

Supplemental Materials

Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP)

Toshitsugu Fujita, Yoshinori Asano, Junko Ohtsuka, Yoko Takada, Kazunobu Saito,
Rieko Ohki, Hodaka Fujii

Supplemental Methods

Supplemental Figure Legends

Supplemental References

Supplemental Figures 1 - 7

Supplemental Table 1

Supplemental Methods

enChIP-PCR. For detection of γ -satellite sequence, enChIP was performed as described in enChIP-Southern blot analysis except that ChIP DNA Clean & Concentrator (Zyme Research) was used for purification of DNA. The purified DNA was used for PCR with AmpliTaq Gold 360 Master Mix (Applied Biosystems). PCR cycles were as follows: denaturing at 95°C for 10 min; 17 cycles of 95°C for 15 sec, 60°C for 1 min. The primers used in this experiment are 5'-gcgagaaaactgaaaatcac-3' and 5'-tcaagtgcgtcaagtggatg-3¹. The PCR amplicons were subjected to electrophoresis in 2% agarose gel and visualized in Gel Doc XR (Bio-Rad). The intensity of visualized bands was analyzed using Scion Imaging Software (Scion Corporation).

Identification of proteins by LC-MS/MS. After staining with Coumassie Brilliant Blue, the each lane was cut into 5 slices. Each slice was digested with trypsin, and the obtained peptides were dried and then dissolved in 0.1% trifluoroacetic acid, 2% acetonitrile prior to LC-MS/MS analysis. Peptides were analysed using a nanoLC-MS/MS system, composed of an LTQ Orbitrap Velos (Thermo Fisher Scientific) coupled with a nanoLC (Advance, Michrom BioResources) and an HTC-PAL autosampler (CTC Analytics). Peptide separation was carried out using the C18 reversed phase analytical column (0.1 mm ID x 15 cm, 3 μ m resin, L-column Micro; CELI). The mobile phases consisted of 0.1% formic acid and 100% acetonitrile. Peptides were eluted by a gradient of 5–35% acetonitrile for 105 min at a flow rate of 500 nl/min. CID spectra were acquired automatically in the data-dependent scan mode with the dynamic exclusion option. Full MS was obtained by Orbitrap in the range of m/z 300–1,500 with resolution 30,000. The nine most intense precursor ions in the full MS spectra were selected for subsequent MS/MS analysis

in an ion trap with the automated gain control mode. The lock mass function was activated to minimize mass error during analysis.

Database searching. Tandem mass spectra were extracted, then charge state was deconvoluted and deisotoped by Proteome Discoverer version 1.2. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.4.1) and X! Tandem (The GPM, thegpm.org; version CYCLONE (2010.12.01.1)). X! Tandem was set up to search the SwissProt_2012 database (535248 entries) also assuming the digestion enzyme as trypsin. Mascot was set up to search the SwissProt_2013_08 database (selected for *Mus musculus*, 16625 entries) assuming the digestion enzyme as trypsin. Mascot and X! Tandem were searched with a fragment ion mass tolerance of 0.80 Da and a parent ion tolerance of 10.0 PPM. Iodoacetamide derivative of cysteine was specified in Mascot and X! Tandem as a fixed modification. S-carbamoylmethylcysteine cyclization (N-terminus) of the n-terminus, oxidation of methionine and acetylation of the n-terminus were specified in Mascot and X! Tandem as variable modifications.

Criteria for protein identification. Scaffold (version Scaffold_3.4.5, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm². Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm³. These thresholds resulted in a protein false discovery rate (FDR) of < 0.1% as calculated by Scaffold. Proteins that

contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. MS data were converted using PRIDE Converter 2 Tool Suite (version 2.0.19)⁴ and deposited in the PRIDE database (<http://www.ebi.ac.uk/pride>)⁵⁻⁷ under accession numbers [31251-31256].

Quantification using spectral counting. Label-free quantification of relative protein abundance was performed by spectral counting⁸ using all spectra matched to a peptide sequence. Relative protein abundance was calculated on the basis of the unweighted spectral count assigned to each identified protein by Scaffold. The unweighted spectral count includes spectra matched to peptides shared between multiple proteins if there is independent evidence that these proteins are present. To normalize the data, spectral counts were expressed as a percentage of the total number of spectra expressed as a percentage of the total number of spectra observed in the entire sample. sd and p-values were calculated from three biological replicates. p-values were calculated by t-test.

