

Figure S1 Proteomic analysis of total and phospho-S953 in LUAD and LUSC: Analysis of (A) total USO1 and (B) phosphoS953 USO1 protein levels in TCGA LUAD and LUSC samples. Analysis was carried out using cProSite: A web based interactive platform for on-line proteomics and phosphoproteomics data analysis. (<https://cprosite.ccr.cancer.gov>).

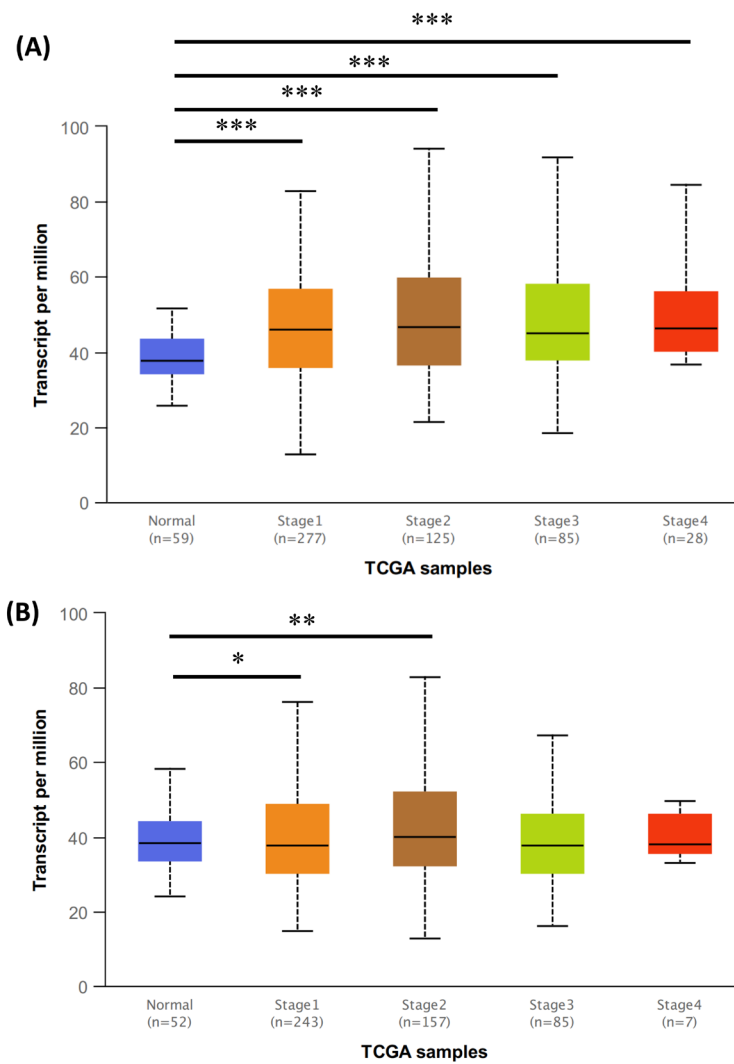


Figure S2 Effects of Staging on *USO1* Expression. Using UALCAN portal *USO1* upregulation occurs at the early stage of cancer development in both (A) LUAD and (B) LUSC, and there are no significant differences between every two stages in LUAD. Graphs are generated using UALCAN. *P<0.05, **P<0.01, ***P<0.001.

