGNRRRWGENE	SEEKEHPTTESVTTSSPLK	VIEI--AINE	----	VDTNVVAEV*	
G3--10		-----FGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AVNE	----	EDVNVVAEV*		
B2--10		GDQDRPMFEMRPFWVNPFG	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VV--AINE	----	EDINVVAEV*	
B3--2		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AIND	----	VAEV*	
B3--4		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AVNE	----	EDVNVVAEV*	
B3--10		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AVNE	----	EDVNVVAEV*	
B5--10		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VV--AINE	----	EDINVVAEV*	
B6--10		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKERPTTESVTTSSPPE	VVEI--AFNE	----	EDVNVVAEV*	
B7--10		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKERPTTESVTTSSPPE	VVEI--AFNE	----	EDVNVVAEV*	
B8--10		GDQDRPMFEMRPF	FWVNPFG	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AVNE	----	EDVNVVAEVL
C2--G		GDQDRPMFEMRPF	FWVNPFG	GRKPF	GDRPF	GRRNTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AVNE	----	EDVNVVAEVY
B8--4		GDQDRPMFESRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AVNE	----	EDVNVVAEV*	
C3--4		GDQDRPMFESRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AFNE	----	EDVNVVAEV*	
C4--G		GDQDRPMFEMRPF	FRNPLGRKPF	GDRPF	GRRNGTEEGSLRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	DVVEI--AVNE	----	EDVNVVAEV*	
C5--4		-----RFNPF	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AVNE	----	EDVNVVAEV*	
D1--2		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
D1--4		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
D1--10		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
D1--10		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
D1d-G		GDQDRPMFESRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSSRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
D1f-G		GDQDRPMFEMRPF	FRNPLGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGR*GENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
D1y-G		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
D5--10		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VV--AINE	----	EDINVVAEV*	
D6--4		GDQDRPMFEMRPF	FWVNPFG	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--VEIAVNE	----	EDVNVVAEV*
D6--10		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	DRRNGTEEGSPRRDGHPH	PHGNRGRWGENE	SEEKEHPTTESVTTSSPLK	VIEI--AINE	----	VDTNVVAEV*	
D7--2		GDQDRPMFEMRPF	FRNPFGRKPF	GGRPF	DRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	V--V--AINE	----	EDINVVAEV*	
E2--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPP*	-----	-----	-----	
E2--4		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEYPTTESVTTSSPP*	-----	-----	-----	
E2--10		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPHYGNRGRWGENE	SEEKEHPTTESVTTYSP*	-----	-----	-----	
E2b-G		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPP*	-----	-----	-----	
E2--G		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPP*	-----	-----	-----	
E3--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPP*	-----	-----	-----	
E3--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPP*	-----	-----	-----	
E3--4		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPP*	-----	-----	-----	
E6--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEYPTTESVTTSSPP*	-----	-----	-----	
E7--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VV--AINE	----	EDINVVAEV*	
E8--2		GDQDRPMFETRPFWVNPFG	GRKPF	GDRPF	DRRNGTEEGSPRRDGHPH	PHGNRGRWGENE	SEEKEHPTTESVTTSSPLK	VIEI--AINE	----	VDTNVVAEV*	
E9--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPHYGNRGRWGENE	SEEKEHPTTESVTTYSP*	-----	-----	-----	
E10-4		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPP*	-----	-----	-----	
F1--2		RDQDRPMFEMRPF	FWVNPFG	GRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESITTSSPPE	VVEI--AVNE	----	EDVNVVAEV*
01--4		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDGHR	PHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AIND	----	VAEV*	
01--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AINE	VDTN	----	VAEV*
01--10		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AIND	----	VAEV*	
01--10		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AIND	----	VAEV*	
01--10		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	VVEI--AIND	----	VAEV*	
01--G		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPHYGNRGRWGENE	SEEKEHPTTESVTTSSPPE	DVVEI--AIND	----	VAEV*	
08--2		GDQDRPMFEMRPF	FRNPFGRKPF	GDRPF	GRRNGTEEGSPRRDG	QRRPYGNRGRWGENE	SEEKEYPTTESVTTSSPP*	-----	-----	-----	

