Supplementary Materials

Supplementary Tables

(See separate Excel file.)

Supplementary Table 1. T2T genome assemblies used in this study.

Supplementary Table 2. Statistics of total isteins and tandem-repeat isteins per genome.

Supplementary Table 3. Top 1,000 isteins by copy number in each genome.

Supplementary Table 4. Statistics of MESSS isteins and istORFs in Homininae species.

Supplementary Table 5. Conserved isteins and their genomic distribution at each taxonomic level.

Supplementary Table 6. Genomic coordinates and annotation of the 179 isteins conserved across primates and mouse.

Supplementary Table 7. dN/dS ratios of the 179 conserved isteins.

Supplementary Table 8. Examples of high-copy-number isteins annotated as bacterial hypothetical proteins.

Supplementary Table 9. Accession numbers for mass spectrometry datasets.

Supplementary Table 10. Accession numbers for ribo-seq datasets.

Supplementary Files

All supplementary files can be found at: https://osf.io/v2hqy/files/. If files cannot be downloaded by clicking the links in the main text, please browse them manually under this URL.

Subfolders:

- ~/AlphaFold
- ~/conserved_isteins
- ~/copy_numbers
- ~/istein
- ~/MESSS
- ~/Proteomics
- ~/Resource
- ~/Tandem_rep_istein
- ~/codes

Supplementary Figure Legends

Supplementary Fig. 1. Extraction of isteins and istORFs by 6-frame *in silico* translation. (A) Schematic of istein and istORF definition. (B) Length distribution of isteins (93–143 aa).

Supplementary Fig. 2. Chromosomal distribution of istORFs.

Supplementary Fig. 3. Complexity distribution of isteins in T2T genomes and the hg38 genome.

Supplementary Fig. 4. Copy number distribution of the top 50 isteins in each primate genome.

Supplementary Fig. 5. Phylogenetic trees based on Jaccard distance-derived ANI. (A) Based on annotated proteins. (B) Based on isteins.

Supplementary Fig. 6. AlphaFold-predicted structures of peptides encoded by the canonical Sat2/3 sequences. (A) Heavy strand. (B) Light strand. Amino acid sequences are provided. (C) Predicted structure of a 162-aa MESSS isoform from chr1.

Supplementary Fig. 7. Repetitiveness profiles of chr1 and chrY in human YAO, hg38, bonobo (PanPan), and gorilla (GorGor).

Supplementary Fig. 8. Amino acid (A) and nucleotide (B) compositions of Sat2/3-derived isteins from different genomic regions.

Supplementary Fig. 9. Correlation between 5-mers identified by mass spectrometry and istein-specific 5-mers at the genomic level. (A) Top 1,000 isteins. (B) MESSS isteins within the top 1,000.

Supplementary Fig. 10. Preparation of MESSS antiserum and prokaryotic expression. (A) MESSS1_446 amino acid sequence, selected immunogenic peptides, and codon-optimized CDS design. (B) ELISA titers of mouse antisera. (C) Western blot detection of recombinant MESSS expressed in *E. coli* using antiserum HNM3-2.

Supplementary Fig. 11. Overexpression and knockdown of MESSS in 293T cells. (A) Anti-Flag Western blot detection of MESSS overexpression. (B) Detection of overexpressed MESSS using HNM3-2 antiserum. (C) CCK-8 assay assessing the effect of MESSS overexpression on cell viability and proliferation. (D) Performance of siRNAs in 293T cells.

Supplementary Fig. 12. Knockdown of MESSS in MDA-MB-231 cells. (A) Morphological changes observed by light microscopy. (B) Annexin-V/PI staining flow cytometry analysis of cell death. (C) Colony formation assay.

Supplementary Fig. 13. Immunohistochemical detection of MESSS expression in liver, small intestine, lung, and ovarian cancer tissues using mouse antiserum.

Supplementary Fig. 14. MEW99 signature peptides detected across datasets using the PepQuery web interface.

Supplementary Fig. 15. AlphaFold-predicted structures of NAMESKL (A) and MEW99 (B).

Supplementary Fig. 1

A

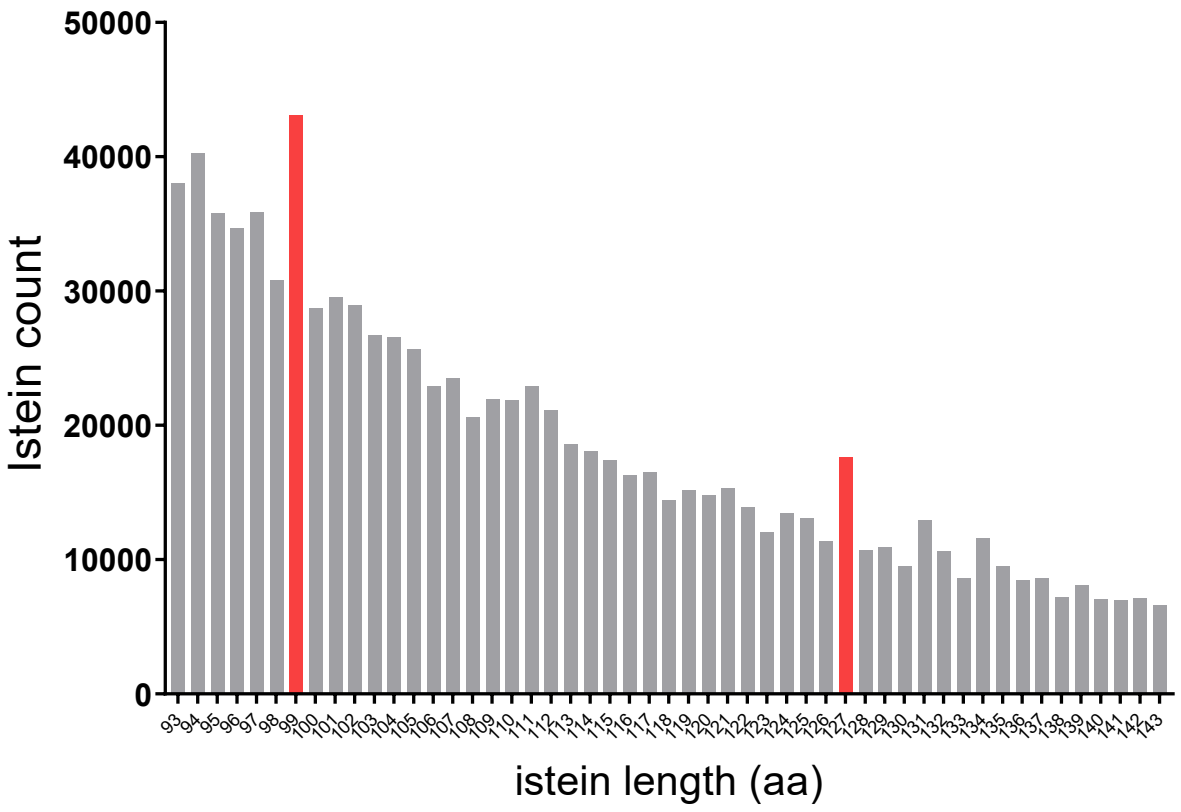
istORF:

AGTGGAGTGGAGTGAAGTGGAGTGGAGTGGAGAGGAGTGGAATGTAGTG
GAATGGAACTGGATTCAATGGAATGGAATGGAATGGACTCGAATGGAAT
GGACTCGAATGGAATGGAAGGGACTTGAATGGAATCATCATCG