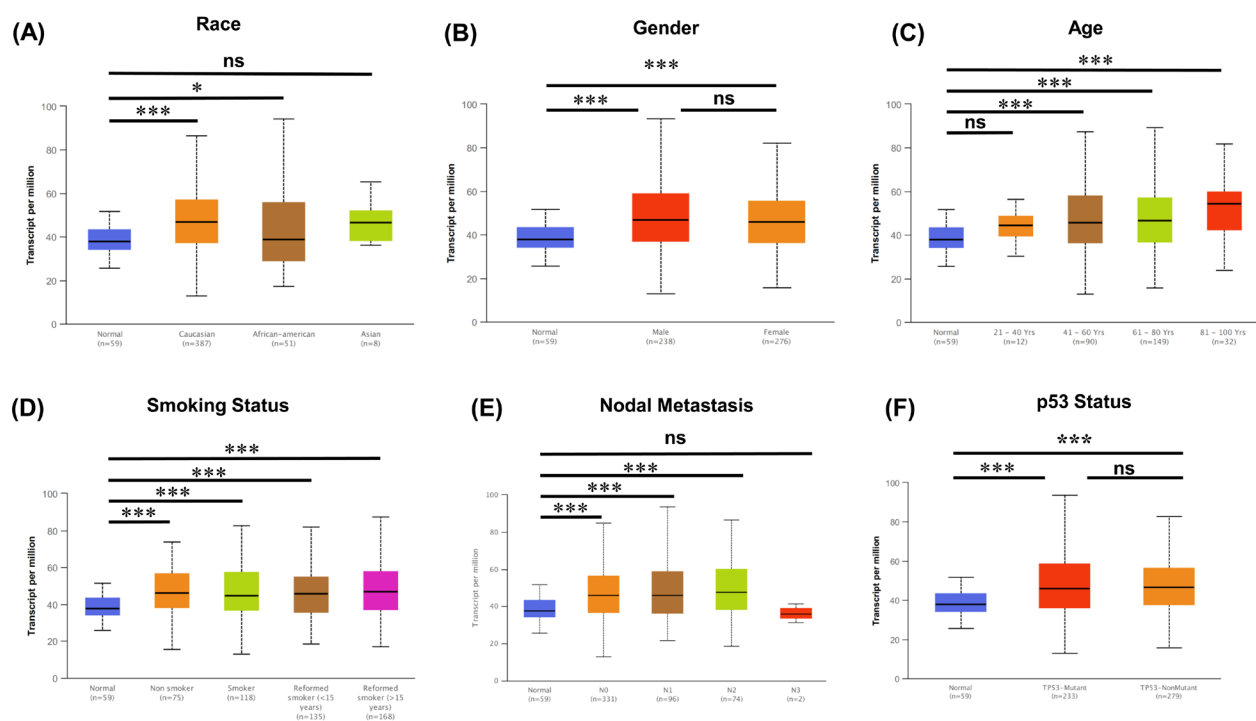


Figure S3 Subgroup expression analysis of *USO1* in LUAD. *USO1* mRNA expression levels were examined in the TCGA LUAD patients grouped according to (A) race, (B) gender, and (C) age. Effects of *USO1* mRNA expression were subsequently examined for (D) smoking status, (E) lymph node metastasis and (F) p53 mutation status. All graphs were generated using UALCAN. *P<0.05, ***P<0.001, ns, not-significant.