Figure S2. The NJGT-PRANK alignment of representative deduced SpTrf protein sequences increases the number of gaps in the alignment. The NJGT-PRANK alignment was done using webPRANK with the deduced SpTrf protein sequences from animals as in Figure S1. The deduced amino acid sequences of the second exon of the *SpTrf* genes are shown. Standard parameters were used in webPRANK (gap rate, 0.05; gap length, 5; K, 2.0) and the sequences were set to align by translated codons and to trust insertions (+F). Substitution scoring was set to relaxed and the neighbor joining guide tree generation was done in webPRANK using ClustalW2. Explanations of protein names, dashes, and asterisks are the same as in Figure S1. Element borders are based on the repeat-based alignment and are indicated by vertical black lines. Horizontal black lines indicate regions where non-matching elements overlap. Notes on the alignment and elements of interest are shown above those regions. Er is the abbreviation for elements in the repeat-based alignment. The sequences are arranged with the L proteins at the top, the O proteins at the bottom, and the remaining S proteins in alphabetical order

[illegible]

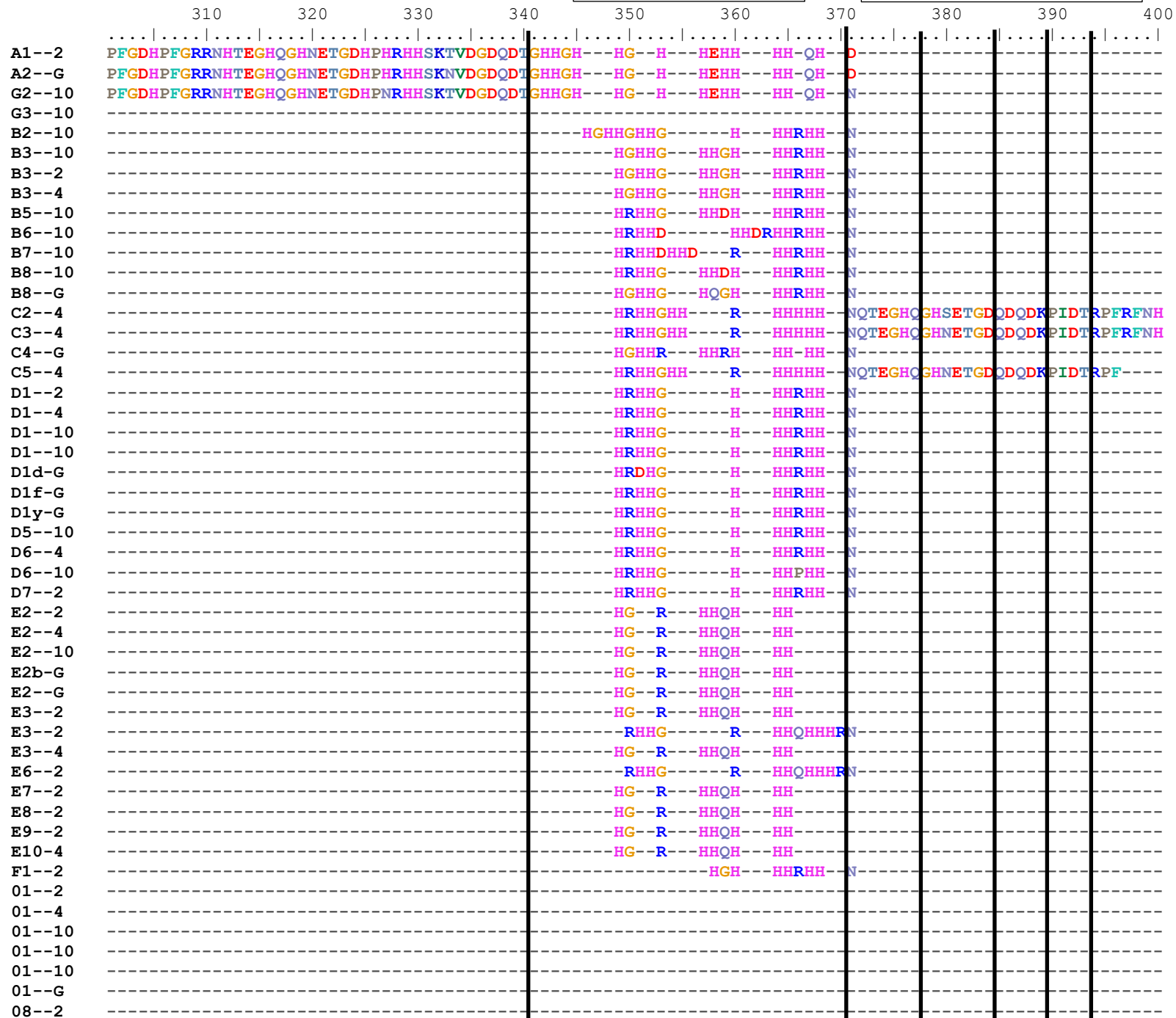
[illegible]

