Identification of LMAN1 and SURF4 dependent secretory cargoes

Vi T. Tang[1,2], Prabhodh S. Abbineni[2,3], Felipe da Veiga Leprevost[4], Venkatesha Basrur[4], Rami Khoriaty[5,6], Brian T. Emmer[5], Alexey I. Nesvizhskii[4,7], and David Ginsburg[2,5,8,9,10,*]

[1]Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, MI 48109, USA
[2]Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109, USA
[3]Department of Microbiology and Immunology, Loyola University Chicago Stritch School of Medicine, Maywood, IL 60153, USA
[4]Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA
[5]Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA
[6]Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA
[7]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA
[8]Department of Human Genetics, University of Michigan, Ann Arbor, MI 48109, USA
[9]Department of Pediatrics and Communicable Diseases, University of Michigan, Ann Arbor, MI 48109, USA
[10]Howard Hughes Medical Institute, University of Michigan, Ann Arbor, MI 48109, USA

*To whom correspondence should be addressed
Room 5028
Life Sciences Institute
210 Washtenaw Avenue
Ann Arbor, MI 48109-2216
734-647-4808
ginsburg@umich.edu

Table of Contents

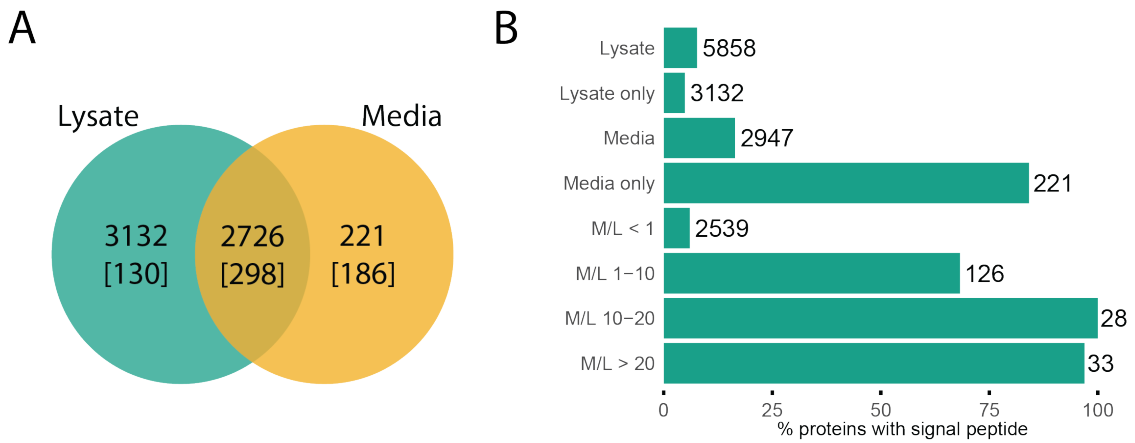

Figure S1: Comparison of protein abundance in the cell media and lysates improves identification of secreted proteins.(A) Venn diagram of proteins detected in conditioned media, cell lysates, or both fractions across all samples (LMAN1-defient, SURF4-deficient, and NT-treated cells, n = 3 per group). Numbers in the bracket represent the number of proteins carrying an annotated signal peptide in each group. **(B)** Proteins were separated into discrete media/lysate (M/L) ratio intervals and the percentage of proteins carrying a signal peptide was calculated. Numbers at the end of each bar indicate the number of proteins in each bin.