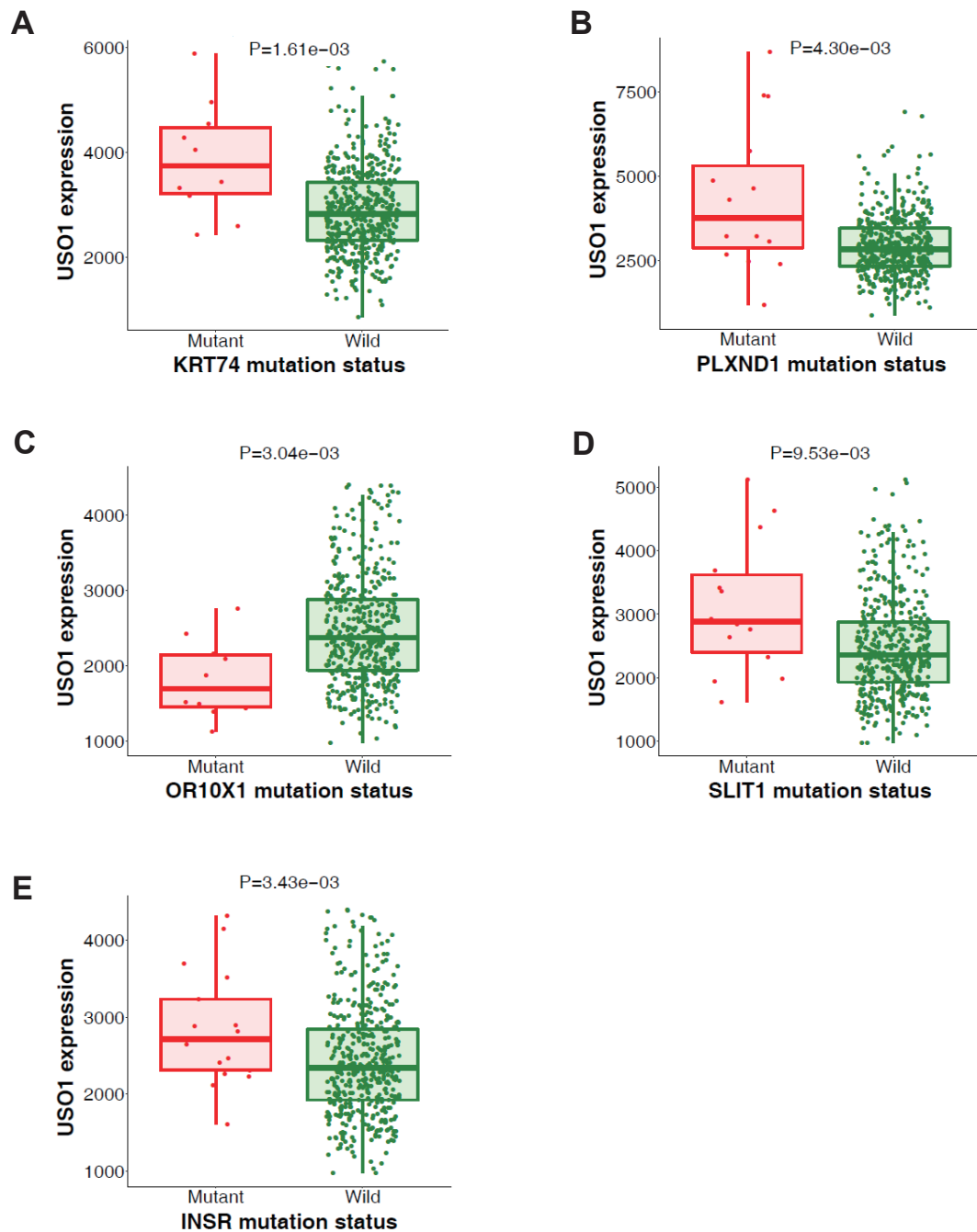
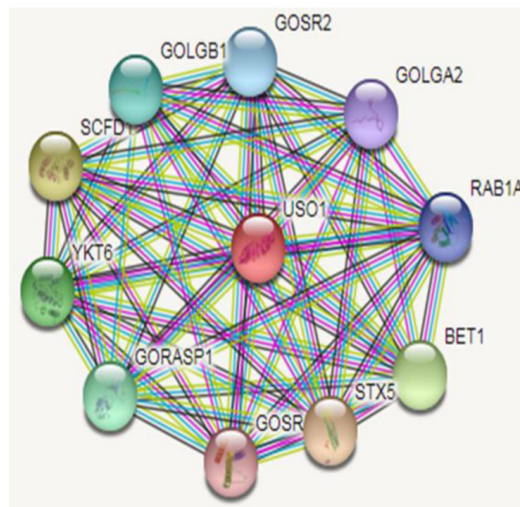


Figure S4 Effects of mutation on *USO1* expression. MuTarget (25) was used to identify and connect mutation status to gene expression changes of *USO1*. Using *USO1* as the Target gene with “all somatic mutations” selected and a cut-off for mutation prevalence set at a minimum of 2%, genes were identified which if mutated significantly altered *USO1* expression compared to their corresponding wild-type. (A) Mutated Keratin74, Type II (KRT74) and (B) Plexin D1 (PLXND1) were identified in LUAD; while mutated Olfactory receptor 10X1 (OR10X1) (C), Slit Guidance Ligand 1 (SLIT1) (D) and Insulin Receptor (INSR) (E) were associated with elevated *USO1* mRNA in LUSC.