L-Er7 & S-Er8 aligned together

S-Er10 to SEr16 are aligned separately

[illegible]

The C proteins align differently for
Er11, 14, 16,17, 21-23



[illegible]

	510	520	530	540	550	560	570	580	590					
A1---2	RPFWVNPFG	GRKPF	GDRPF	DRRNGTE	EEGSP	RRDG	HPH	PHGNR	GRWGENE	SEEKEHPTTESVTTSSPLK	VEIAIAINEVD	TNV----	VAEV*	
A2---G	RPFWVNPFG	GRKPF	GDRPF	DRRNGTE	EEGSP	RRDG	HPH	PHGNR	GRWGENE	SEEKEHPTTESVTTSSPLK	VEIAIAINEVD	TNV----	VAEV*	
G2---10	RPFWVNPFG	GRKPF	GDRPF	DRRNGTE	EEGSP	RRDG	HPH	PHGNRRR	WGENE	SEEKEHPTTESVTTSSPLK	VEIAIAINEVD	TNV----	VAEV*	
G3---10	-----	FGRKPF	GDRPF	RRNGTE	EEGSP	RRDG	QRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
B2---10	RPFWVNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
B3---10	RPFWVNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
B3---2	RPFWVNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAIND	-----	VAEV*	
B3---4	RPFWVNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
B5---10	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
B6---10	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAFNEED	VNV----	VAEV*	
B7---10	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
B8---10	RPFWVNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRHPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEVL	
B8---G	RPFWVNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRHPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEVY	
C2---4	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
C3---4	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAFNEED	VNV----	VAEV*	
C4---G	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
C5---4	---	RFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*
D1---10	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D1---10	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D1---2	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D1---4	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D1d---G	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D1f---G	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	QRRPY	GNRGR	*	-----	-----	-----	-----	
D1y---G	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D5---10	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D6---4	RPFWVNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEV*	
D6---10	RPFWVNPFG	GRKPF	GDRPF	DRRNGTE	EEGSP	RRDG	HPH	PHGNR	GRWGENE	SEEKEHPTTESVTTSSPLK	VEIAIAINEVD	TNV----	VAEV*	
D7---2	RPFRFNPFG	GRKPF	GGRPF	DRRNGTE	EEGSP	RRDG	HRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	VEIAVNEED	VNV----	VAEVK	
E2---10	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	QRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	-----	-----	-----	
E2---2	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	QRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	*	-----	-----	
E2---4	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	QRRPY	GNRGR	WGENE	SEEKEYPTTESVTTSSPP	*	-----	-----	
E2b---G	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	QRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	*	-----	-----	
E2---G	RPFRFNPFG	GRKPF	GDRPF	GRRNGTE	EEGSP	RRDG	QRRPY	GNRGR	WGENE	SEEKEHPTTESVTTSSPP	*	-----	-----	
E3---2</														