Supplemental Figure Legends

Supplemental Figure 1. Nucleotide and amino acid sequences of 3xFN-Tel-TAL. Sequences of 3xFLAG-tag, V5 epitope-tag and NLS are highlighted.

Supplemental Figure 2. The full-length blot of Figure 2b including molecular size marker.

Supplemental Figure 3. The full-length blot of Figure 2d including molecular size marker.

Supplemental Figure 4. The amounts of γ -satellite repeats in enChIP samples. γ -satellite repeats in enChIP samples were quantified by PCR.

Supplemental Figure 5. SDS-PAGE and CBB staining. The stained regions were divided into 5 parts (2-4 mm height each), excised, and subjected to in-gel tryptic digestion. The digested peptides were analyzed in LC-MS/MS.

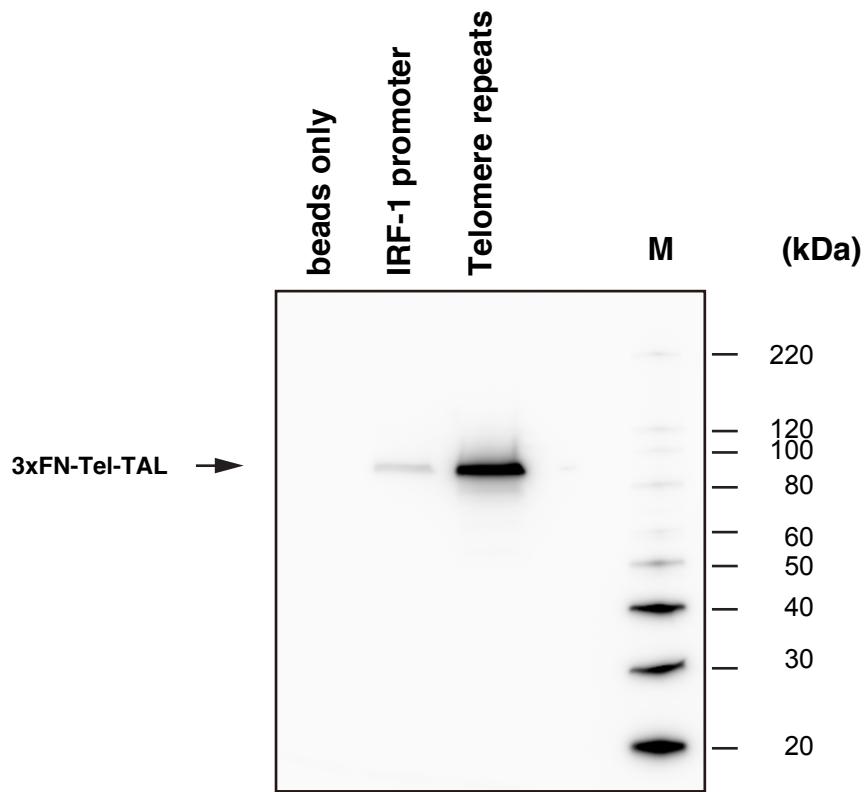
Supplemental Figure 6. Localization of candidate proteins at telomeres. U2OS cells were transfected with the expression vector of ECFP-TRF2 together with those of Halo-tagged proteins. Cells were treated with TMR ligand (5 μ M) for 15 min at 37°C 23 h post transfection, fixed with 4% paraformaldehyde, and subjected to immunofluorescence microscopy. Thirty cells expressing both ECFP-TRF2 and Halo-tagged proteins were randomly picked up and analyzed as Figure 2.

Supplemental Figure 7. The full-length gel image of Figure 4.