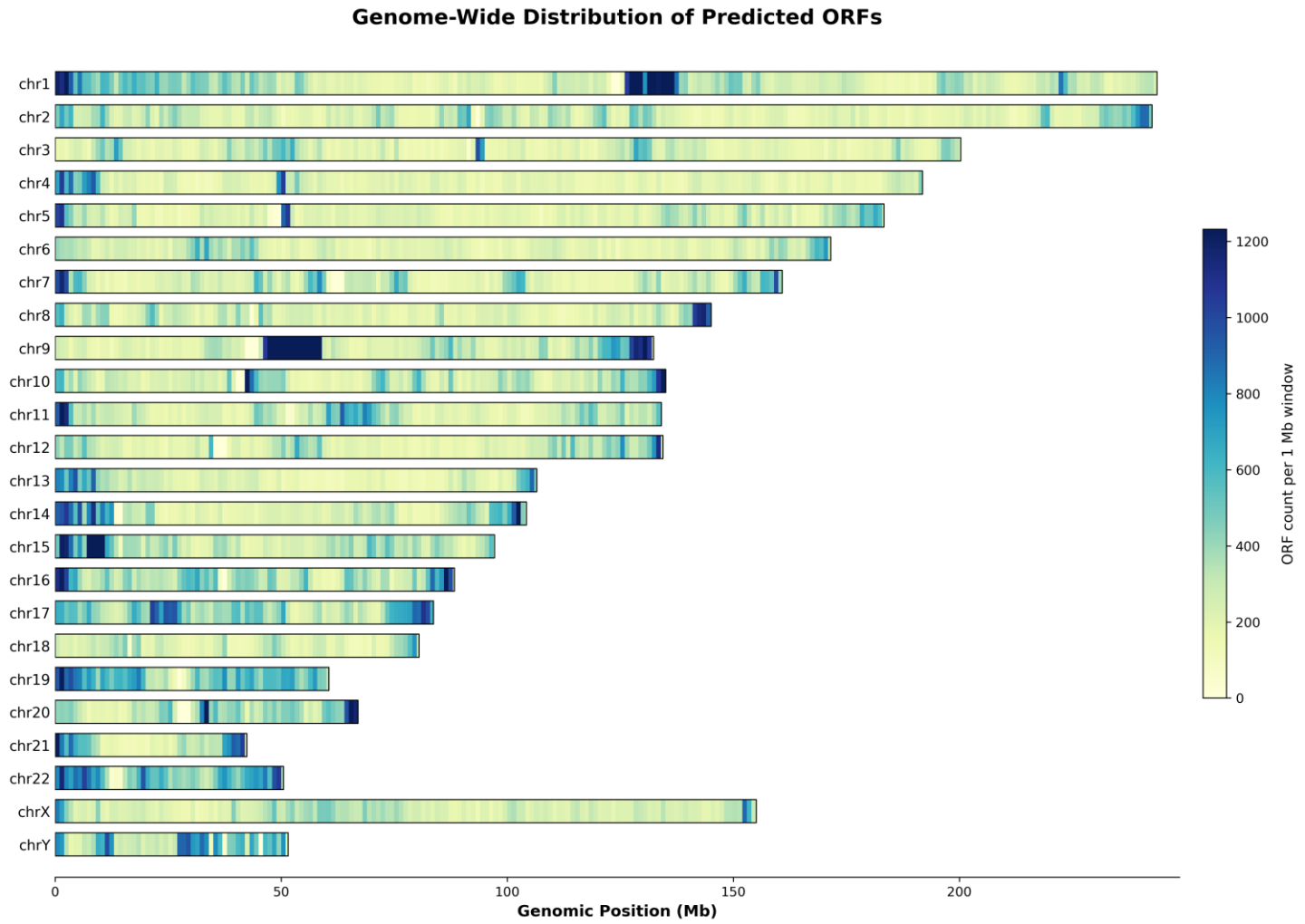
istein:

SGVEZSGVEWRGVECSGMELDSMEWNGMDSNGMDSNGMEGTZMESSS

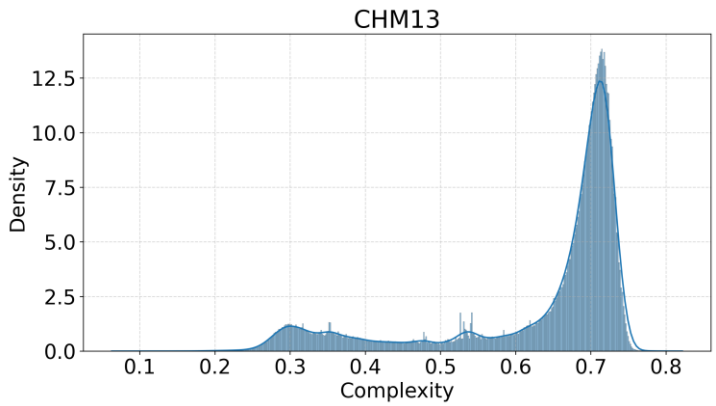
B

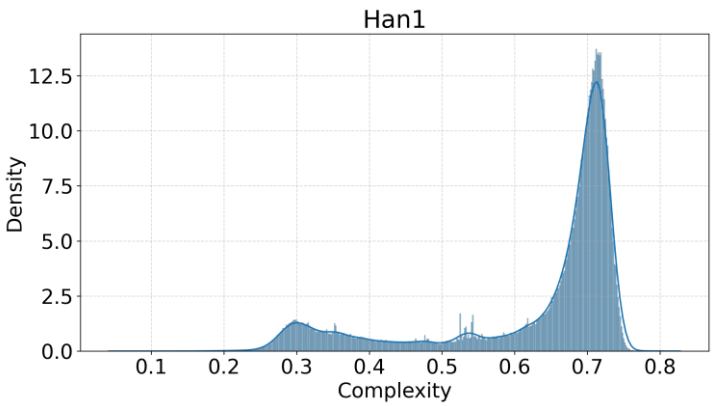


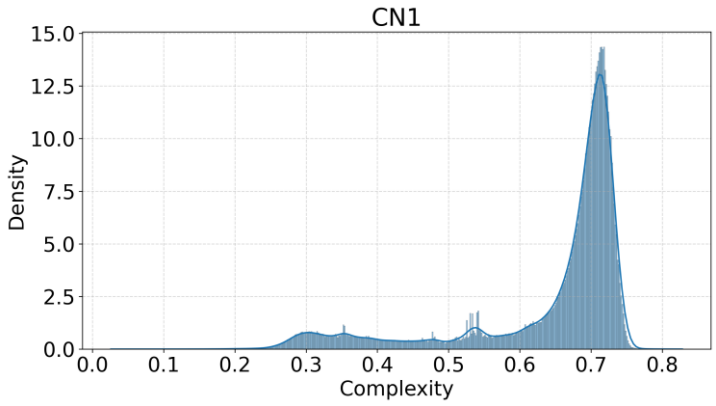
Supplementary Fig. 2

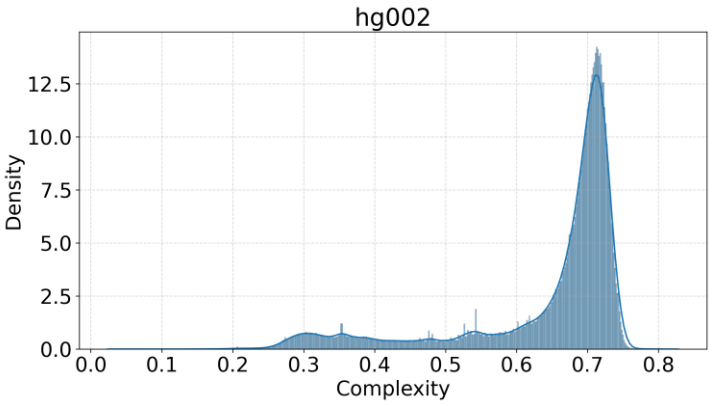


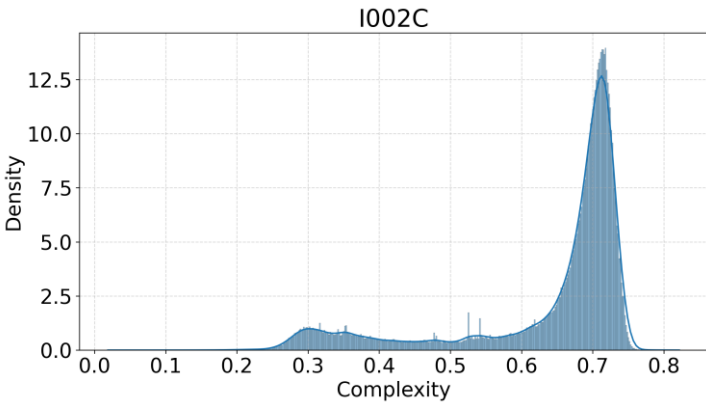
Supplementary Fig. 3

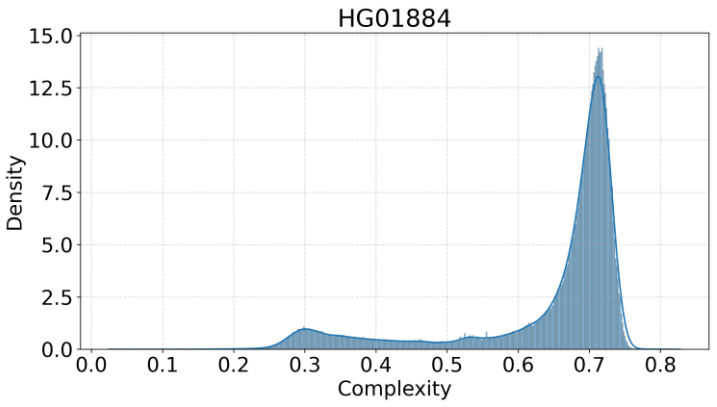


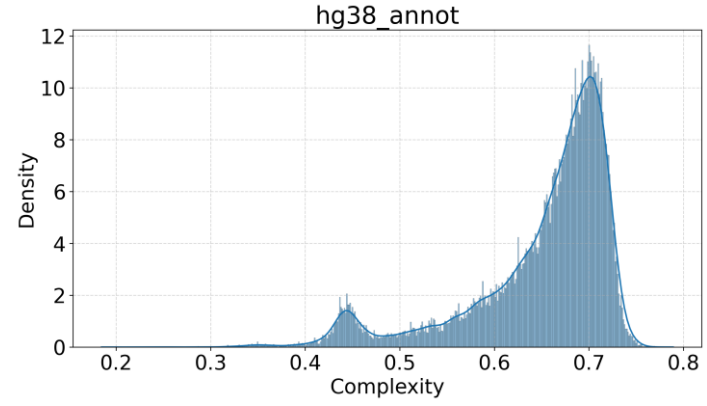


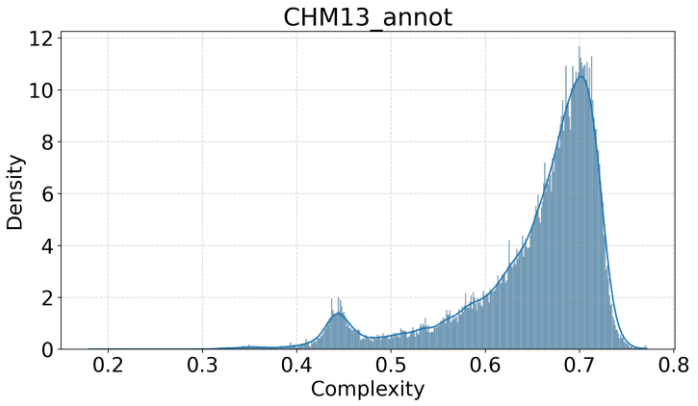




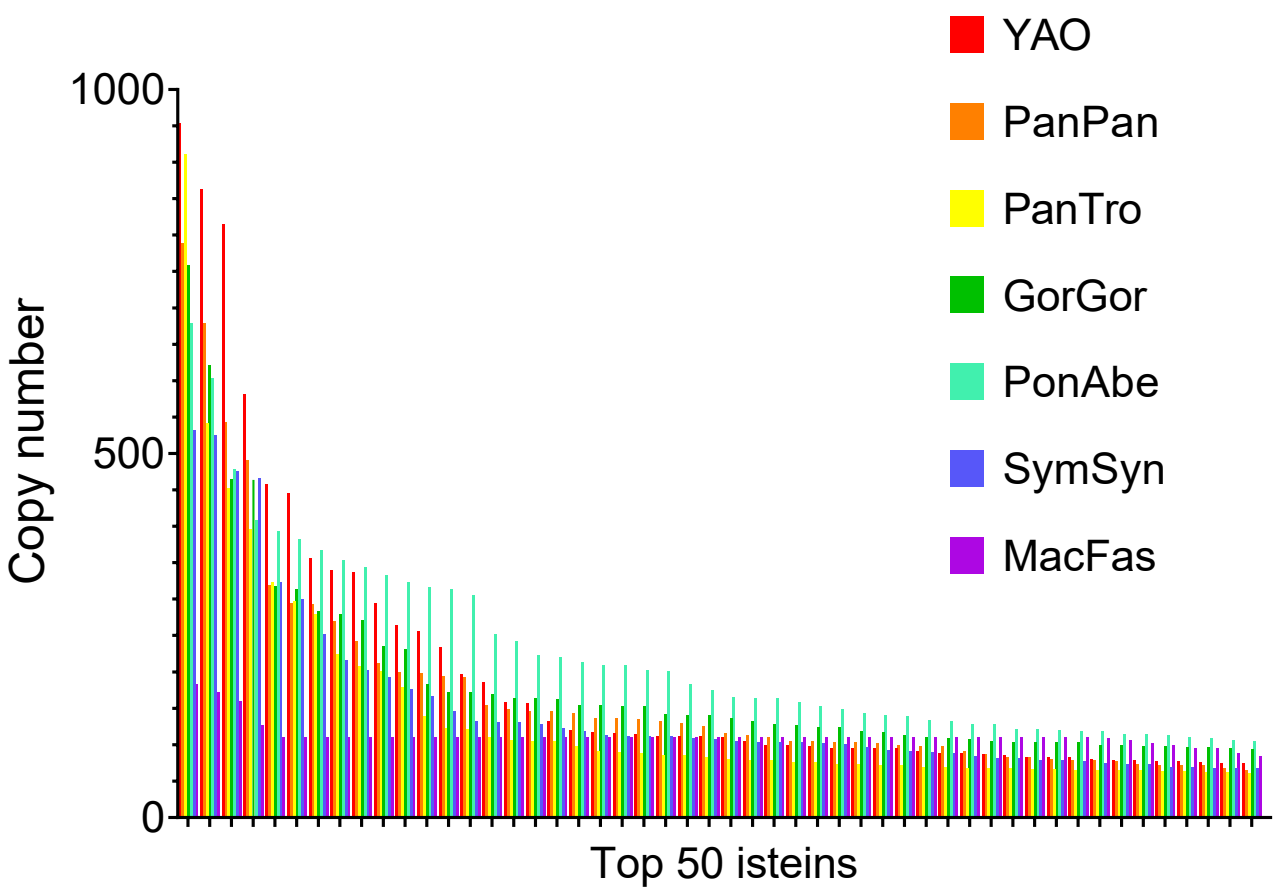




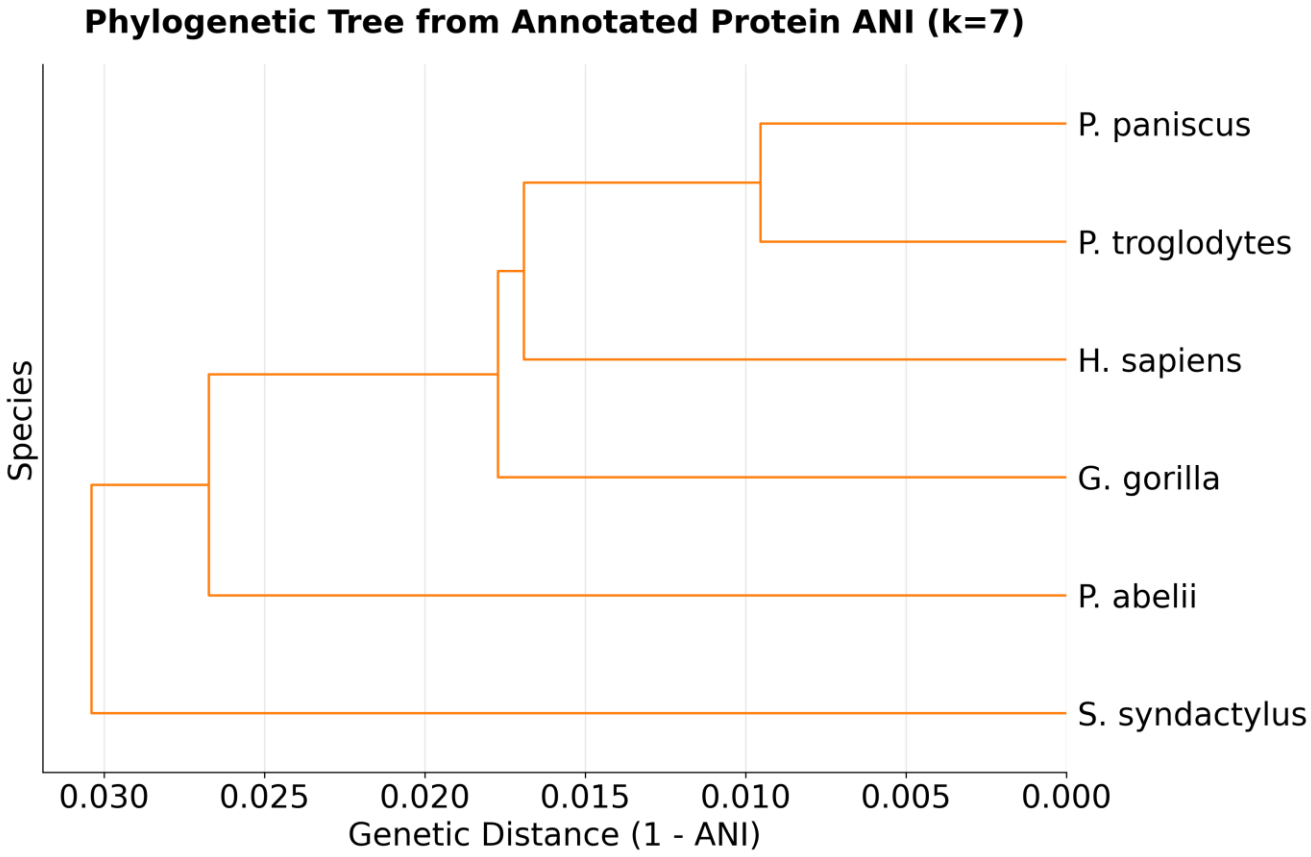


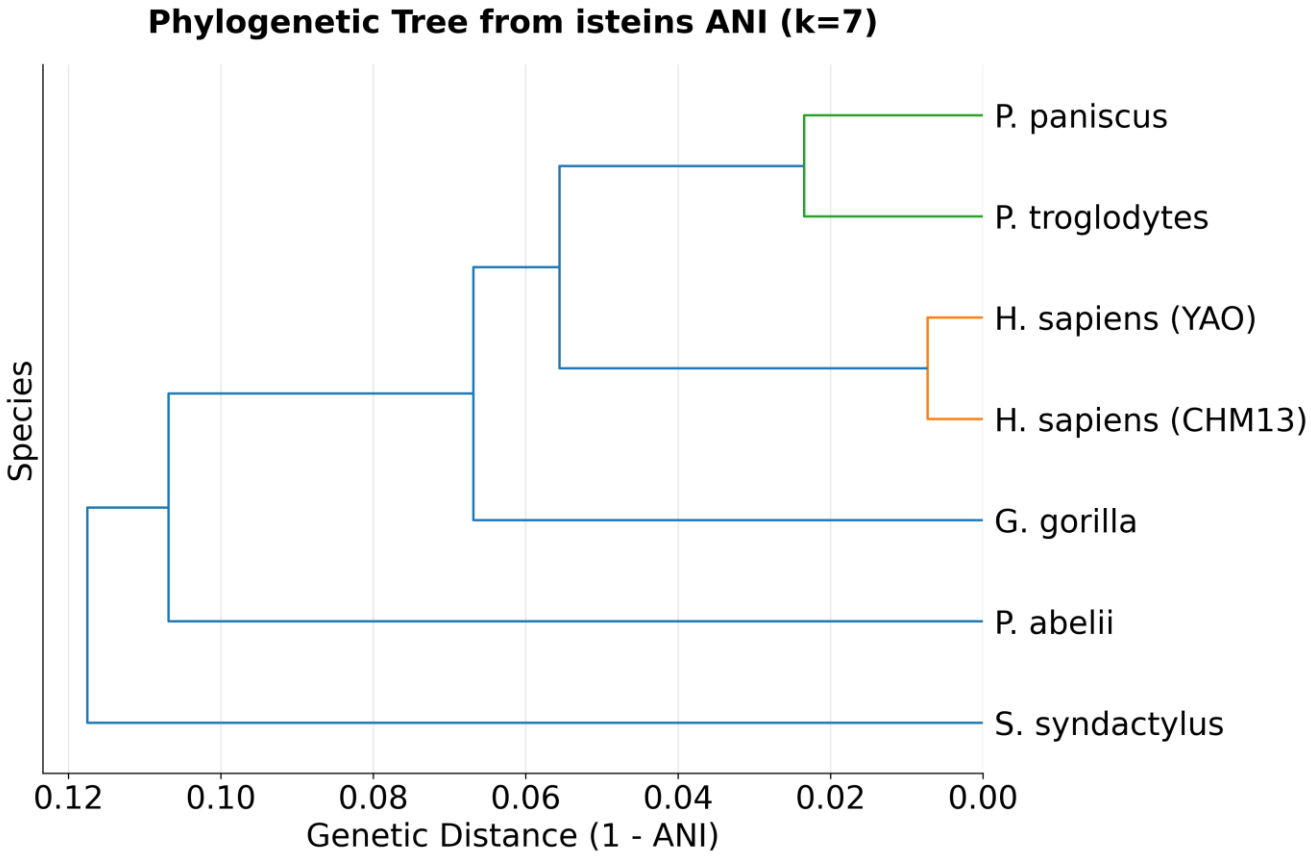


Supplementary Fig. 4



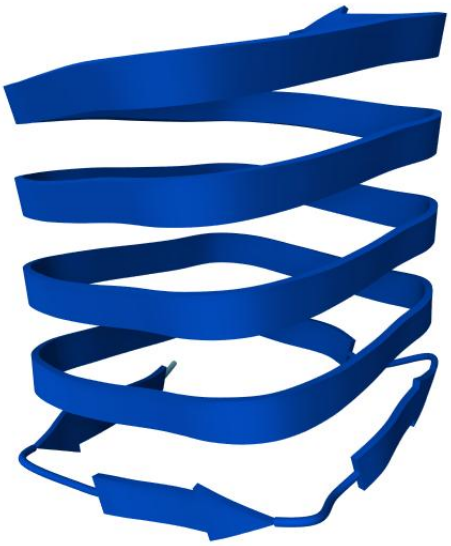
Supplementary Fig. 5