Figure S3. The MLGT-PRANK alignment of representative SpTrf protein sequences with a provided maximum likelihood guide tree results in regions of misaligned sequence. The MLGT-PRANK alignment was done using webPRANK with the deduced protein sequences as described in Figure S1. Explanations of protein names, dashes, and asterisks are the same as in Figure S1. Element borders based on the repeat-based alignment are indicated by vertical black lines. Horizontal black lines indicate regions where non-matching elements overlap. Notes in the alignment and elements of interest are labeled above those regions. Er is used as an abbreviation to indicate elements in the repeat-based alignment. The maximum likelihood tree generated from the manual repeat-based alignment (Figure 1B in the main paper) was provided as a guide tree for this PRANK alignment. Standard parameters were used in webPRANK (gap rate, 0.05; gap length, 5; K, 2.0) and the sequences were set to align by translated codons and to trust insertions (+F). Substitution scoring was set to relaxed.

	10	20	30	40	50	60	70	80	90	100				
A1--2	AHAQSDFNERRK	ENGRER	GD	DRF	GGRPD	GMGGPRQD	DGGPMG	GRRFD	GP	FGA	PMGGPRQD	DGGPMG	GRRFD	DG
A2--G	AHAQSDFNERRK	ENGRER	GD	DRF	GGRPD	GMGGPRQD	DGGPMG	GRRFD	GP	FGA	PMGGPRQD	DGGPMG	GRRFD	DG
G2--10	AHARRDFNELR	GKENG	RER	GQ	GRFG	GRPD	GMGGPRQD	DGGPMG	GRRFD	GP	FGA	PMGGPRQD	DGGPMG	GRRFD
G3--10	AHARRDFNELR	GKENG	RER	GQ	GRFG	GRPD	GMGGPRQD	DGGPMG	GRRFD	GP	FGA	PMGGPRQD	DGGPMG	GRRFD
B2--10	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	DSG	AP	MDGRRQ
B3--10	AHARRDYNE	RRGNEN	GRER	GQ	GRFG	GRPG	GMGGPRQD	DGGPMG	GRRFD	GP	---	---	---	---
B3--2	AHARRDYNE	RRGNEN	GRER	GQ	GRFG	GRPG	GMGGPRQD	DGGPMG	GRRFD	GP	---	---	---	---
B3--4	AHARRDFNE	RRGNEN	GRER	GQ	DRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
B5--10	AHARRDFNE	RRGNEN	GRER	GQ	GRFG	GRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
B6--10	AHAQRD	FNELR	GKENG	RER	GQ	GRFG	GRPG	MLM	GSP	RQD	DGGSM	GRRFD	GP	---
B7--10	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
B8--10	AHARRDFNE	RRGNEN	GRER	GQ	GRFG	GRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
B8--G	AHARRDFNE	RRGNEN	GRER	GQ	GRFG	GRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
C2--4	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
C3--4	AHAQRDYNE	RRGNEN	GRER	GQ	GRFG	GGRPG	QMD	GRRQD	DGGPMG	GRRFD	GP	---	---	---
C4--G	AHARRDFNE	RRGNEN	GRER	GQ	GGFG	GRPD	GM	---	---	---	---	---	---	---
C5--4	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
D1--10	AHAQRDYNE	LRGNEN	GRER	GQ	GRFG	GGRPG	GMGGPRQD	DGGPMG	GRRFD	GP	---	---	---	---
D1--10	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
D1--2	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
D1--4	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
D1d-G	AHAQRDYNE	RRGNEN	GRER	GQ	GRFG	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
D1f-G	AHAQRDYNE	RRGNEN	GRER	GQ	GRFG	GGRPG	MTG	GPRQD	DGGPMG	GRRFD	GP	---	---	---
D1y-G	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
D5--10	AHAQRD	FNELR	GKENG	RER	GQ	GRFG	GRPG	MLM	GSP	RQD	DGGSM	GRRFD	GP	---
D6--4	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
D6--10	AHARRDYNE	RRGNEN	GRER	GQ	GRFG	GGRPG	GMGGPRQD	DGGPMG	GRRFD	GP	---	---	---	---
D7--2	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
E2--10	AHAQRDYNE	RRGNEN	GRER	GQ	GGFG	GGRPG	GMGGPRQD	DGGPMG	GRRFD	GP	---	---	---	---
E2--2	AHAQRDYNE	RRGNEN	GRER	GQ	GGFG	GGRPG	GMGGPRQD	DGGPMG	GRRFD	GP	---	---	---	---
E2--4	AHAERDFNE	RRGNEN	GRER	GQ	GGFG	GGRPG	MTG	SPRQD	DGGPMG	GRRFD	GP	---	---	---
E2b-G	AHARRDFNE	RRGNEN	GRER	GQ	GGFG	GGRPG	MTG	SPRQD	---	---	---	---	---	---
E2--G	AHAQRDYNE	RRGNEN	GRER	GQ	GGFG	GGRPG	QMG	SPRQD	DGGPMG	GRRFD	GP	---	---	---
E3--2	AHAQRDYNE	LRGNK	ENGR	ERG	QGRF	GGRPG	GMGGS	RQD	DGGPMG	GRRFD	GP	---	---	---
E3--2	AHARRDFNE	RRGNEN	GRER	GQ	GGFG	GGRPG	MTG	SPRQD	DGGPMG	GRRFD	GP	---	---	---
E3--4	AHAQRDYNE	RRGNEN	GRER	GQ	GGFG	GGRPG	GMGGPRQD	DGGPMG	GRRFD	GP	---	---	---	---
E6--2	AHAQRDYNE	LRGNK												

