TRANSPIRE: A computational pipeline to elucidate intracellular protein movements from spatial proteomics datasets

Michelle A. Kennedy, William A. Hofstadter, and Ileana M. Cristea*

Department of Molecular Biology, Princeton University, Washington Road, Princeton, NJ 08544

* Corresponding author

Address reprint requests to:

Ileana M. Cristea

210 Lewis Thomas Laboratory

Department of Molecular Biology

Princeton University

Princeton, NJ 08544

Email: icristea@princeton.edu

Tel: 6092589417

Fax: 609258457

SUPPLEMENTAL FIGURES

FIGURE S1

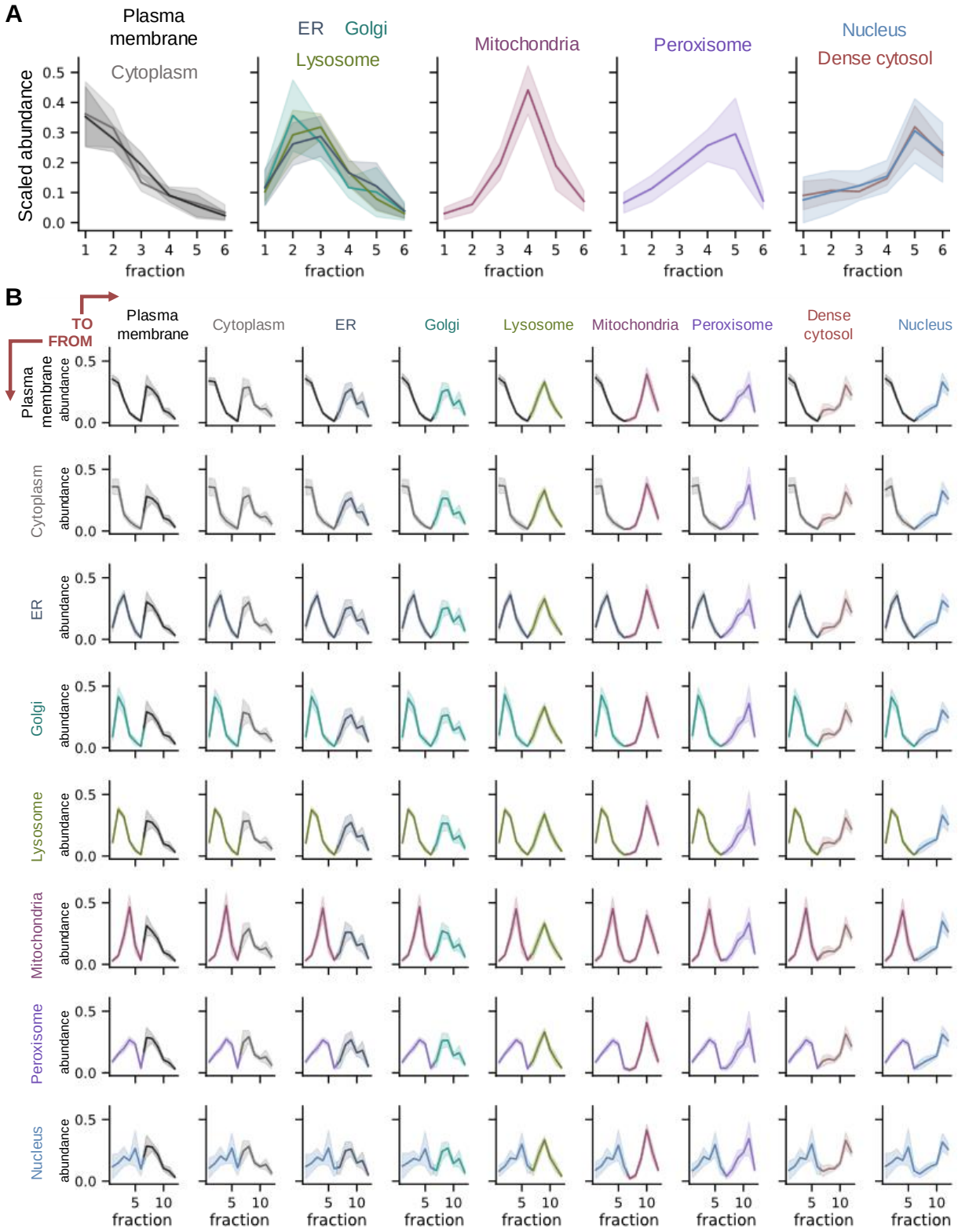

A
Plasma membrane
Cytoplasm
ER
Golgi
Lysosome
Mitochondria
Peroxisome
Nucleus
Dense cytosol
Scaled abundance
fraction
B
TO
FROM
Plasma membrane
Cytoplasm
ER
Golgi
Lysosome
Mitochondria
Peroxisome
Dense cytosol
Nucleus
abundance
fraction