Entrez ID	Symbol	Function
5861	RAB1A	GTPase - key regulators of intracellular membrane trafficking, from the formation of transport vesicles to their fusion with membranes
10282	BET1	Required for vesicular transport from the ER to the Golgi complex. Functions as a SNARE involved in the docking process of ER-derived vesicles with the cis-Golgi membrane
6811	STX5	Mediates endoplasmic reticulum to Golgi transport. Together with p115/USO1 and GM130/GOLGA2, involved in vesicle tethering and fusion at the cis-Golgi membrane to maintain the stacked and inter-connected structure of the Golgi apparatus.
9527	GOS-28 (GOSR1)	Involved in transport from the ER to the Golgi apparatus as well as in intra-Golgi transport. Belongs to the SNARE family
64689	GORASP1	Key structural protein required for maintenance of the Golgi apparatus integrity. Plays an important role in assembly and membrane stacking of the Golgi cisternae, and in the reassembly of Golgi stacks after breakdown during mitosis.
10652	YKT6	As part of a SNARE complex comprising GOSR1, GOSR2 and STX5, it functions in endoplasmic reticulum to Golgi transport.
23256	STXBP1L2 (SCFD1)	Plays a role in SNARE-pin assembly and Golgi-to-ER retrograde transport.
2804	GOLGB1	Postulated to form intercisternal cross-bridges of the Golgi complex.
9570	GOSR2	Transport of proteins from the cis/medial-Golgi to the trans-Golgi network.
2801	GOLGA2	Together with USO1 and STX5, involved in vesicle tethering and fusion at the cis-Golgi membrane to maintain the stacked and inter-connected structure of the Golgi apparatus. Plays a central role in mitotic Golgi disassembly.

Figure S5 Identification of USO1 first-neighbour protein-protein interactions. STRING (29) was used to identify protein-protein interactions with USO1. First Neighbours and their associated functions are indicated underneath as shown. Ras-Associated Protein RAB1 (RAB1); BET1 Golgi Vesicular Membrane-Trafficking Protein (BET1); Syntaxin 5 (STX5); Golgi Snare, 28-KD (GOS-28)/Golgi Snap Receptor Complex Member 1 (GOSR1); Golgi Reassembly Stacking Protein 1 (GORASP1); YKT6 v-SNARE Homolog (YKT6); SEC1 Family Domain-Containing Protein 1 (SCFD1/STXBP1L2); Golgi Autoantigen, Golgin Subfamily B, 1 (GOLGB1); Golgi Snap Receptor Complex Member 2 (GOSR2); Golgin A2 (GOLGA2).