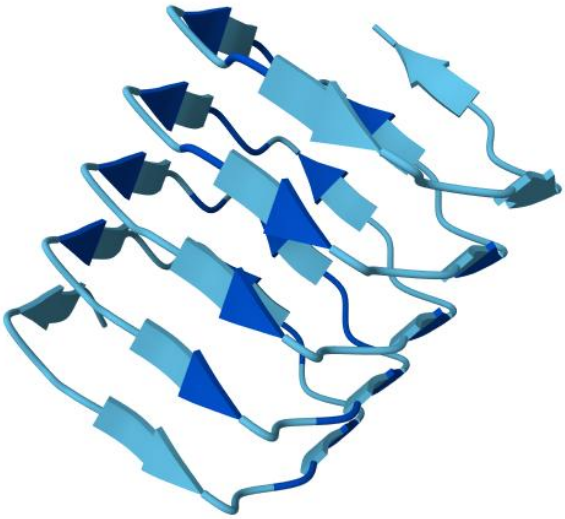


Supplementary Fig. 6

A



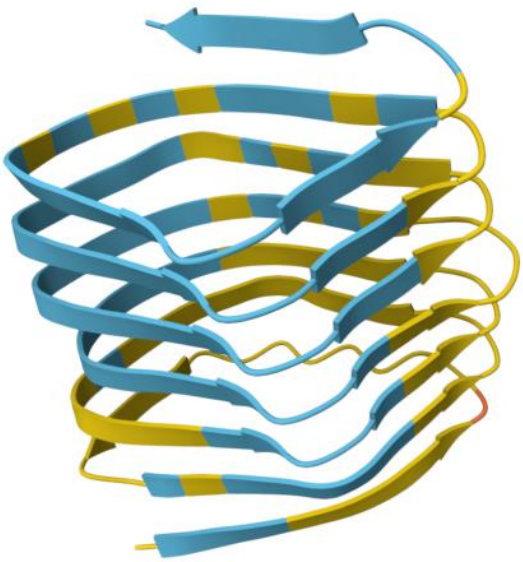
B



Sat2/3 heavy chain coding aa sequence:
GMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGMEWNGME
WNGMEWNGMEWNGMEWNGMEWNGMEWN

Sat2/3 light chain coding aa sequence:
IPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPFHSIPF
HSIPFHSIPFHSIPFHSIPFHSIPFHS

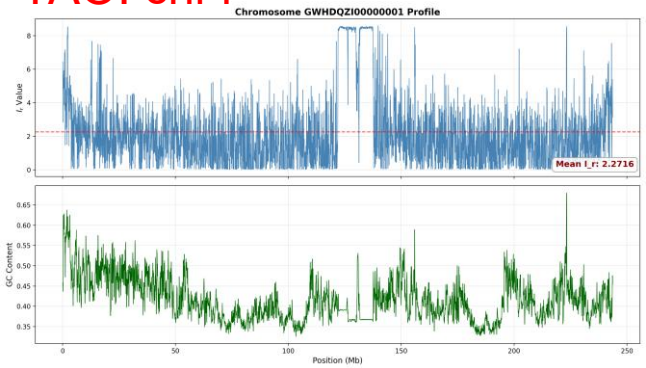
C



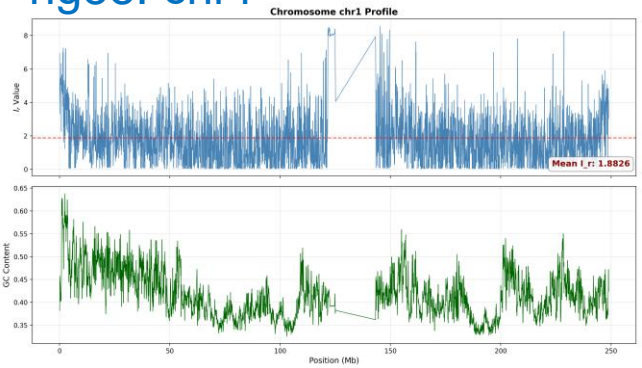
MESSS_162aa-
MKLKWMDSSSNGFEWNHRIKLIEIIIEWNRMVSLNGIEWNHHQMEMNGIVIEWNRMDSLNGIRWNHRMDWNGIIEW
TRKGSLLNGIEWNHRMVSIGIIIKWNRMDSPNRIEWNNHRMDSNGIILKWNRMELSNAIEWNYRMQSNRIIEWTRMES
SNGMEWNSQ