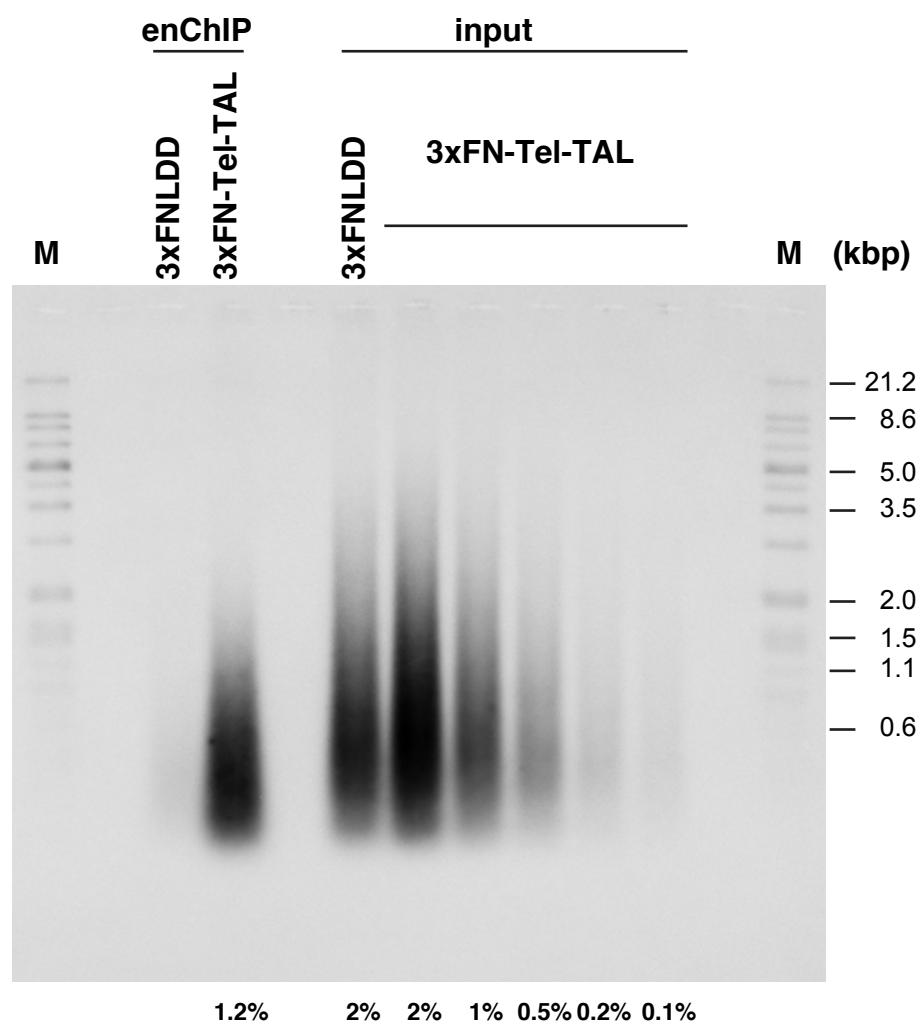
Supplemental References

1. Kishi, Y., Kondo, S. & Gotoh, Y. Transcriptional activation of mouse major satellite regions during neuronal differentiation. *Cell Struct. Funct.* **37**, 101-110 (2012).
2. Keller, A., Nesvizhskii, A.I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**, 5383-5392 (2002).
3. Nesvizhskii, A.I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* **75**, 4646-4658 (2003).
4. Barsnes, H., Vizcaíno, J.A., Eidhammer, I. & Martens, L. PRIDE Converter: Making proteomics data-sharing easy. *Nat. Biotechnol.* **27**, 598-599 (2009).
5. Vizcaíno, J.A. et al. The Proteomics Identifications database: 2010 update. *Nucleic Acids Res.* **38**, D736-D742 (2010).
6. Old, W.M. et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* **4**, 1487-1502 (2005).
7. Zyballov, B., Coleman, M.K., Florens, L. & Washburn, M.P. Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal. Chem.* **77**, 6218-6224 (2005).
8. Liu, H.B., Sadygov, R.G. & Yates, J.R. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **76**, 4193-4201 (2004).