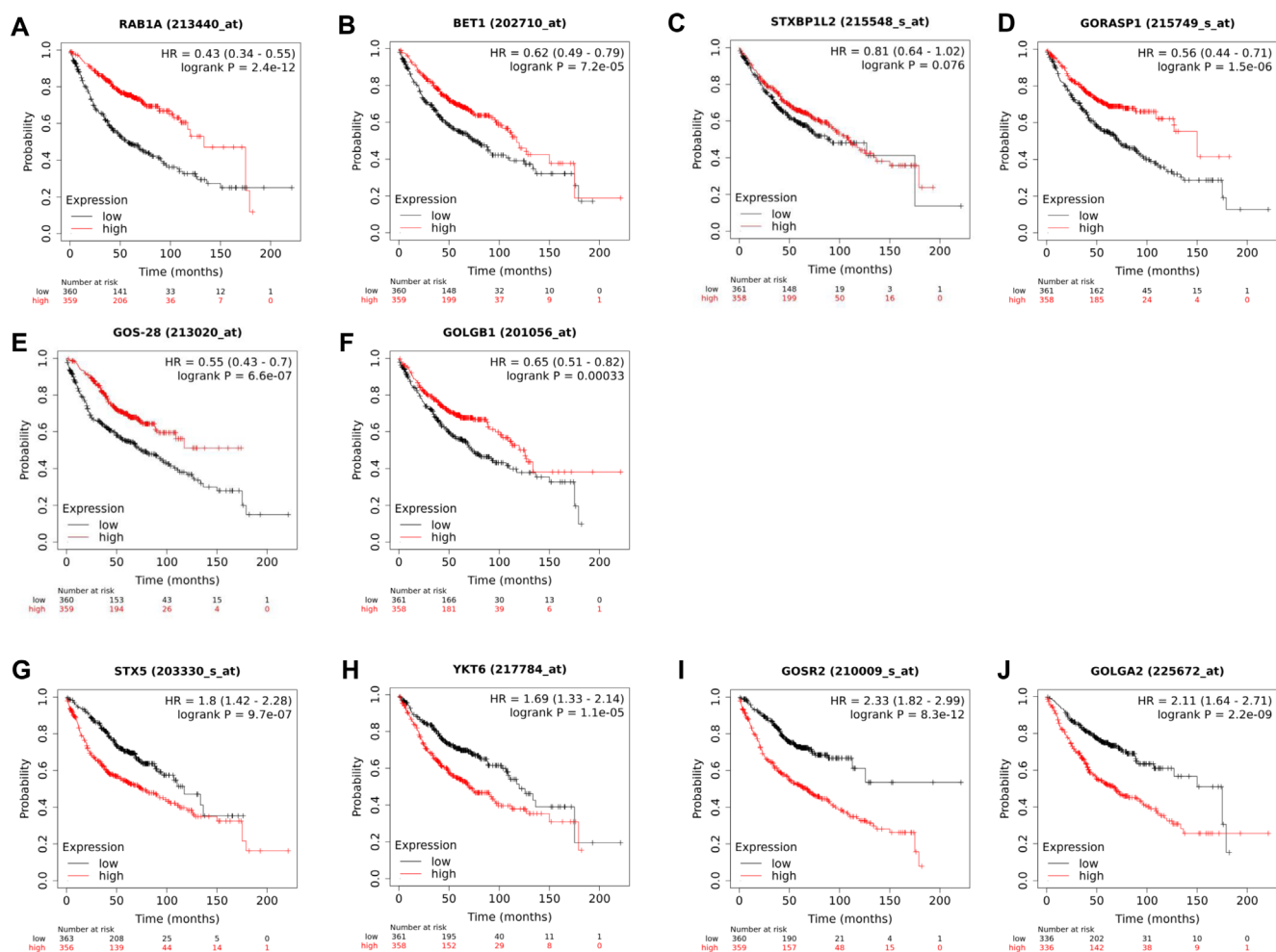


Figure S6 Overall survival (OS) for identified first neighbours in LUAD. KM plotter (22) was used for Kaplan-Meier analysis of the prognostic value (OS) of mRNA expression (stratified at the median) of *USO1*'s identified first neighbours in LUAD. High mRNA expression of (A) *RAB1A*; (B) *BET1*; (C) *STXBP1L2* (*SCFD1*); (D) *GORASP1*; (E) *GOS-28* (*GOSR1*); and (F) *GOLGB1* are associated with a better overall survival ($P < 0.05$); while high expression of (G) *STX5*; (H) *YKT6*; (I) *GOSR2* and (J) *GOLGA2* are associated with a worse OS ($P < 0.05$).