Supplementary Fig. 7

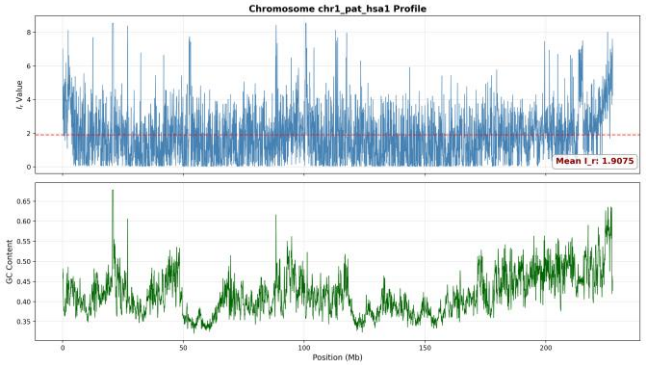
YAO: chr1



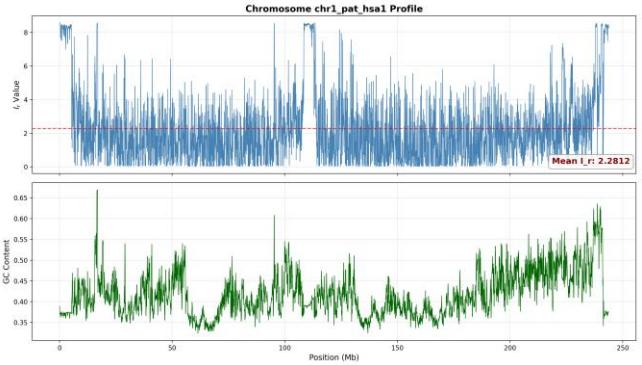
hg38: chr1



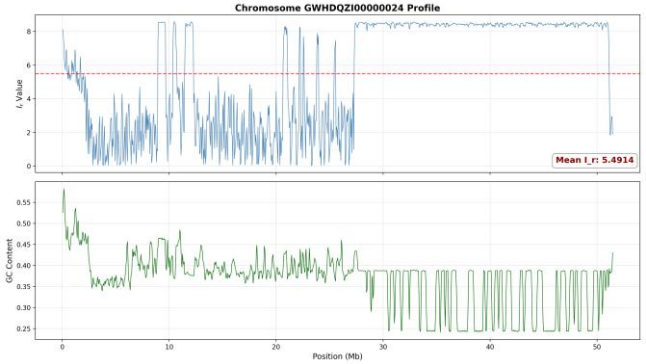
PanPan: chr1



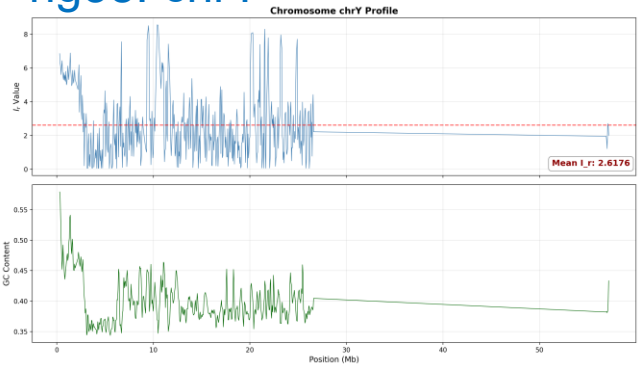
GorGor: chr1



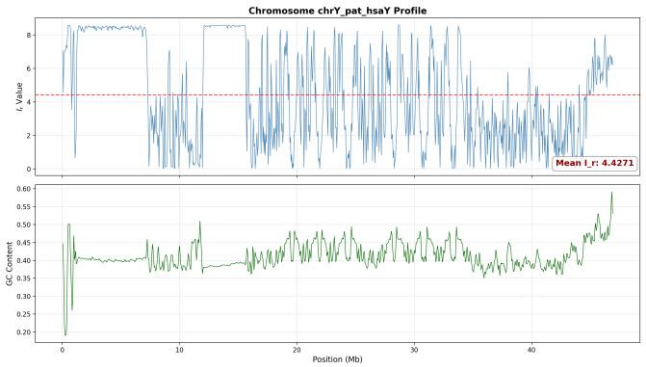
YAO: chrY



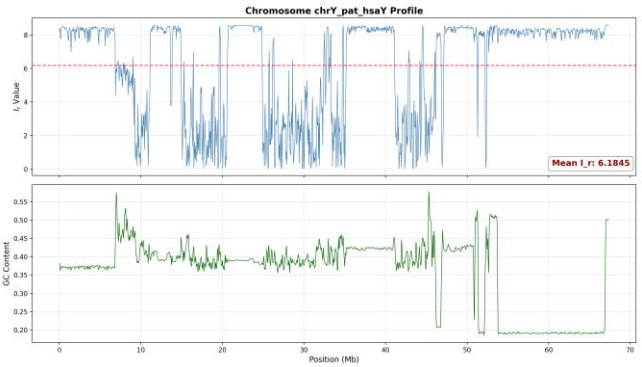
hg38: chrY



PanPan: chrY



GorGor: chrY



Supplementary Fig. 8

A

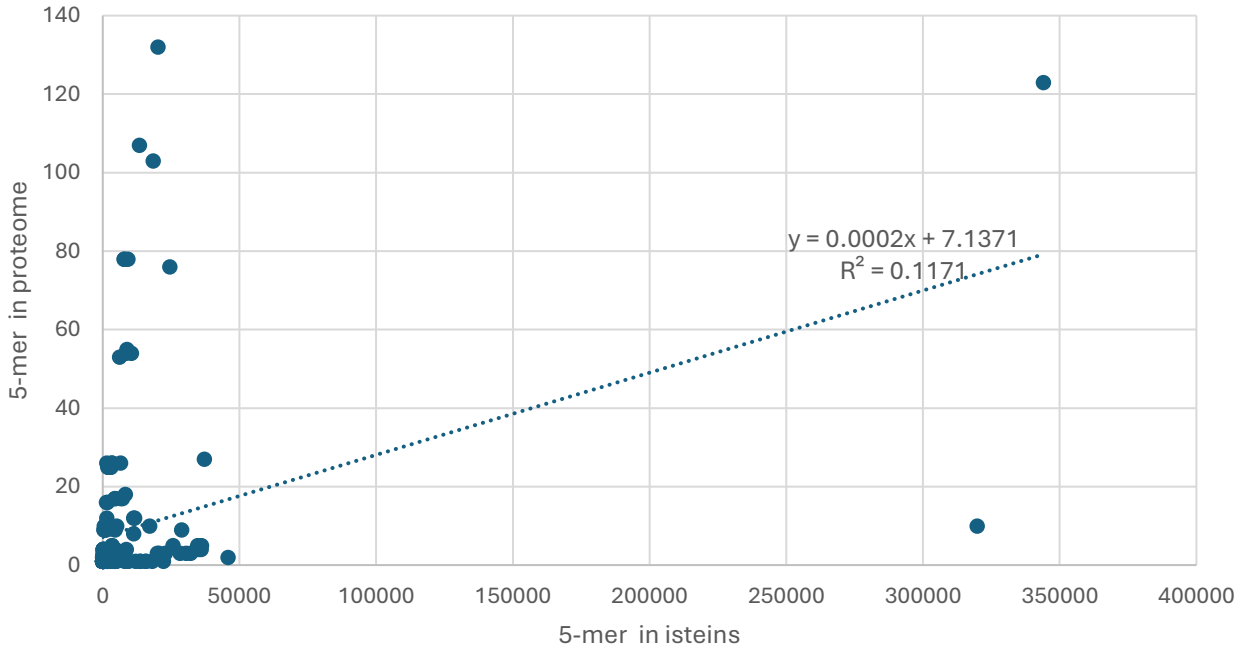
Amino Acid (%)	Theoretical Sat2/3 heavy chain	chr9_Sat3	MEW99	MESSS_chr16	MESSS_chr1
A	0	1.52	3.03	0.52	0.69
C	0	1.51	0.00	1.05	0.16
D	0	0.67	2.02	2.47	1.93
E	20	15.93	15.15	10.91	11.93
F	0	0.37	2.02	0.19	0.44
G	20	14.66	13.13	7.53	6.05
H	0	0.46	2.02	4.71	3.54
I	0	3.48	4.04	11.73	11.67
K	0	2	10.10	3.59	3.48
L	0	1.54	2.02	3.68	2.37
M	20	12.73	5.05	8.01	8.49
N	20	16.51	12.12	13.05	13.4
P	0	1.78	0.00	1.69	1.11
Q	0	2.41	2.02	2.52	1.56
R	0	4.04	3.03	5.64	9.18
S	0	3.25	5.05	12.25	13.11
T	0	2.58	3.03	2.6	2.41
V	0	1.62	3.03	1.15	1
W	20	12.57	9.09	5.99	6.72
Y	0	0.37	0	0.71	0.77