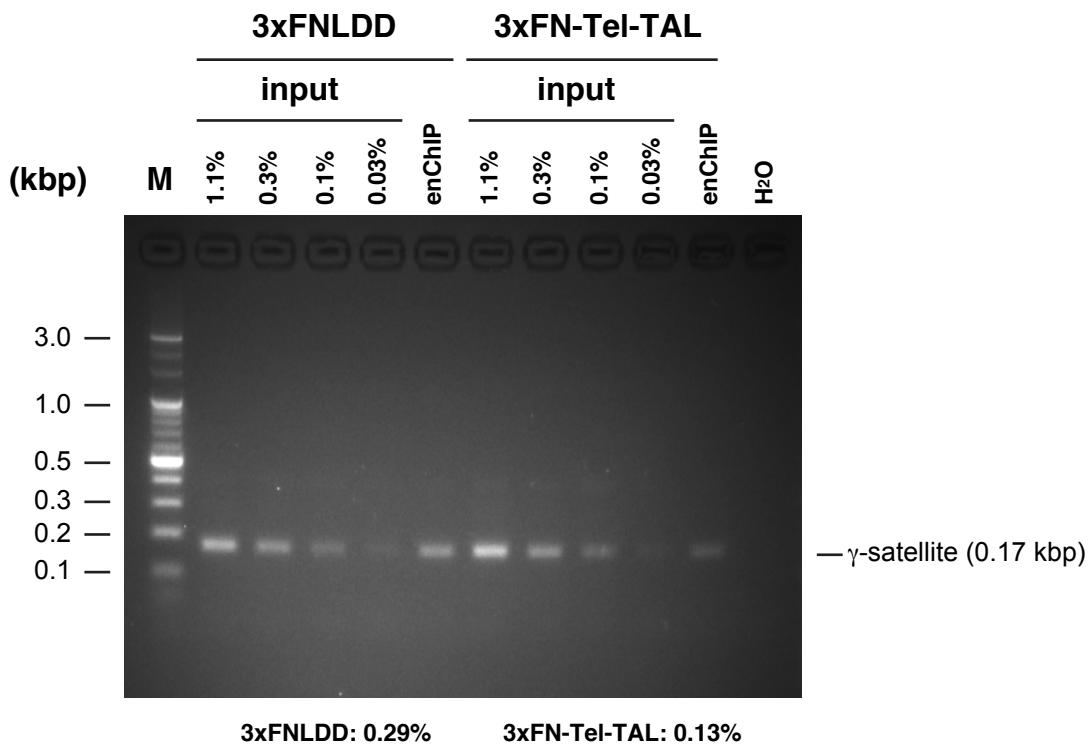
Supplemental Figure 2



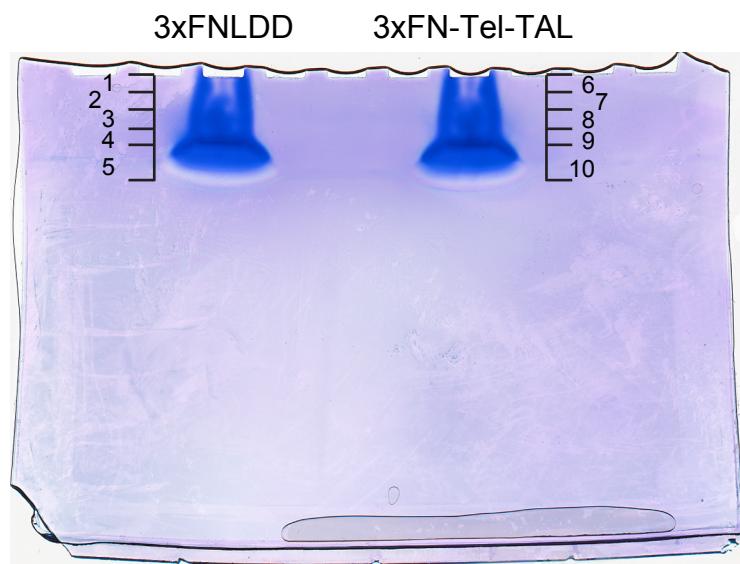
Supplemental Figure 3



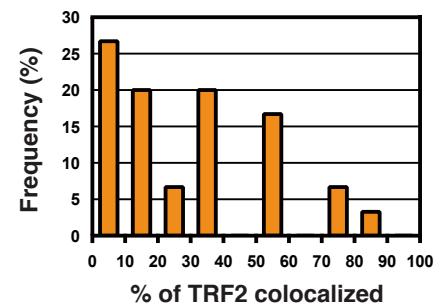