FIGURE S1

Figure S1: Density fractionation of cellular organelles results in overlap between organelle and translocation profiles. A) Spatial profiles of organelle markers across timepoints throughout HCMV infection. **B)** Synthetic translocation profiles generated using the organelle markers from above. Note that fractions 1-6 correspond to the uninfected state, while fractions 7-12 correspond to the HCMV-infected state. Solid lines and shaded region represent the mean and standard deviation, respectively, of marker profiles for a given organelle or spatial profiles for a given translocation.

FIGURE S2

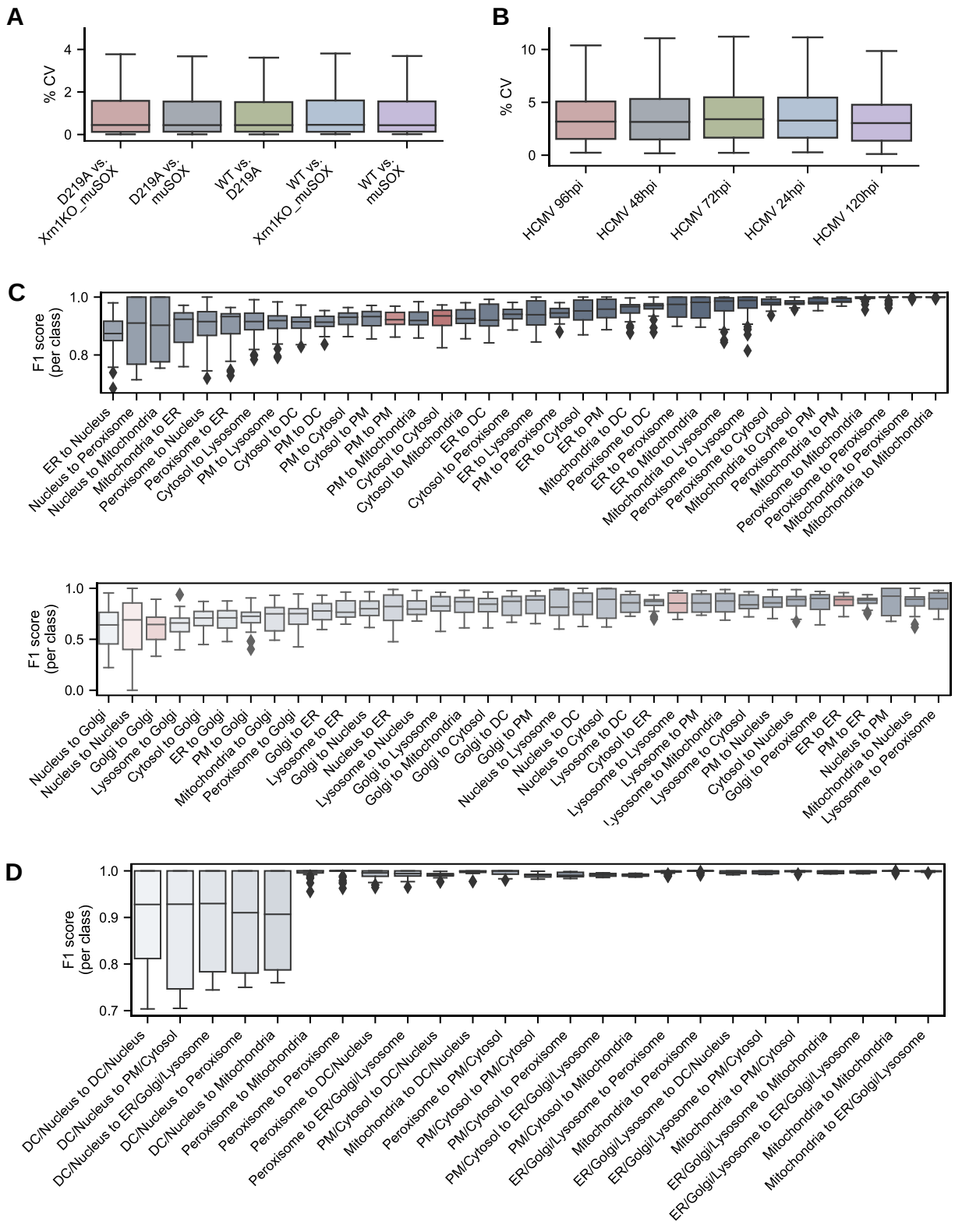


Figure S2: Assessing classifier performance and reliability. Variance in classifier predictions due to use of the soft-max likelihood function, which can only provide a