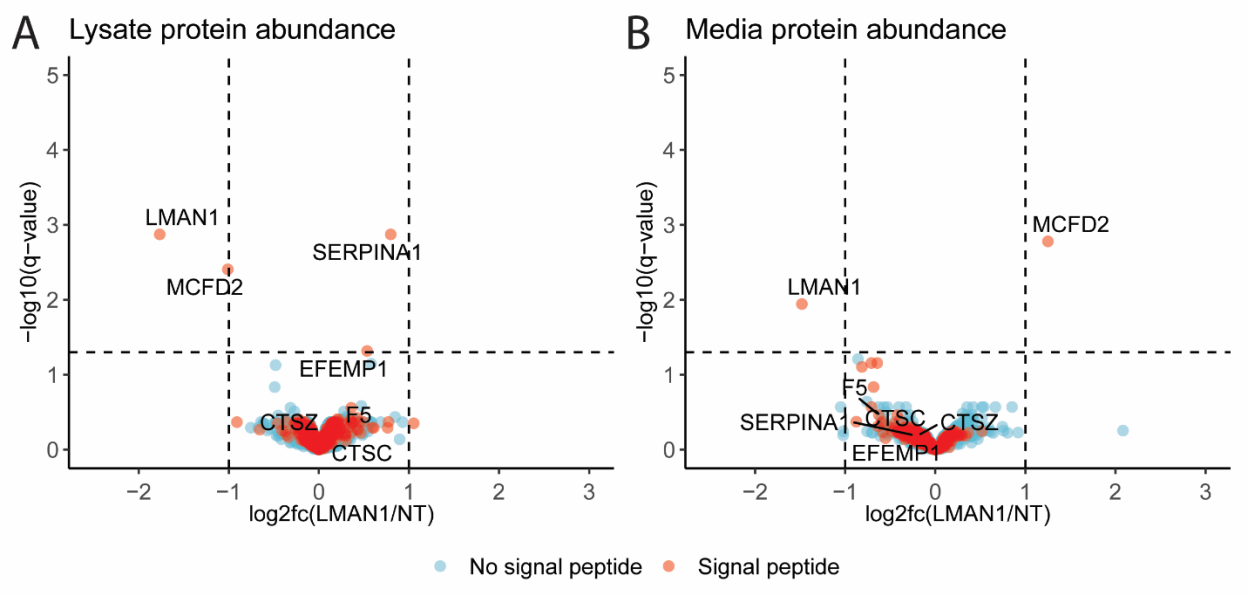

Figure S2: Comparison of proteins abundance in the lysates and media between LMAN1 deficient and control cells. Volcano plots representing changes in protein abundance in the cell lysates **(A)** or conditioned media **(B).** The log2 fold change (log2fc) and statistical significance are plotted on the x and y-axis, respectively. Proteins with a signal peptide are colored in red and proteins without a signal peptide are colored in blue. Dashed vertical lines represent the log2fc of 1 and -1. Dashed horizontal lines represent the -log10(q-value of 0.05).

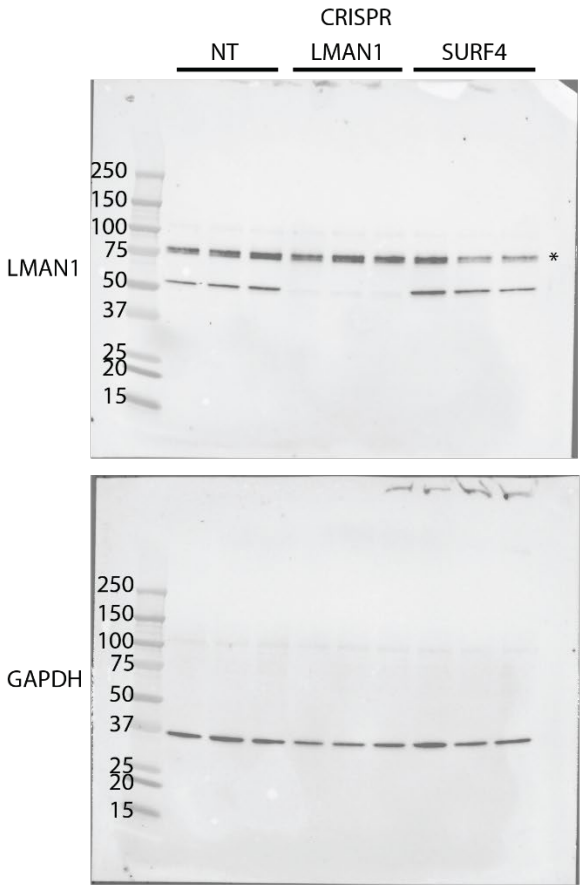

Figure S3: Efficient depletion of LMAN1 in cells treated with *LMAN1* targeting gRNA.

Immunoblots of cell lysates collected from controls (NT) and *LMAN1* or *SURF4* deleted cells. Numbers represent protein molecular weights in kDa. Asterisk indicates nonspecific binding of LMAN1 antibody.