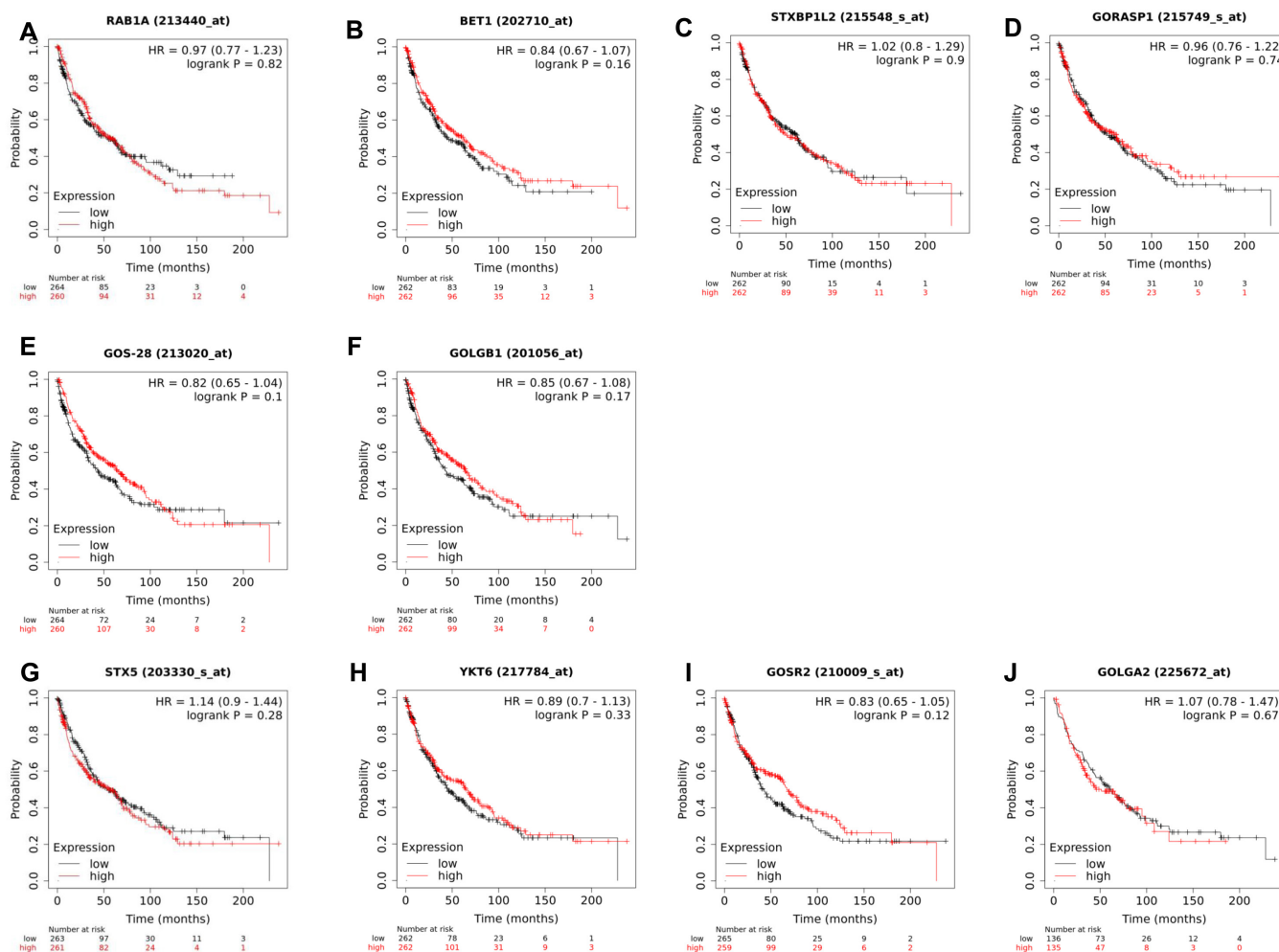


Figure S7 Overall survival (OS) for identified first neighbours in LUSC. KM plotter (22) was used for Kaplan-Meier analysis of the prognostic value (OS) of mRNA expression (stratified at the median) of USO1's identified first neighbours in LUSC. No significant OS benefit was observed for (A) *RAB1A*; (B) *BET1*; (C) *STXBP1L2* (*SCFD1*); (D) *GORASP1*; (E) *GOS-28* (*GOSR1*); (F) *GOLGB1* (G) *STX5*; (H) *YKT6*; (I) *GOSR2* and (J) *GOLGA2*.