stochastic estimate of model variational expectations. Boxplots represent the distribution of 100 individual evaluations of the trained classifier on either the **(A)** Gilbertson et al. 2018 or **(B)** Jean Beltran et al. 2016 datasets. **C)** Classifier performance on the test data partition from analysis of Jean Beltran et al. 2016. Unweighted F1 scores are reported on a per-class basis for classifier predictions across all conditions analyzed in the dataset. Red boxes highlight non-translocating classes, while blue boxes denote translocation classes. **D)** Unweighted F1 scores when ambiguous organelles are grouped together.

FIGURE S3

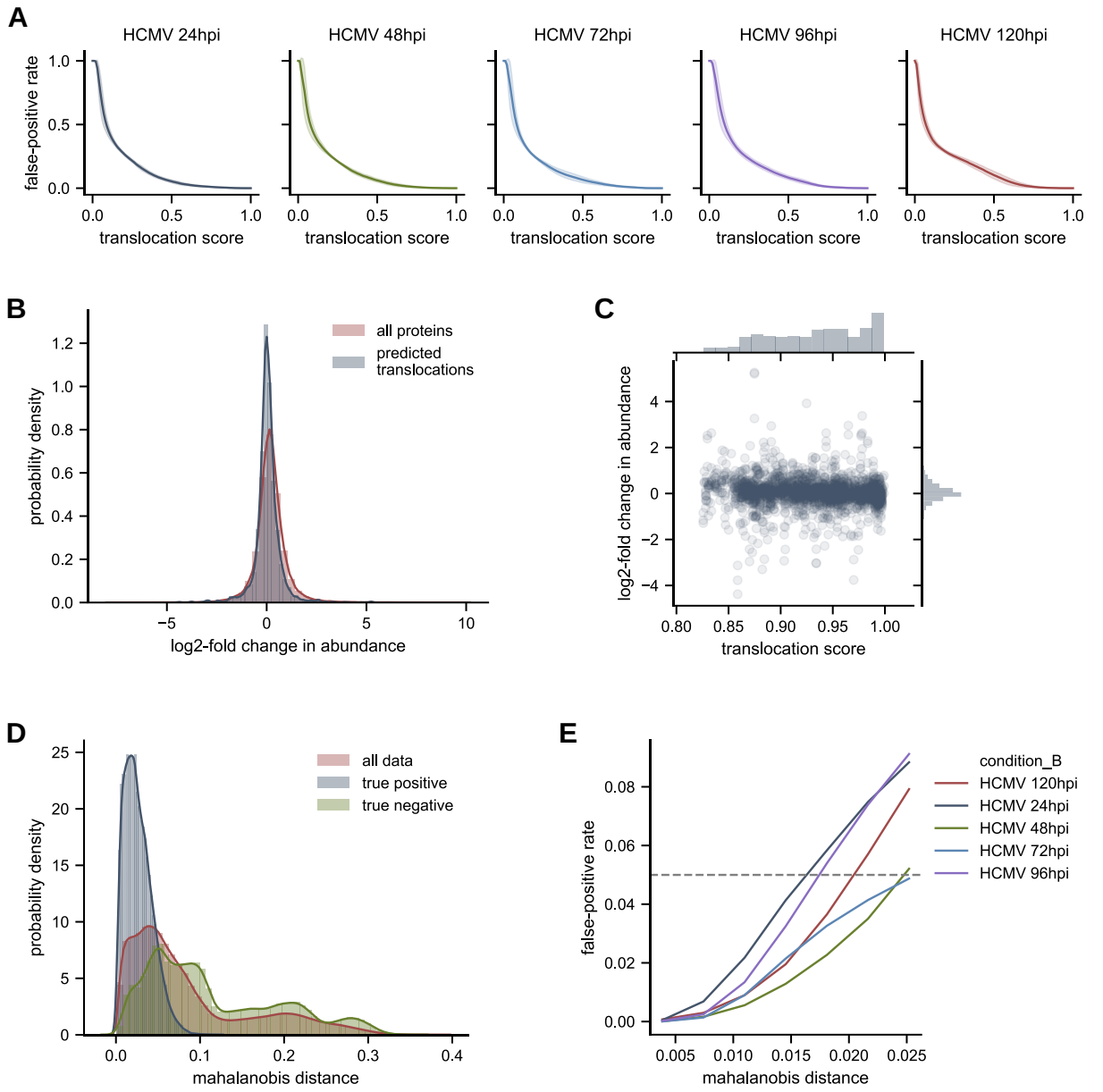
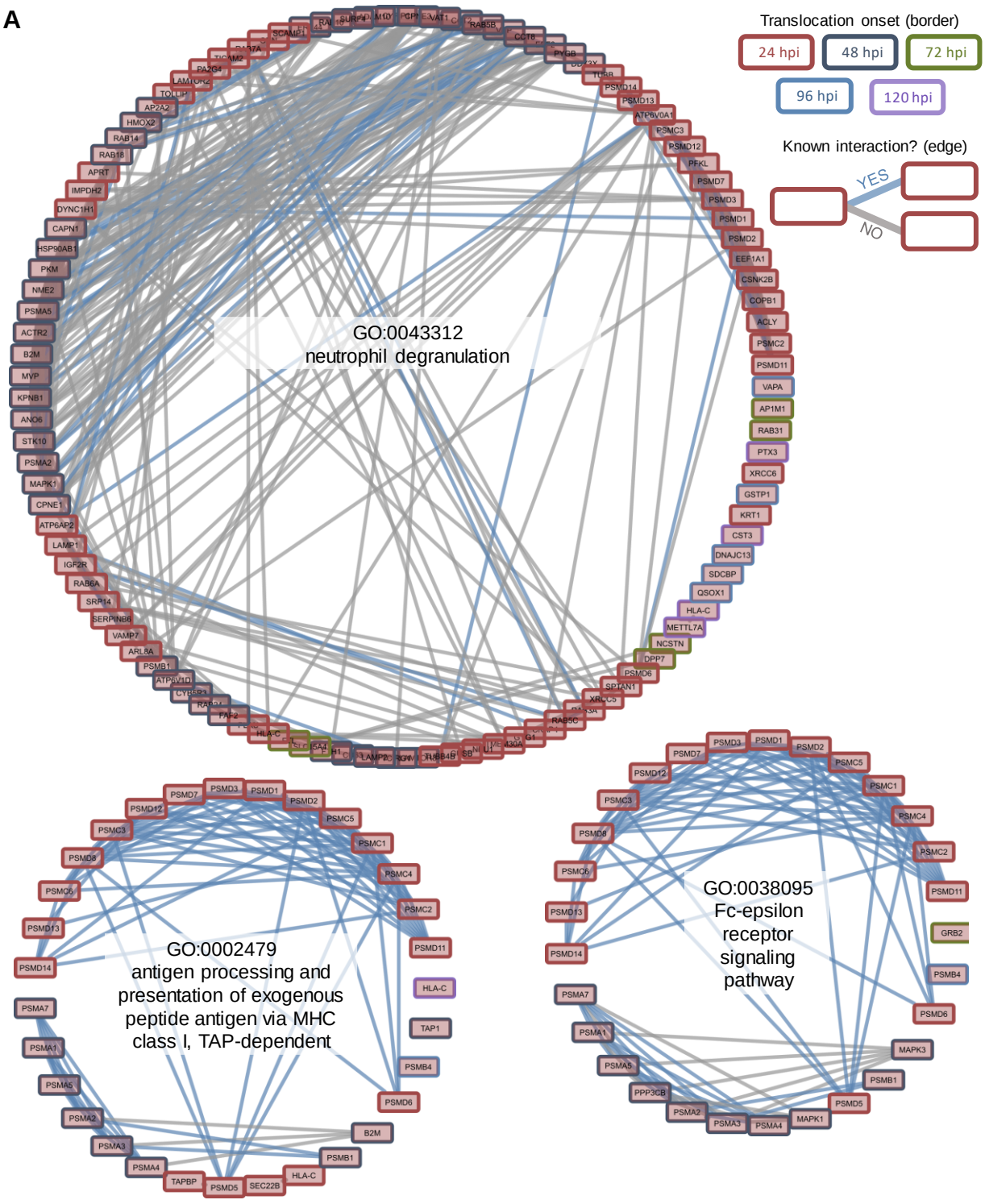