B

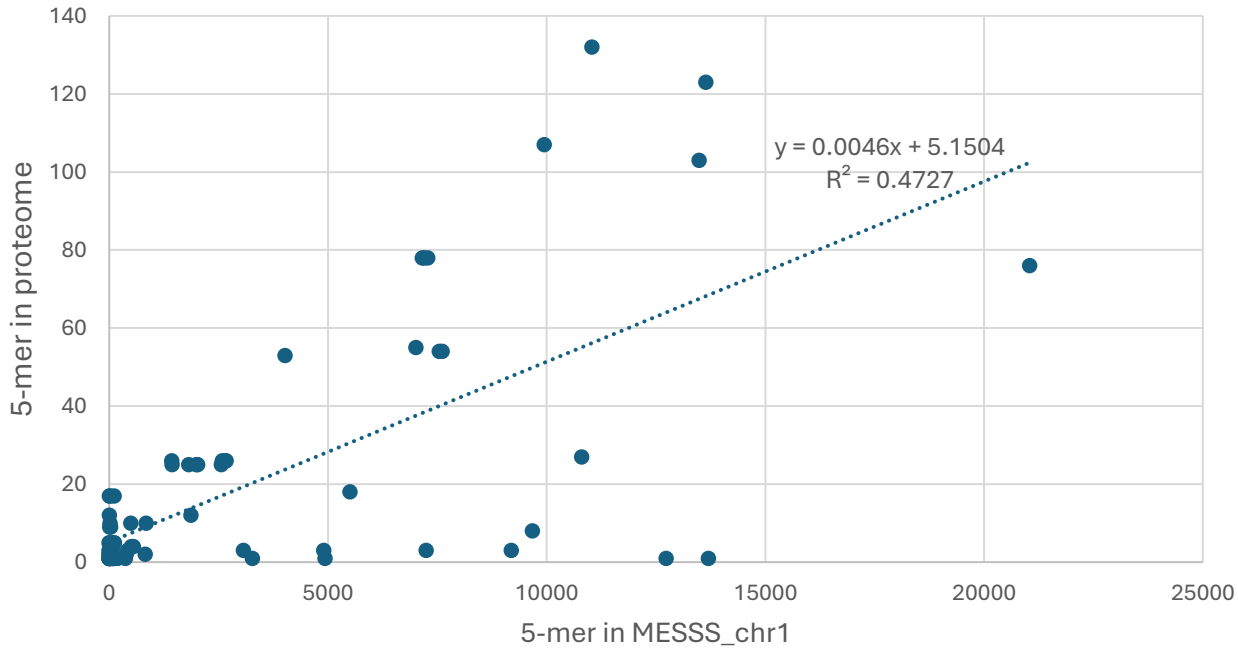
Nucleotide	Theoretical Sat2/3 heavy chain	Chr9_Sat2/3	MEW99	MESSS_chr16	MESSS_chr1
A (%)	40	39.07	45.5	38.85	39.31
T (%)	20	18.91	15	23.35	23.38
G (%)	40	34.1	32	22.19	22.78
C (%)	0	7.91	7.5	15.61	14.53
GC (%)	40	42.01	39.5	37.81	37.31
AT (%)	60	57.98	60.5	62.19	62.69
AG (%)	80	73.17	77.5	61.04	62.09
CT (%)	20	26.83	22.5	38.96	37.91

Supplementary Fig. 9

A



B



Supplementary Fig. 10

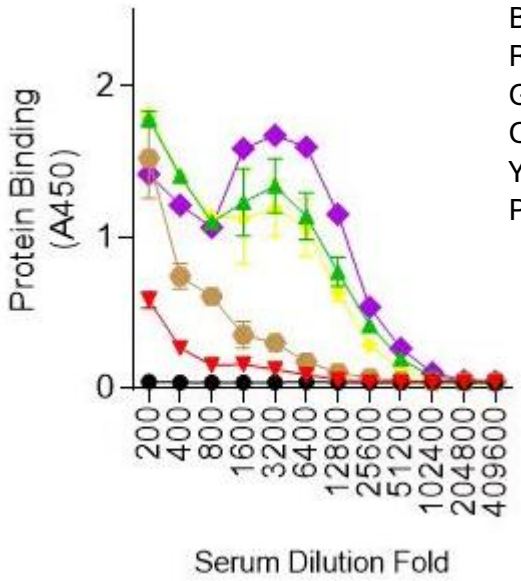
A

>YAO_MESSS_1_446
MESSSNELNAIIEWSRMESSSNGKEWNHRIESNGIIIEWNRMVSTPNGKKRNYRMESKRIFERTRMESSNGMEWN
NPWTRMQSSSNGIEWNHRMDSNGIIIERNRMESSSDGNEWNHHRMEMKGVII

Yellow labeled: peptide used for immunizing mice

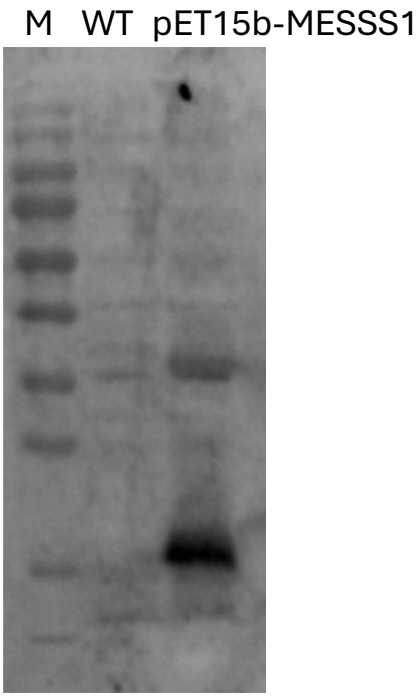
Designed CDS-
ATGGAGAGCTCATCTAATGAGCTTAACGCTATTATCGAATGGAGTAGAATGGAAAGTAGCTCCAATGGTAAAGAGTGGAACC
ATAGAATCGAGTCAAACGGGATTATCATAGAATGGAATCGGATGGTTTCTACTCCCAACGGAAAAAAGCGGAATTACAGGAT
GGAGAGTAAAAGAATTTTTGAGCGGACAAGGATGGAGAGCAGTAACGGAATGGAGTGGAACAATCCTTGGACACGCATGCAG
TCTTCCAGTAATGGGATTGAGTGGAACCATCGGATGGACTCCAATGGCATAATCATTGAACGGAATAGAATGGAGTCTAGCT
CAGATGGTAACGAATGGAACCATCACAGAATGGAGATGAAAGGGGTTATCATCTAA

B

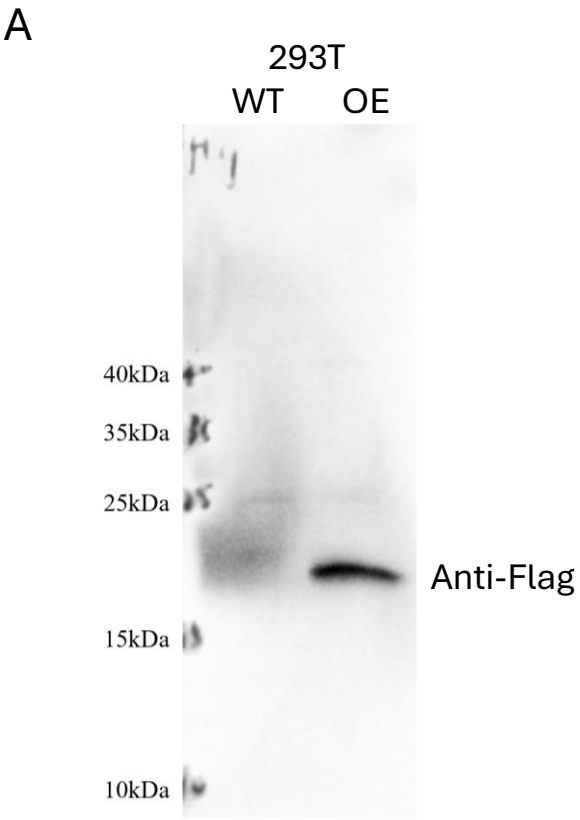


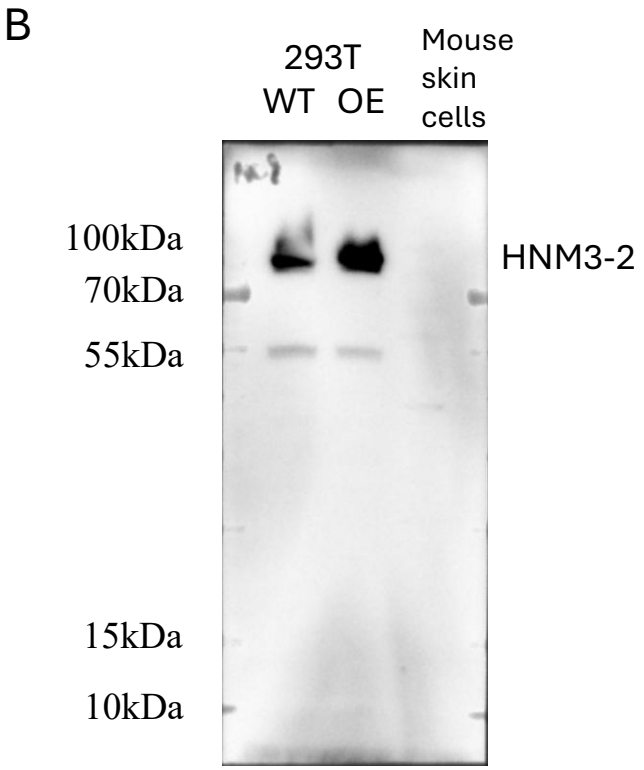
Black: control
Red: HNM3-1
Green: HNM3-2
Orange: HNM3-3
Yellow: HNM3-4
Purple: HNM3-5