	110	120	130	140	150	160	170	180	190	200			
A1--2	PGFGA	PMGGPRQ	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGEGGRRDGEETD	AAQQIGDGLGGPGRCDGP	GHGHYGHHQAGR	FFGNPPP	FNPEQ					
A2--G	PGFGA	PMGGPRQ	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGEGGRRDGEETD	AAQQIG	-----	PGRFDGP	GHGHYGHHQAGR	FFGNPPP	FNPEQ			
G2--10	PGFGA	PMGGPRQ	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGEGGRRDGEETD	AAQQIGDGLGGPGRFDGP	GHGHYGHHQAGR	FFGNPPP	FNPEQ					
G3--10	PGFGA	PMGGPRQ	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGEGGRRDGEETD	AAQQIGDGLGGPGRFDGP	GHGHYGHHQAGR	FFGNPPP	FNPEQ					
B2--10	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
B3--4	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRGR	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
B3--10	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
B3--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
B5--10	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
B6--10	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
B7--10	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGS	DRFDGLRRGH	RQGP	-----	PQDRPEEQ	PF				
B8--10	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
B8--G	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	ASQQIGDGLGGRQ	QFDGHGRRHGH	RQGP	-----	PQDRPEEQ	PF				
C2--4	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGS	DRFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF		
C3--4	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGS	DRFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF			
C4--G	-----	GGRRFDGPGFGGSR	PDGAG--R	PFFGQGGRRDGEETD	AAQQIGDGLGGS	DRFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF		
C5--4	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGS	DRFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF			
D1--10	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PVGAGGR	LFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF	
D1--10	PGIGA	EMDGT	EKNGGPMGGR	ELDGPFGGSR	PDGGGGR	PFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF
D1--2	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF		
D1--4	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF		
D1d-G	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PVGAGGR	PFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF	
D1f-G	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PVGAGGR	PVFQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF	
D1y-G	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF		
D5--10	P-----	-----	CGGGR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF		
D6--4	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF		
D6--10	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF		
D7--2	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDQAEQ	PF		
E10--4	PGFGA	PMGGPRQ	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGEGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF		
E2--10	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	DDQKMGDGP	GRGQFDGH	GRGHGH	RQGP	-----	PQDRPEEQ	PF			
E2--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRD	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF				
E2--4	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQLIGDGLGGRQ	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF				
E2b-G	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGSQ	QFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF			
E3--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGSQ	QFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF			
E3--4	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGPGGRRDGEETD	AALLIGDVLGGRQ	QFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF			
E3--2	P-----	-----	CGGGR	PDGAGGRPFFGQGGRRDGEETD	AAQQMGDGLGGRQ	QFDGPRR	GHGH	RQGP	-----	PQDRPEEQ	PF		
E6--2	PGFGA	EMDGR	QNGGPMGRRFDGPGFGGSR	PVGAGGR	PFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF		
E7--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF				
E8--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGPGRQ	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF				
E9--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGPGRQ	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF				
F1--2	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF				
E2--G	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGRQ	QFDGHGRGHGH	RQGP	-----	PQDRPEEQ	PF				
01--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRRH	-----	-----	-----	-----	-----			
01--4	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRRH	-----	-----	-----	-----	-----			
01--10	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRRH	-----	-----	-----	-----	-----			
01--10	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRRH	-----	-----	-----	-----	-----			
01--10	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRRH	-----	-----	-----	-----	-----			
01--G	-----	NGGPMGRRFDGPGFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRRH	-----	-----	-----	-----	-----			
08--2	-----	NGGPMGRRFDGPRFGGSR	PDGAGGRPFFGQGGRRDGEETD	AAQQIGDGLGGPQ	QFDGPRRH	-----	-----	-----	-----	-----			