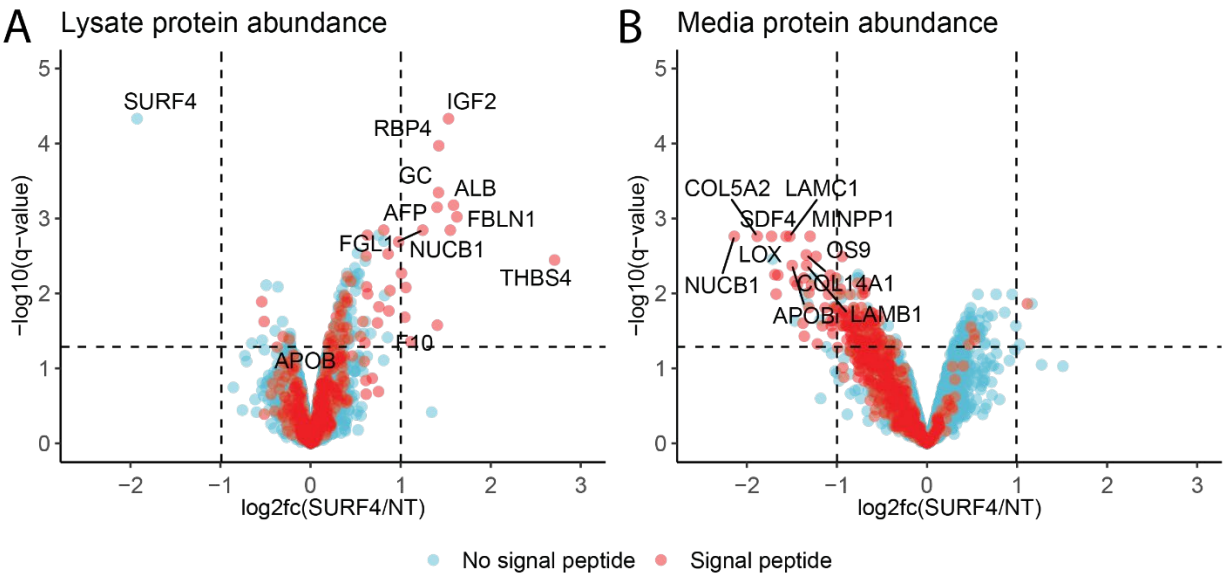

Figure S4: Comparison of proteins abundance in the lysates and media between SURF4 deficient and control cells. Volcano plots representing changes in protein abundance in the cell lysates **(A)** or conditioned media **(B).** The log2 fold change (log2fc) and statistical significance are plotted on the x and y-axis, respectively. Proteins with a signal peptide are colored in red and proteins without a signal peptide are colored in blue. Dashed vertical lines represent the log2fc of 1 and -1. Dashed horizontal lines represent the -log10(q-value of 0.05).