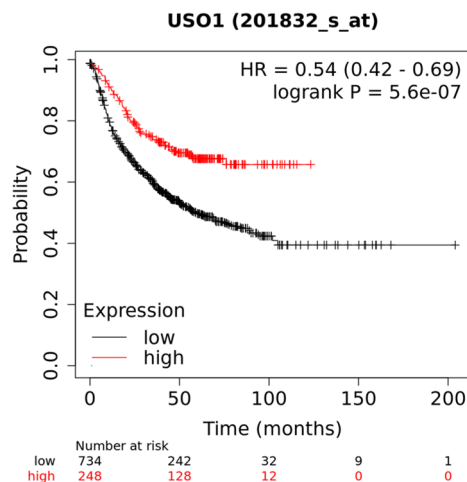


Figure S8 OS survival analysis for *USO1* mRNA in gastric cancer: KM plotter (22) was used for Kaplan-Meier analysis of the prognostic value for overall survival (OS) of *USO1* mRNA expression (stratified at the median) in Gastric Cancer.

```
>NP_001276978.1 general vesicular transport factor p115 isoform 1 [Homo sapiens]
MNFLRGVMGGQSAGPQHTEAETIQKLCDRVASSTLLDDRRNAVRALKSLSKYRLEVGIQAMEHLIHVLQ
TDRSDSEIIGYALDTLYNIIISNEEEEEVDVEEENSTRQSEDLGSQFTEIFIKQQENVTLTLLSLEEFD
HVRWPGVKLLTSLKQLGPVQVQIILVSPMGVSRMLDLDLADSREVIRNDGVLLQLALTRSNQAIQKIVAF
ENAFERLLDIISEEGNSDGGIVVEDCLILLQNLKNNNSNQNFKEGSYIQRMKPWFEVGDENSGWSAQK
VTNLHMLQLVRVLVSPNPPGATSSCQKAMFQCGLLQQLCTILMATGVPADILTETINTVSEVIRGCQV
NQDYFASVNAFSPNPPRPAIVVLLMSMVNERQPFVLRCAVLYCFQCFLYKNQKGQGEIVSTLLPSTIDATG
NSVSAGQLLCGGLFSTDLSLNWCAVALAHALQENATQKEQLLRVQLATSIGNPPVSLQQCTNLSQGD
KIDRRGSKIQTRVGLMLLCTWLSNCP IAVTHFLHNSANVPFLTGTQIAENLGEEQLVQGLCALLLGISI
YFNDNSLESYMKELKQLIEKRIGKENFIEKLGFIKHELYSRASQKPPNFPSPPEYIMFDHEFTKLVE
LEGVITKAIYKSSEEDKKEEVKKTLEQHDNIVTHYKNMIREQDLQLEELRQVSTLKCQNEQLQTAVTQ
QVSQIQQHKDQYNLLKIQLGKDNQHQQSYSEGAQMNGIQPEEIGRLREEIEELKRNQELLQSQLTEKDSM
IENMKSSQTSQSGTNEQSSAIVSARDSEQVAELKQELATLKSQLSQSVEITKLQTEKQELLQKTEAFKSV
EVQGETETIIATKTTDVEGRSALLQETKELKNEIKALSEERTAIKEQLDSSNSTIAILQTEKDKLELEI
TDSKKEQDDLLVLLADQDQKILSLKNKLDLGHFVVEEEDLESGDQDEDEDESEDPGKDLDDHI
S953
```

```
>NP_003706.2 general vesicular transport factor p115 isoform 2 [Homo sapiens]
MNFLRGVMGGQSAGPQHTEAETIQKLCDRVASSTLLDDRRNAVRALKSLSKYRLEVGIQAMEHLIHVLQ
TDRSDSEIIGYALDTLYNIIISNEEEEEVEENSTRQSEDLGSQFTEIFIKQQENVTLTLLSLEEFDHVRW
PGVKLLTSLKQLGPVQVQIILVSPMGVSRMLDLDLADSREVIRNDGVLLQLALTRSNQAIQKIVAFENAF
ERLLDIIISEEGNSDGGIVVEDCLILLQNLKNNNSNQNFKEGSYIQRMKPWFEVGDENSGWSAQKVTNL
HMLQLVRVLVSPNPPGATSSCQKAMFQCGLLQQLCTILMATGVPADILTETINTVSEVIRGCQVNDQY
FASVNAFSPNPPRPAIVVLLMSMVNERQPFVLRCAVLYCFQCFLYKNQKGQGEIVSTLLPSTIDATGSKI
AGQLLCGGLFSTDLSLNWCAVALAHALQENATQKEQLLRVQLATSIGNPPVSLQQCTNLSQGSQIQT
RVGLMLLCTWLSNCP IAVTHFLHNSANVPFLTGTQIAENLGEEQLVQGLCALLLGISYIFNDNSLESY
KEKELKQLIEKRIGKENFIEKLGFIKHELYSRASQKPPNFPSPPEYIMFDHEFTKLVELEGVITKAIYK
SSEEDKKEEVKKTLEQHDNIVTHYKNMIREQDLQLEELRQVSTLKCQNEQLQTAVTQVVSQIQQHKDQ
YNLLKIQLGKDNQHQQSYSEGAQMNGIQPEEIGRLREEIEELKRNQELLQSQLTEKDSMIENMKSSQTS
TNEQSSAIVSARDSEQVAELKQELATLKSQLSQSVEITKLQTEKQELLQKTEAFKSVQGETETIIA
TKTTDVEGRSALLQETKELKNEIKALSEERTAIKEQLDSSNSTIAILQTEKDKLELEITDSKKEQDDLL
VLLADQDQKILSLKNKLDLGHFVVEEEDLESGDQDEDEDESEDPGKDLDDHI
S942
```