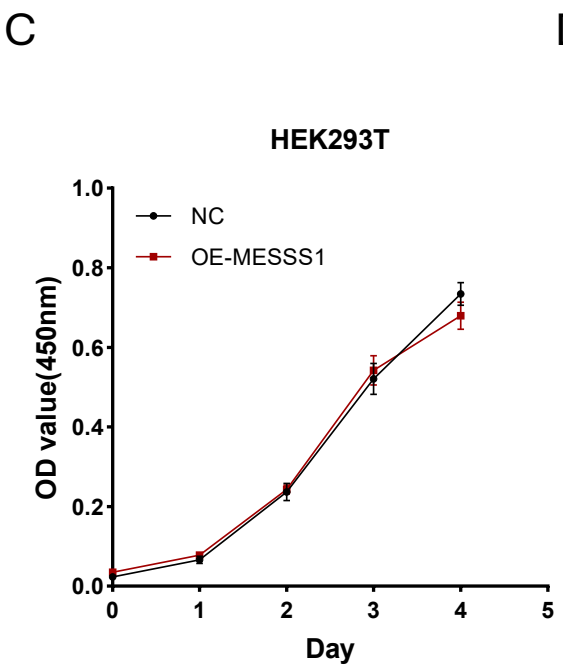
C

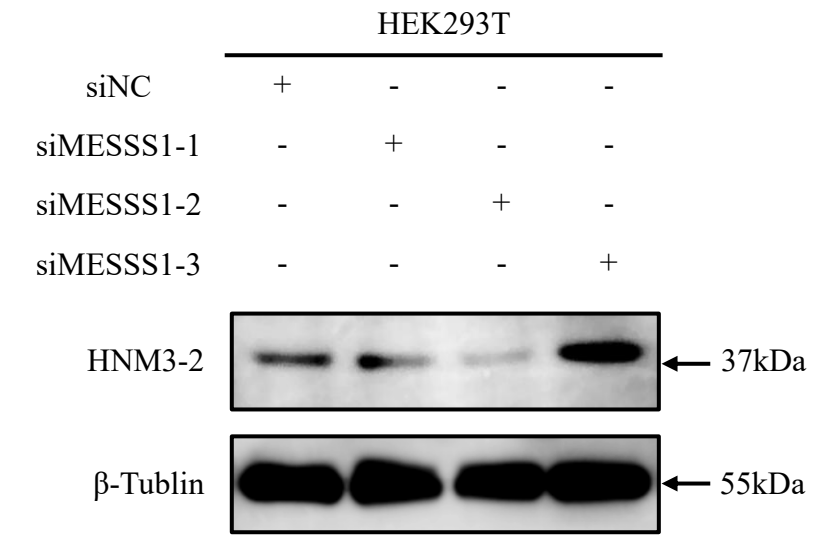


Supplementary Fig. 11









Supplementary Fig. 12

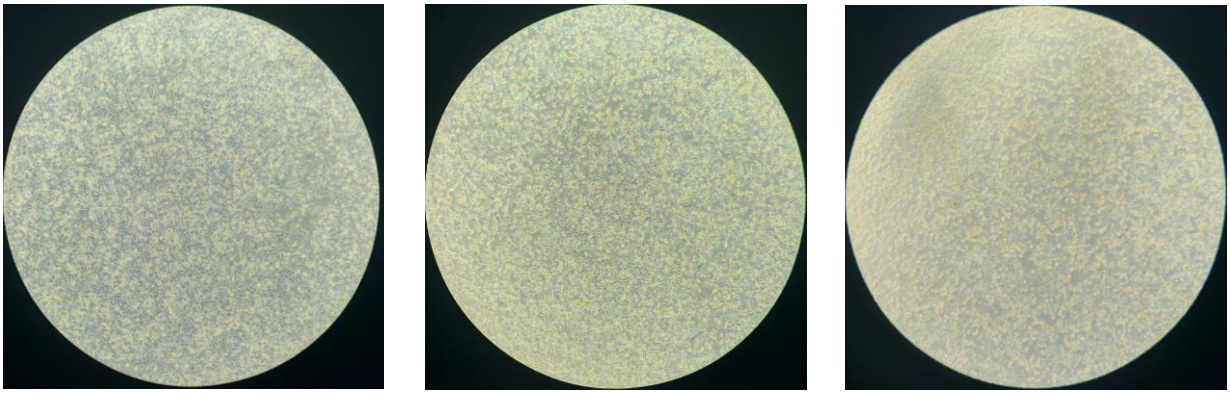
MDA-MB-231 siMESSS

A

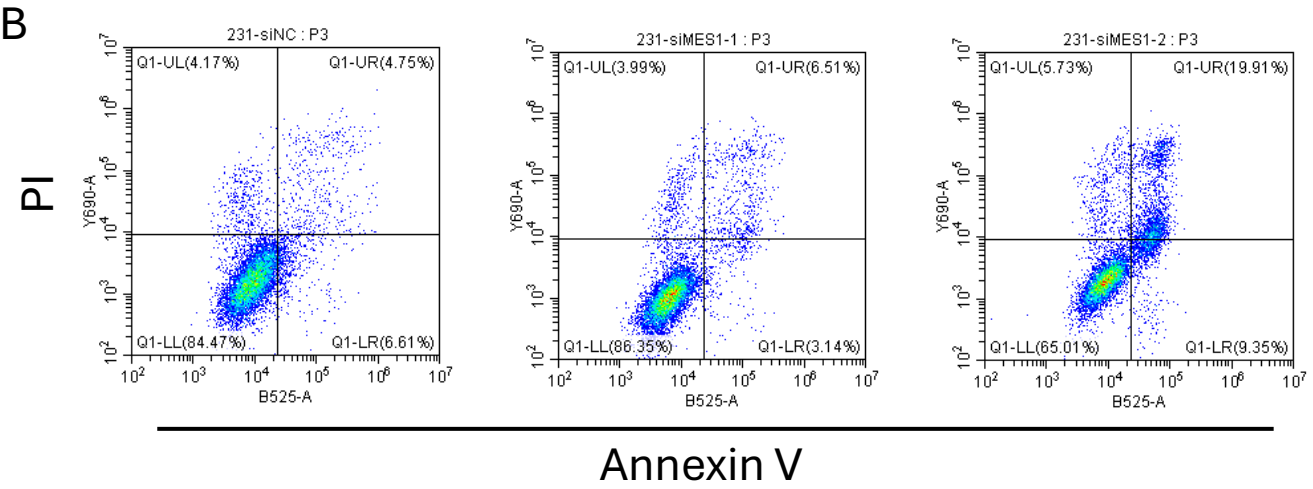
siNC
50nM

siMESSS1-1
50nM

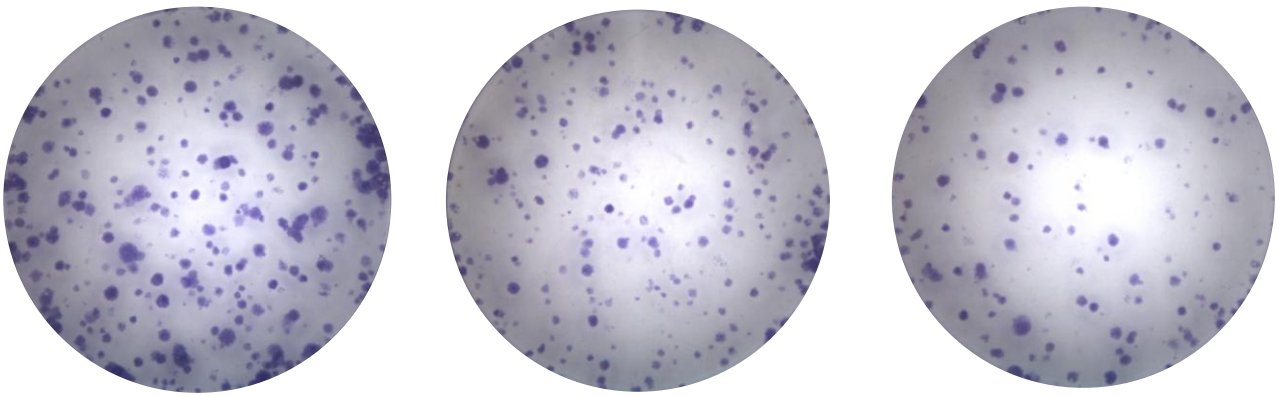
siMESSS1-2
50nM



B



C