L-Er10 and Er11 aligned together

Er13

Er14

Er15

10

[illegible]

	310	320	330	340	350	360	370	380	390	400
A1--2	GDQDTGHHGHHGHHEHHHHQH	HDHREGHQDHD	DRPMFGMRPF	FRFNPFGRKPF	GDHPFGFRNHT	EGHQGHNETG	DHPRHHSKT	GDGQ	QDRPMFETRPF	FWVNP
A2--G	GDQDTGHHGHHGHHEHHHHQH	HDHREGHQDHD	DRPMFEMRPF	FRFNPLGRKPF	GDHPFGFRNHT	EGHQGHNETG	DHPRHHSKT	GDGQ	QDRPMFETRPF	FWVNP
G2--10	GDQDTGHHGHHGHHEHHHHQH	NHREGHQDHD	DRPMFGMRPF	FRFNPFGRKPF	GDHPFGFRNHT	EGHQGHNETG	DHPRHHSKT	GDGQ	QDRPMFETRPF	FWVNP
G3--10	-----P									
B2--10					FG	RNHTEGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFEMRPF	FWVNP
B3--10					FG	RNHTEGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFETRPF	FWVNP
B3--2					FG	RNHTEGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFETRPF	FWVNP
B3--4					FG	RNHTEGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFETRPF	FWVNP
B5--10					FG	RNHTEGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF	FRFNP
B6--10					--DHPRHHNKT					
B7--10					--DHPRHHNKT					
B8--10					FG	RNHTEGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFEMRPF	FWVNP
B8--G					FG	RNHTEGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFEMRPF	FWVNP
C2--4					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFESRPF
C3--4					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFESRPF
C4--G					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFESRPF
C5--4					-----P					
D1--10					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
D1--10					FG	RKPFGRDRTFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
D1--2					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
D1--4					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
D1d-G					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFESRPF
D1f-G					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
D1y-G					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
D5--10					FG	RKPFGRDRTFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
D6--4					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFEMRPF
D6--10					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHSKT	GDGQ	QDRPMFETRPF
D7--2					FG	RKPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
E2--10					--RNRT					
E2--2					--RNHT					
E2--4					--RNHT					
E2b-G					--RNHT					
E2--G					--RNHT					
E3--2					--RNHT					
E3--2					--RNHT					
E3--4					--RNHT					
E6--2					--RNHT					
E7--2					--RNHT					
E8--2					--RNRT					
E9--2					--RNRT					
E10-4					--RNHT					
F1--2					FG	RNHTEGHQGHNETG	DHPRHHNKT	GDGQ	QDRPMFEMRPF	FWVNP
01--2					GR	KPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
01--4					GR	KPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
01--10					GR	KPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
01--10					GR	KPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
01--10					GR	KPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
01--G					GR	KPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF
08--2					GR	KPFGRDPFGFRNHT	EGHQGHNETG	DHPRHHNKT	RDGQ	QDRPMFEMRPF