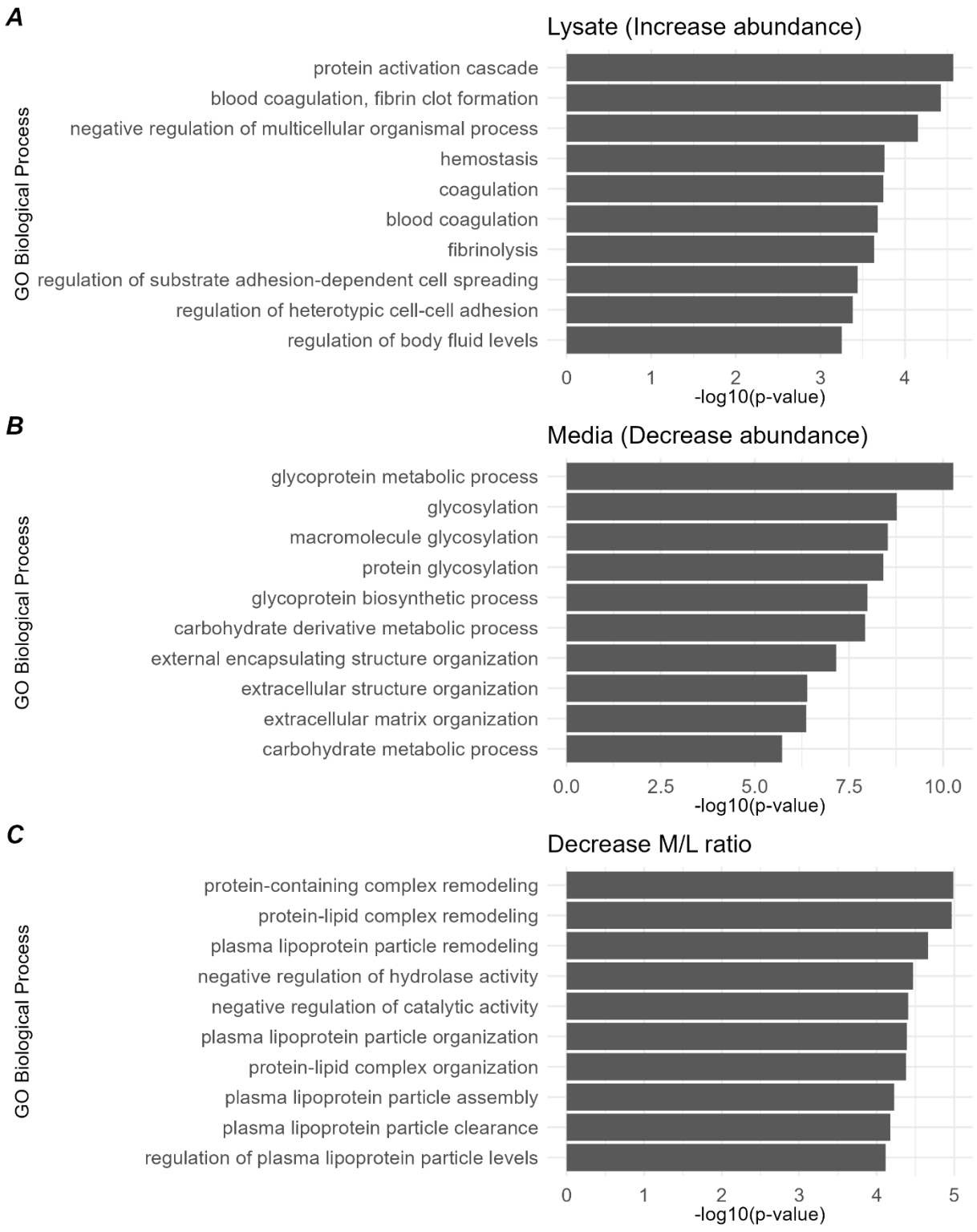

Figure S5: Gene Ontology (GO) enrichment analysis for biological process terms using a list of proteins that demonstrated increase lysate abundance **(A)**, decrease media abundance **(B)**, and decrease M/L ratio **(C)**

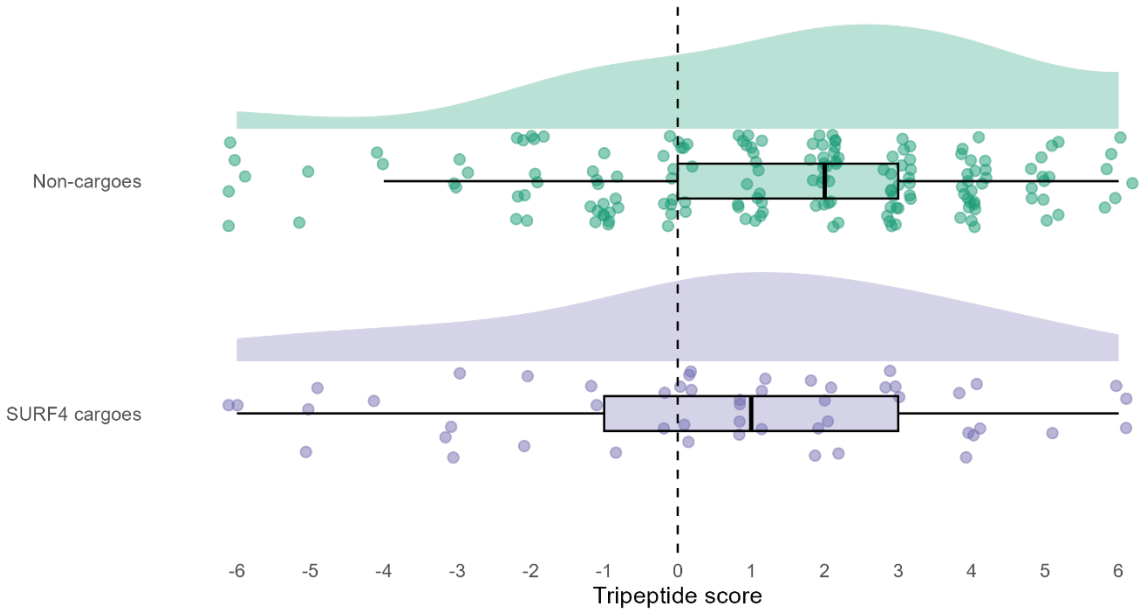

Figure S6: Proteins that are not SURF4 cargoes on average carry better ER-ESCAPE motif than putative SURF4 cargoes. To determine whether ER-ESCAPE motifs (as defined by Yin et al. [29]) are enriched in putative SURF4 cargoes, we assigned a score for each amino acid residue in the tripeptide motif (-2 for “bad” residue, -1 for “not good” residue, 0 for “neutral” residue, 1 for “good” residue, and 2 for “very good” residue, see **Figure 3**) and calculate a tripeptide score for each protein by summing up the score for each individual residue. The tripeptide scores range from -6 (a tripeptide with 3 “bad” residues) to 6 (a tripeptide with 3 “very good” residues).