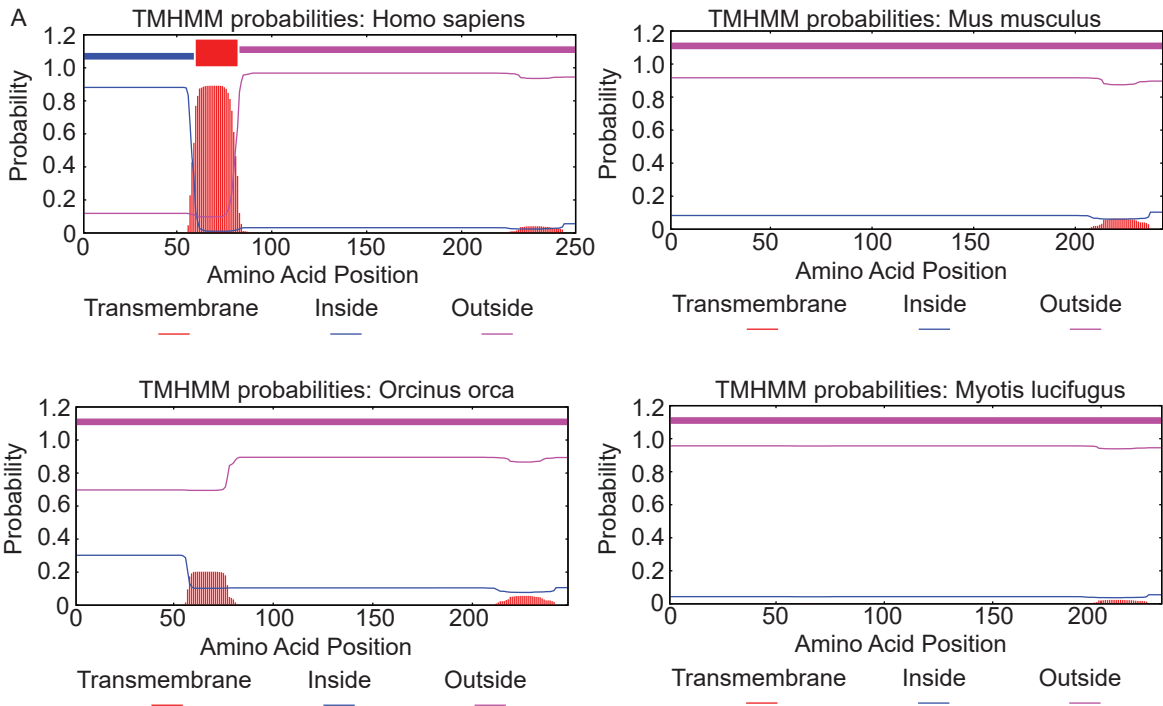


Supplementary Figure 5



B

GHPRGAGPGVRRGGLL

ELM Name	Instances (Matched Sequence)	Positions	ELM Description	Cell Compartment	Probability
CLV_NRD_NRD_1	RRG	11-13	N-Arg dibasic convertase (NRD/Nardilysin) cleavage site (X- -R-K or R- -R-X).	extracellular, golgi apparatus, cell surface	7.465x10 ⁻³
CLV_PCSK_KE_X2_1	RRG	11-13	Yeast kexin 2 cleavage site (K-R- -X or R-R- -X).	extracellular, Golgi apparatus	7.973x10 ⁻³
DOC_ANK_TNK_S_1	PRGAGP GV	3-10	The Tankyrase binding motif interacts with the ankryrin repeat domain region in Tankyrase-1 and Tankyrase-2 to facilitate the PARsylation of the target proteins.	nucleus, cytosol, chromosome, telomeric region, spindle pole, exon-exon junction complex, beta-catenin destruction complex	3.538x10 ⁻⁴
TRG_ER_diArg_1	VRR	10-12	The di-Arg ER retention motif is defined by two consecutive arginine residues (RR) or with a single residue insertion (RXR). The motif is completed by an adjacent hydrophobic/arginine residue which may be on either side of the Arg pair.	endoplasmic reticulum membrane, integral protein, ER-golgi transport vesicle membrane, endoplasmic reticulum membrane, golgi -ER transport vesicle membrane, rough endoplasmic reticulum, endoplasmic reticulum cisterna, cytosol	5.369x10 ⁻³
TRG_LysEnd_A_PsAcLL_1	RRGGLL	11-16	Sorting and internalisation signal found in the cytoplasmic juxta-membrane region of type I transmembrane proteins. Targets them from the Trans Golgi Network to the lysosomal-endosomal-melanosomal compartments. Interacts with adaptor protein (AP) complexes	cytosol, endocytic vesicle	2.758x10 ⁻³