410 420 430 440 450 460 470

A1--2 FGRKPF GDRPF DRRNGT EESSPRRDGH PPHGNRGRWGGENESEEKEHPTTESVTTSSPLK VIEIAINEVD TNVVAEV

A2--G FGRKPF GDRPF DRRNGT EESSPRRDGH PPHGNRGRWGGENESEEKEHPTTESVTTSSPLK VIEIAINEVD TNVVAEV

G2--10 FGRKPF GDRPF DRRNGT EESSPRRDGH PPHGNRRRWGGENESEEKEHPTTESVTTSSPLK VIEIAINEVD TNVVAEV

G3--10 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEV

B2--10 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

B3--10 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEV

B3--2 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAIND----VAEV

B3--4 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEV

B5--10 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

B6--10 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAFNEEDVHVVAEV

B7--10 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKERPTTESVTTSSPPEVVEIAFNEEDVNVVAEV

B8--10 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEVL

B8--G FGRKPF GDRPF GRRNRTEESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEVL

C2--4 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEV

C3--4 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAFNEEDVNVVAEV

C4--G LGRKPF GDRPF GRRNGT EESSLRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEV

C5--4 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEV

D1--10 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

D1--10 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

D1--2 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

D1--4 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

D1d-G FGRKPF GDRLF GRRNGT EESSSRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

D1f-G LGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGR*-----

D1y-G FRRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV*

D5--10 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

D6--4 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAVNEEDVNVVAEV

D6--10 FGRKPF GDRPF DRRNGT EESSPRRDGH PPHGNRGRWGGENESEEKEHPTTESVTTSSPLK VIEIAINEVD TNVVAEV

D7--2 FGRKPF GGRPF DRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVV--AINEED INVVAEV

E10--4 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E2--10 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E2--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E2--4 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEYPTTESVTTSSP*-----

E2b-G FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E2--G FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E3--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E3--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E3--4 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

E6--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEYPTTESVTTSSP*-----

E7--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVAINNEED INVVAEV---

E8--2 FGRKPF GDRPF DRRNGT EESSPRRDGH PPHGNRGRWGGENESEEKEHPTTESVTTSSPLK VIEIAINEVD TNVVAEV

E9--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSP*-----

F1--2 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESITTSSPPEVVEIAVNEEDVNVVAEV

01--10 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAINDV----AEV

01--10 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAINDV----AEV

01--10 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAINDV----AEV

01--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAINEVD TNVVAEV

01--4 FGRKPF GDRPF GRRNGT EESSPRRDGHRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAINDV----AEV

01--G FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEHPTTESVTTSSPPEVVEIAINDV----AEV

08--2 FGRKPF GDRPF GRRNGT EESSPRRDGQRRPYGNRGRWGGENESEEKEYPTTESVTTSSP*-----

Figure S4. The ClustalW alignment without manual correction includes a non-matching region of the deduced SpTrf proteins. The alignment was done using ClustalW in BioEdit with the deduced protein sequences from animals as in Figure S1. Standard parameters were used with full multiple alignment mode. Substitution scoring was set to relaxed. Explanations of protein names, dashes, and asterisks are the same as for Figure S1. Element borders based on the repeat-based alignment are indicated by vertical black lines. Horizontal black lines indicate regions where non-matching elements overlap. Notes on the alignment and elements of interest are labeled above those regions. Er indicates elements in the repeat-based alignment.

References

1. Cameron RA, Mahairas G, Rast JP, Martinez P, Biondi TR, Swartzell S, et al. A sea urchin genome project: Sequence scan, virtual map, and additional resources. *Proc Natl Acad Sci USA*. (2000) 97(17):9514–8. doi:10.1073/pnas.160261897.
2. Oren M, Barela Hudgell MA, D’Allura B, Agronin J, Gross A, Podini D, et al. Short tandem repeats, segmental duplications, gene deletion, and genomic instability in a rapidly diversified immune gene family. *BMC Genomics*. (2016) 17(1):900. doi:10.1186/s12864-016-3241-x.
3. Barela Hudgell MA, Smith LC. Sequence diversity, locus structure, and evolutionary history of the *SpTransformer* genes in the sea urchin genome. *Front Immunol*. (2021) 12:744783. doi:10.3389/fimmu.2021.744783.