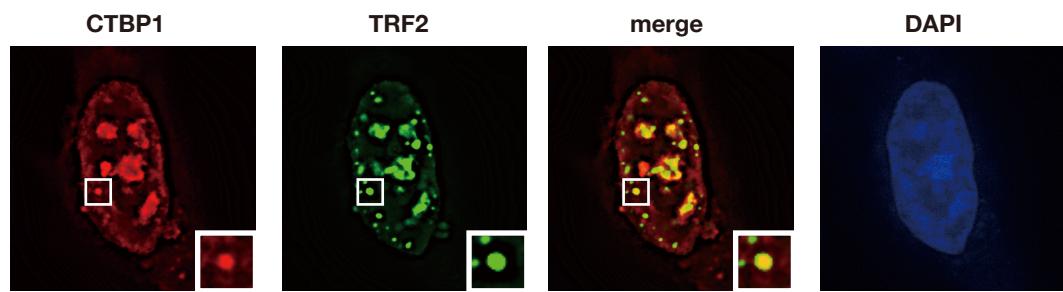
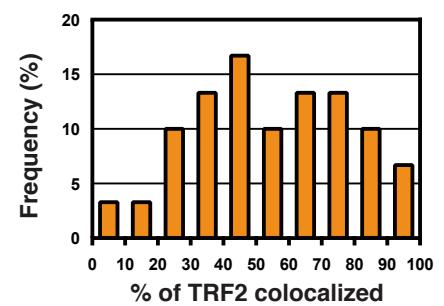
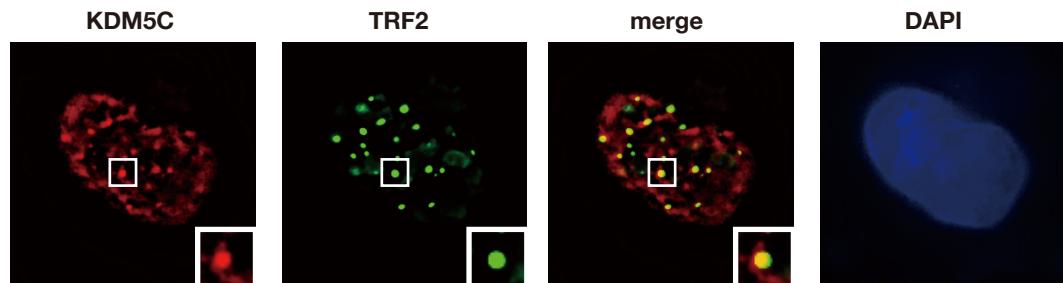
Supplemental Figure 4