Figure S9 Mapping of phospho-Serine residues critical for USO1 association with the Golgi. Originally a site for phosphorylation at position Serine 942 (S942) was identified as being critical for regulating USO1s association with the Golgi membrane (39). The phosphorylation site identified in the CPTAC analysis was identified at position S953. When the two existing isoforms of USO1 are aligned, the mapped position Phospho S942 on isoform 2 (NP_003706.2), equates to position S935 on isoform 1 (NP_001276978.1).

Table S1 Systematic Analysis results for *USO1* and associated first neighbours in NSCLC as determined using Lung Cancer Explorer (Meta-analysis of Standardized Mean Difference of Tumour-Normal gene expression)

Entrez ID	Symbol	SMD	SMD lower	SMD upper	P value	P.adj	Tumour-normal standardized expression difference
8615	USO1	0.63	0.44	0.81	5.3E-11	3.3E-10	Lung adenocarcinoma (LUAD)
		0.23	0.00	0.46	0.047	0.075	Lung squamous cell carcinoma (LUSC)
5861	RAB1A	0.6	0.27	0.93	0.00033	0.00081	LUAD
		0.75	0.46	1.04	5.1E-7	2.2E-6	LUSC
10282	BET1	1.08	0.51	1.66	0.00022	0.00056	LUAD
		0.8	0.62	0.97	1.2E-18	2.4E-17	LUSC
6811	STX5	0.04	-0.13	0.2	0.66	0.72	LUAD
		-0.31	-0.49	-0.12	9E-4	0.0021	LUSC
9527	GOS-28 (GOSR1)	0.57	0.4	0.74	3.7E-11	2.4E-10	LUAD
		0.13	-0.04	0.30	0.14	0.19	LUSC
64689	GORASP1	-1.31	-1.49	-1.13	1.5E-47	1.5E-45	LUAD
		-2.24	-3.09	-1.38	2.8E-7	1.2E-6	LUSC
10652	YKT6	1.59	1.09	2.09	4.9E-10	2.8E-9	LUAD
		1.91	1.35	2.47	2.6E-11	2.2E-10	LUSC
23256	STXBP1L2 (SCFD1)	0.54	0.2	0.89	0.0022	0.0046	LUAD
		0.34	0.08	0.6	0.01	0.019	LUSC
2804	GOLGB1	0.12	-0.18	0.41	0.44	0.52	LUAD
		0.08	-0.29	0.46	0.67	0.73	LUSC
9570	GOSR2	0.85	0.61	1.10	1E-11	7E-11	LUAD
		0.52	0.18	0.86	0.0028	0.0059	LUSC
2801	GOLGA2	0.72	0.49	0.95	1.5E-9	8.2E-9	LUAD
		0.16	-0.26	0.59	0.46	0.53	LUSC

Entrez Identification Number (ID), National Center for Biotechnology Information (NCBI) designated Gene ID; Symbol, Gene Symbol; SMD, tumour-normal standardized mean difference; SMD.lower, lower bound of 95% confidence interval for SMD; SMD.upper, upper bound of 95% confidence interval for SMD; P.adj, multiple comparison adjusted P value by Benjamini Hochberg procedures; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.