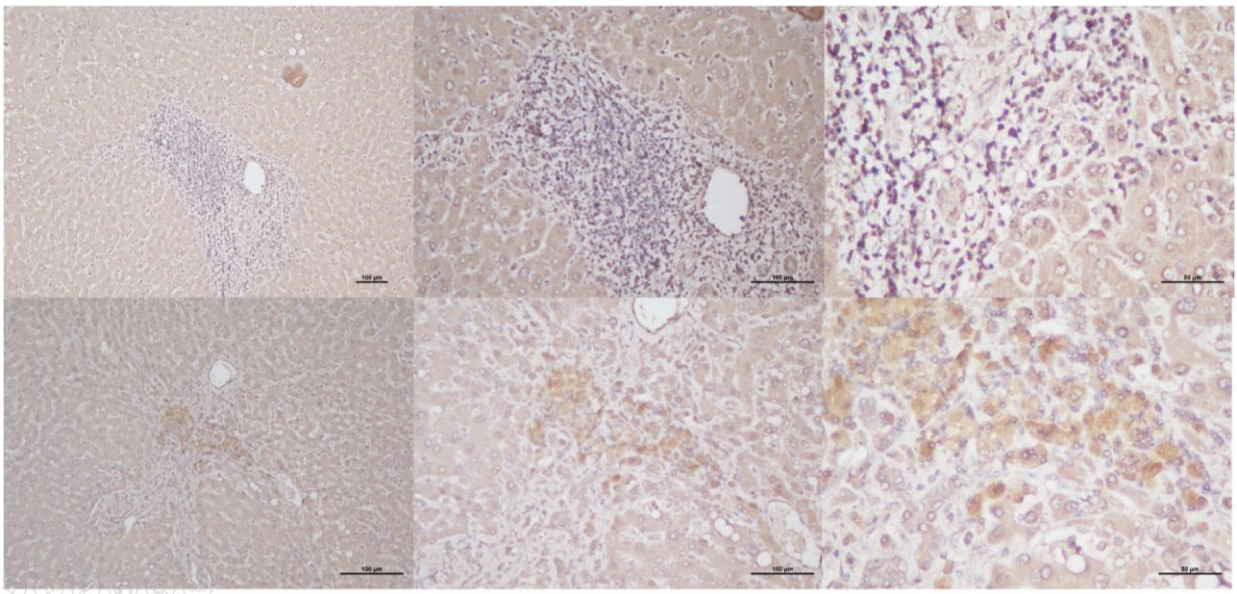
N=464

N=350

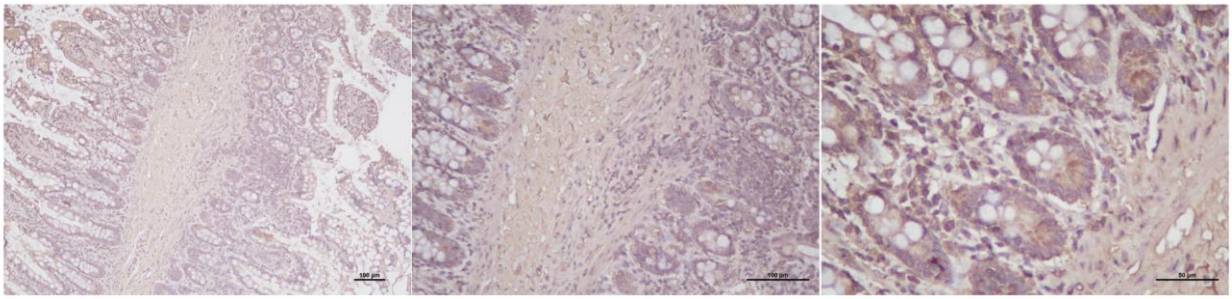
N=219

Supplementary Fig. 13

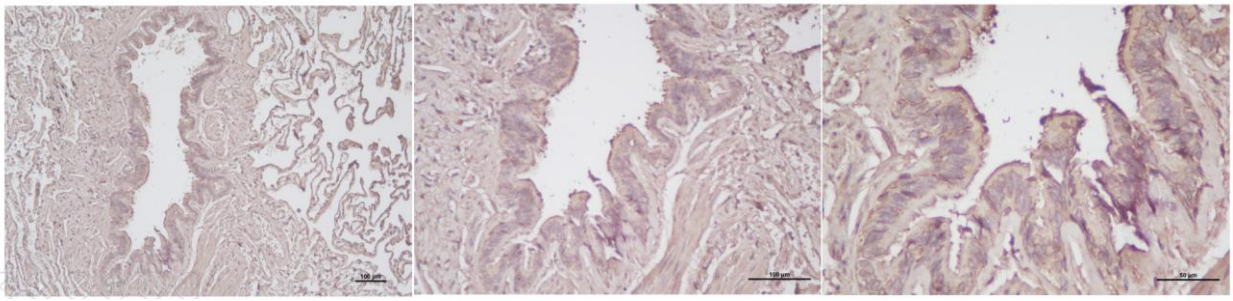
Liver



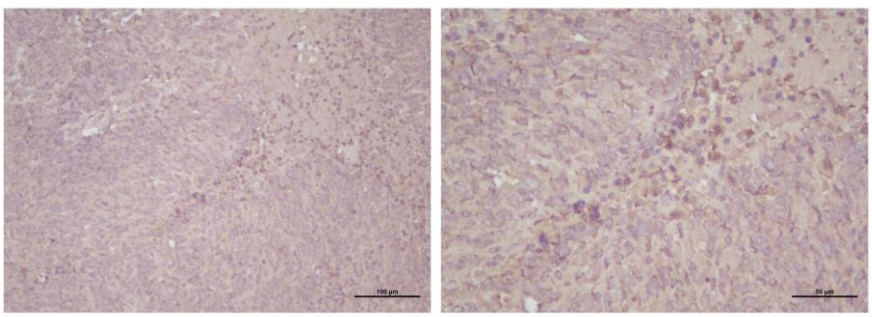
Intestine



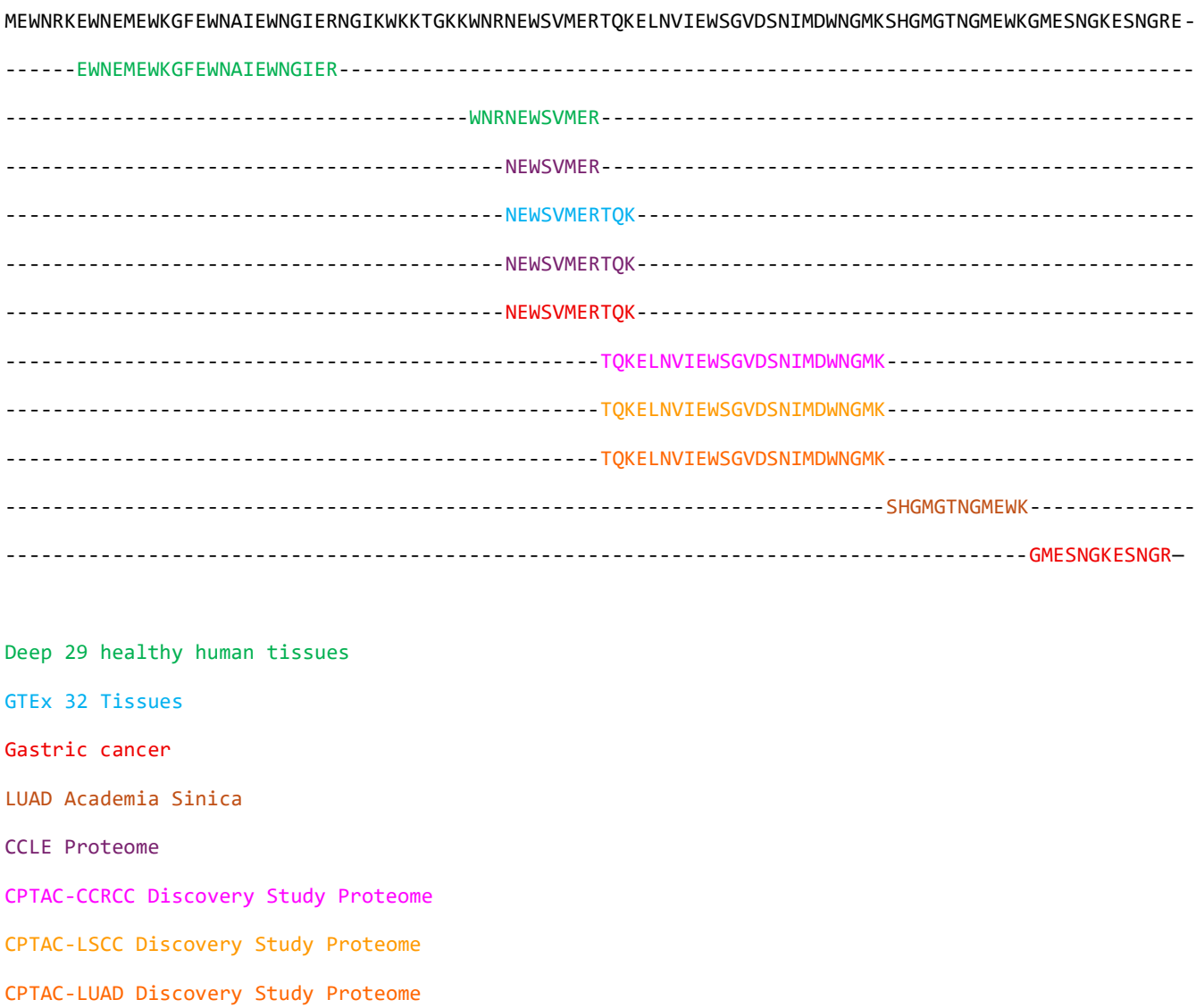
Lung



Ovarian cancer



Supplementary Fig. 14



Supplementary Fig. 15