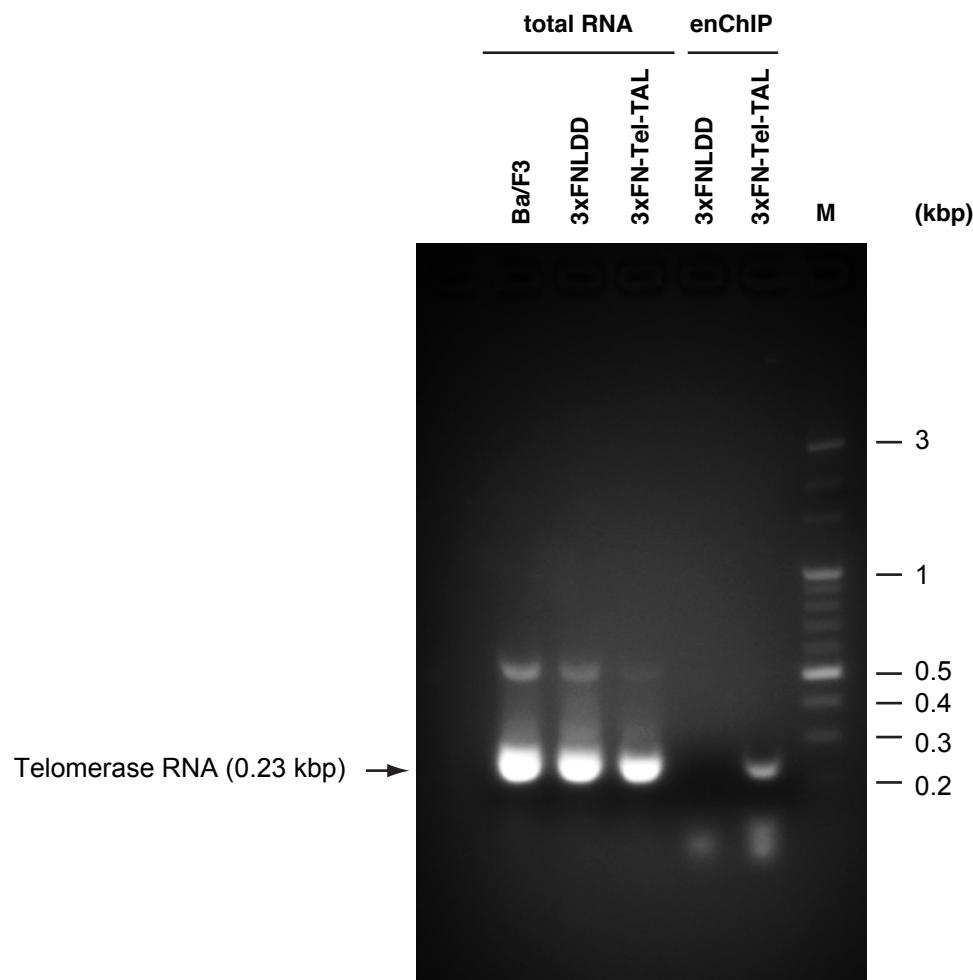
Supplemental Figure 5



Supplemental Figure 6



Supplemental Figure 7



740	26S proteasome non-ATPase regulatory subunit 1 OS-Mus musculus GN-Punrl PE=1 SV=1	PUNRL MOUSE	100 kDa	1.57	2.73	0	0	0		0%	(0.037)
741	T-complex protein 1 subunit epsilon OS-Mus musculus GN-Cct5 PE=1 SV=1	TCP5 MOUSE	60 kDa	1.51	2.73	0	0	0		0%	(0.037)
742	Staphylococcal nucleic acid-binding protein I OS-Mus musculus GN-Snbd PE=1 SV=1	SNBD MOUSE	102 kDa	1.6	1.45	0	0	0		0%	(0.13)
743	Protein SOR OS-Mus musculus GN-Sor PE=1 SV=2	SOR MOUSE	266 kDa	1.62	1.5	0	0	0		0%	(0.14)
744	Transcriptional repressor CTCF OS-Mus musculus GN-Ctfb PE=1 SV=2	CTCF MOUSE	84 kDa	1.64	1.47	0	0	0		0%	(0.13)
745	Nucleolar protein 10 OS-Mus musculus GN-Nucp10 PE=1 SV=1	NUCP10 MOUSE	100 kDa	1.65	2.85	0	0	0		0%	(0.37)
746	Microtubule-associated protein 2 OS-Mus musculus GN-Tacop2 PE=1 SV=1	TACO2 MOUSE	65 kDa	1.65	2.85	0	0	0		0%	(0.37)
747	Nascent polypeptide-associated complex subunit alpha, muscle-specific form OS-Mus musculus GN-Naca PE=1 SV=1	NACAM MOUSE	220 kDa	1.65	2.85	0	0	0		0%	(0.37)
748	Importin subunit alpha-2 OS-Mus musculus GN-Kipn2 PE=1 SV=2	KIPN2 MOUSE	98 kDa	1.74	3.01	0	0	0		0%	(0.37)
749	Keratin, type II cytoskeletal 4 OS-Homo sapiens GN-Krt4 PE=1 SV=1	KRT4 HUMAN	57 kDa	1.74	3.01	0	0	0		0%	(0.37)
750	Putative protein OS-Mus musculus GN-Pnp1 PE=2 SV=2	PNP1 MOUSE	44 kDa	1.8	3.42	0	0	0		0%	(0.37)
751	Put and SPRR-interacting protein OS-Mus musculus GN-Prrsp1 PE=1 SV=2	PRRSP1 MOUSE	100 kDa	1.86	1.73	0	0	0		0%	(0.10)
752	Keratin, type I cuticular Ha1 OS-Homo sapiens GN-Krt14 PE=2 SV=2	KRT14 HUMAN	49 kDa	2.08	3.61	0	0	0		0%	(0.37)
753	HRNA (quartet26-N(2'))-dimethyltransferase OS-Mus musculus GN-Tm1 PE=1 SV=2	TM1 MOUSE	72 kDa	2.2	3.82	0	0	0		0%	(0.37)
754	Ribosome biogenesis regulatory protein homolog OS-Mus musculus GN-Rrs1 PE=2 SV=1	RRS1 MOUSE	42 kDa	2.28	2.49	0	0	0		0%	(0.19)
755	Serine/threonine-protein kinase PRKA homolog OS-Mus musculus GN-Prkab PE=1 SV=3	PRKAB MOUSE	117 kDa	2.31	0.59	0	0	0		0%	(0.0022)
756	Protein kinase C theta, delta, epsilon, zeta, eta, tau, iota, tau-like protein I OS-Mus musculus GN-Smkn1 PE=1 SV=1	SMKN1 MOUSE	208 kDa	2.36	4.45	0	0	0		0%	(0.37)
757	Keratin, type I cuticular Ha1 OS-Homo sapiens GN-Krt13 PE=2 SV=3	KRT13 HUMAN	47 kDa	2.78	4.82	0	0	0		0%	(0.37)
758	Keratin, type II cuticular Ha1 OS-Homo sapiens GN-Krt81 PE=1 SV=3	KRT81 HUMAN	65 kDa	2.78	4.82	0	0	0		0%	(0.37)
759	Probable ATP-dependent RNA helicase DDX47 OS-Mus musculus GN-Ddx47 PE=2 SV=2	DDX47 MOUSE	91 kDa	3	2.69	0	0	0		0%	(0.13)
760	Phosphotriesterase family/glycosaminoglycan synthase OS-Mus musculus GN-Pte1 PE=2 SV=2	PTE1 MOUSE	145 kDa	2.81	2.61	0	0	0		0%	(0.13)
761	Protein kinase C theta, delta, epsilon, zeta, eta, tau, iota, tau-like protein A OS-Mus musculus GN-Erta PE=1 SV=5	ERTA MOUSE	100 kDa	3.21	1.54	0	0	0		0%	(0.0041)
762	Keratin, type II cuticular Ha1 OS-Homo sapiens GN-Krt86 PE=1 SV=1	KRT86 HUMAN	55 kDa	3.47	6.02	0	0	0		0%	(0.37)
763	SAFB-like transcription modulator OS-Mus musculus GN-Stm PE=1 SV=1	SLTM MOUSE	117 kDa	4.1	4.17	0	0	0		0%	(0.16)
764	Keratin, type I cuticular Ha1 OS-Mus musculus GN-Krt11 PE=2 SV=2	KRT11 HUMAN	47 kDa	4.17	7.22	0	0	0		0%	(0.37)
765	Polymerase delta-interacting protein 3 OS-Mus musculus GN-Pold3 PE=2 SV=1	POLD3 MOUSE	46 kDa	4.4	3.88	0	0	0		0%	(0.12)
766	Importin subunit beta OS-Mus musculus GN-Ibp1 PE=1 SV=1	IBP1 MOUSE	55 kDa	4.43	4.19	0	0	0		0%	(0.13)
767	Acetylacetate polymerase 2, isoforms beta, delta, gamma, gamma OS-Mus musculus GN-Trepo PE=1 SV=4	LAPB2B MOUSE	100 kDa	4.52	7.82	0	0	0		0%	(0.37)
768	Apoptotic chromatin condensation inducer in the nucleus OS-Mus musculus GN-Asn1 PE=1 SV=3	ANCR1 MOUSE	151 kDa	4.81	5.95	0	0	0		0%	(0.23)
769	Keratin, type II cuticular Ha5 OS-Mus musculus GN-Krt85 PE=2 SV=2	KRT85 MOUSE	96 kDa	5.45	5.23	0	0	0		0%	(0.15)
770	Heterochromatin protein 1-binding protein 3 OS-Mus musculus GN-Hbp3 PE=1 SV=1	HBP3 MOUSE	81 kDa	6.82	4.86	0	0	0		0%	(0.072)
771	Importin subunit beta OS-Mus musculus GN-Ibp2 PE=2 SV=2	IBP2 MOUSE	50 kDa	8.8	7.25	0	0	0		0%	(0.0048)
772	Histone H2A type I OS-Mus musculus GN-H2afnb1 PE=1 SV=3	H2AFNB1 MOUSE	15 kDa	15.43	35.11	0	0	0		0%	(0.37)
773	Histone H2A type I OS-Mus musculus GN-H2afnb1 PE=1 SV=3	H2AFNB1 MOUSE	14 kDa	42.98	38.84	0	0	0		0%	(0.13)
774	LexA repressor OS-Escherichia coli Q139H26 (strain E2437A - ETEC) GN-lexA PE=3 SV=1	LEXA ECO24 (>26) MOUSE	85.59	33.97	0	0	0	0xPNLDO		0%	(0.012)