Figure S3: Determining an appropriate translocation and co-translocation score/distance cutoffs and investigating cellular abundance of translocating proteins. A) False-positive rates for translocation prediction based on classifier performance on the test data. **B)** Change in protein abundance upon infection for all proteins in the HCMV dataset (red) vs. TRANSPIRE predicted translocations (blue). **C)** Plot of translocation score vs. change in abundance for proteins predicted to translocate. **D)** Mahalanobis distance calculations between all predicted translocations ("all data," red), translocating proteins annotated to the same CORUM complex ("true positive," blue), and marker proteins for different subcellular organelles ('true negative," green). **E)** False-positive rate vs. mahalanobis distance as calculated from the distributions in D on a per-condition basis.

FIGURE S4



(continued on following page)

FIGURE S4

A *(continued from previous page)*

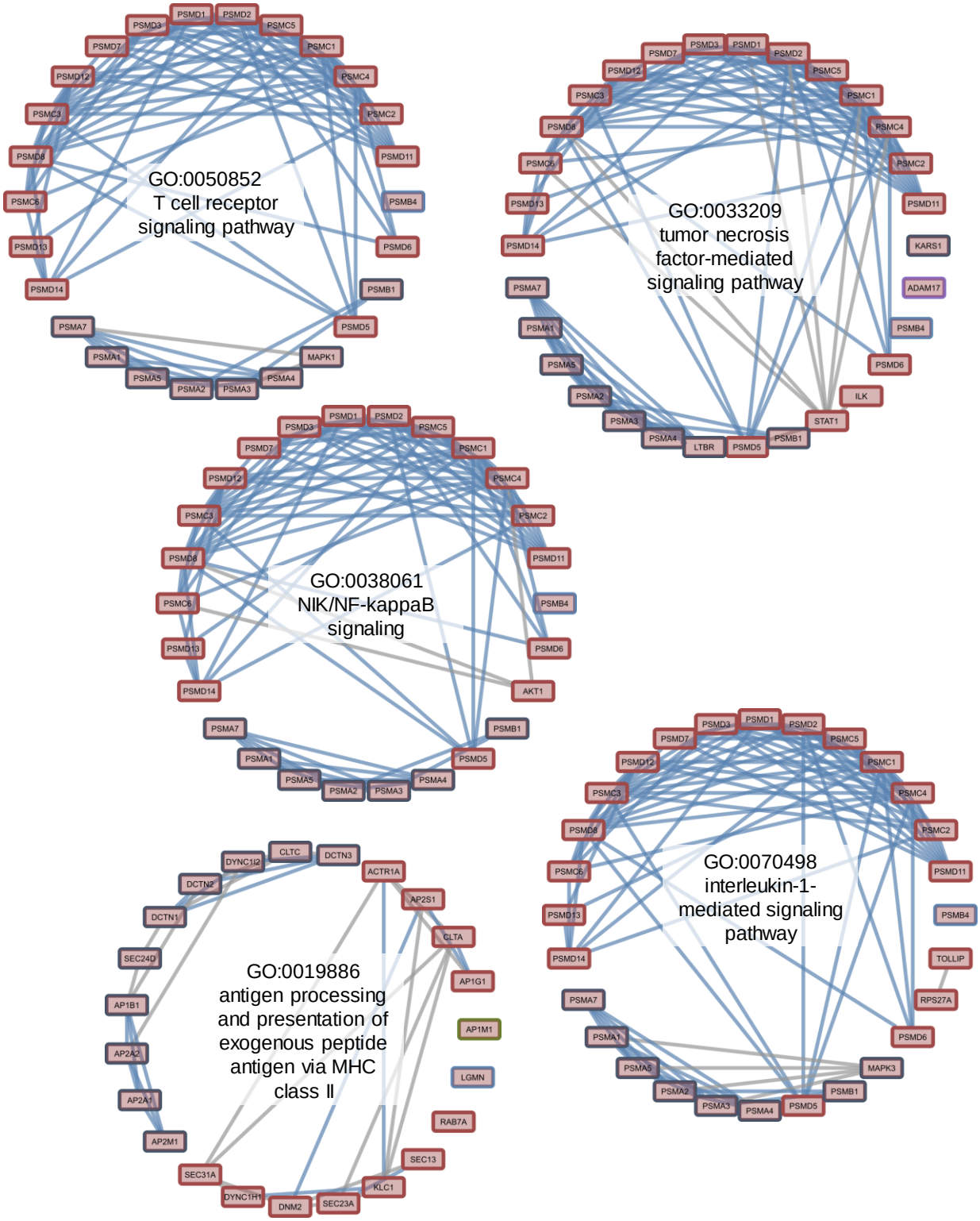
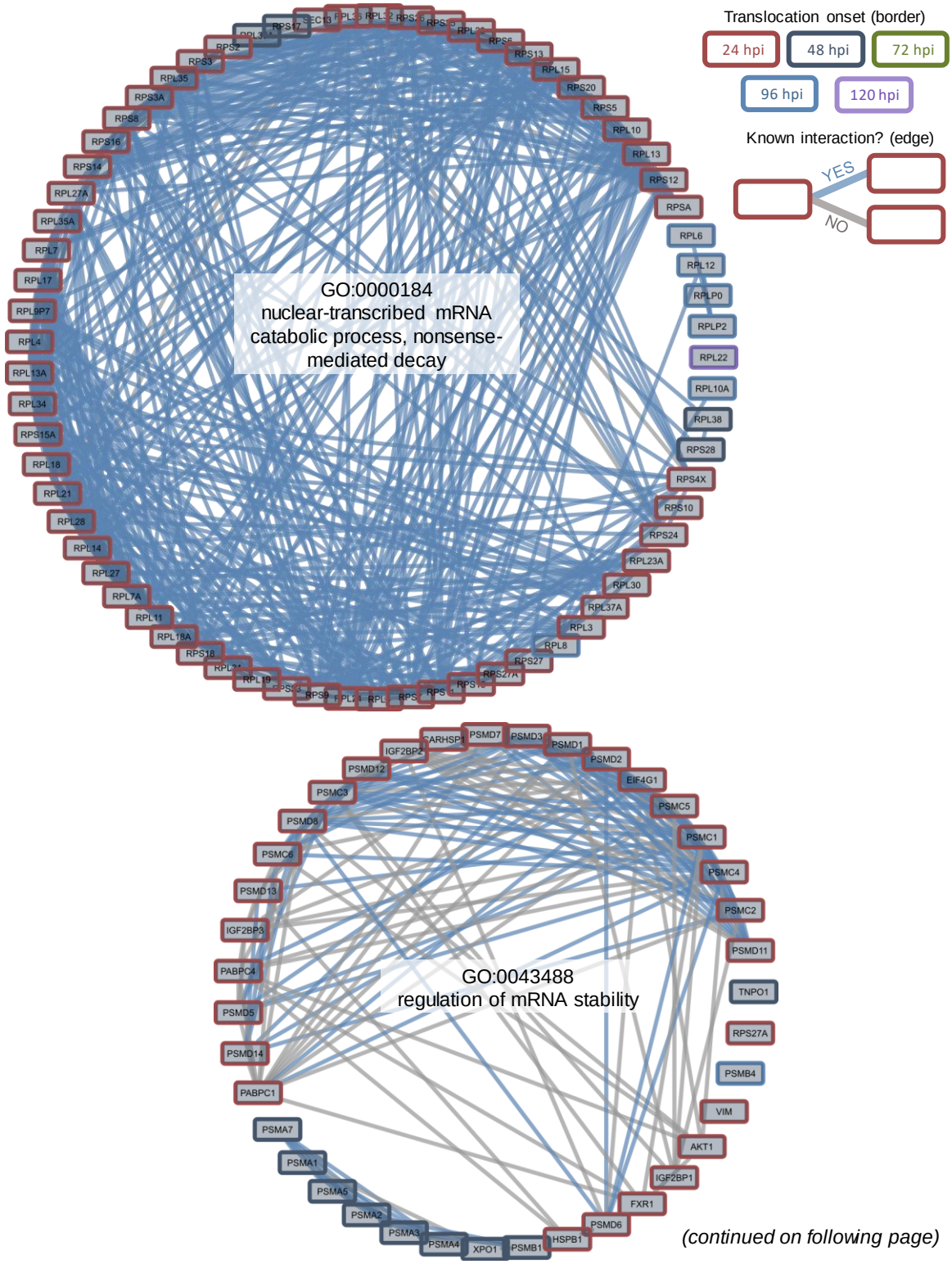


FIGURE S4

B



(continued on following page)

FIGURE S4

B *(continued from previous page)*

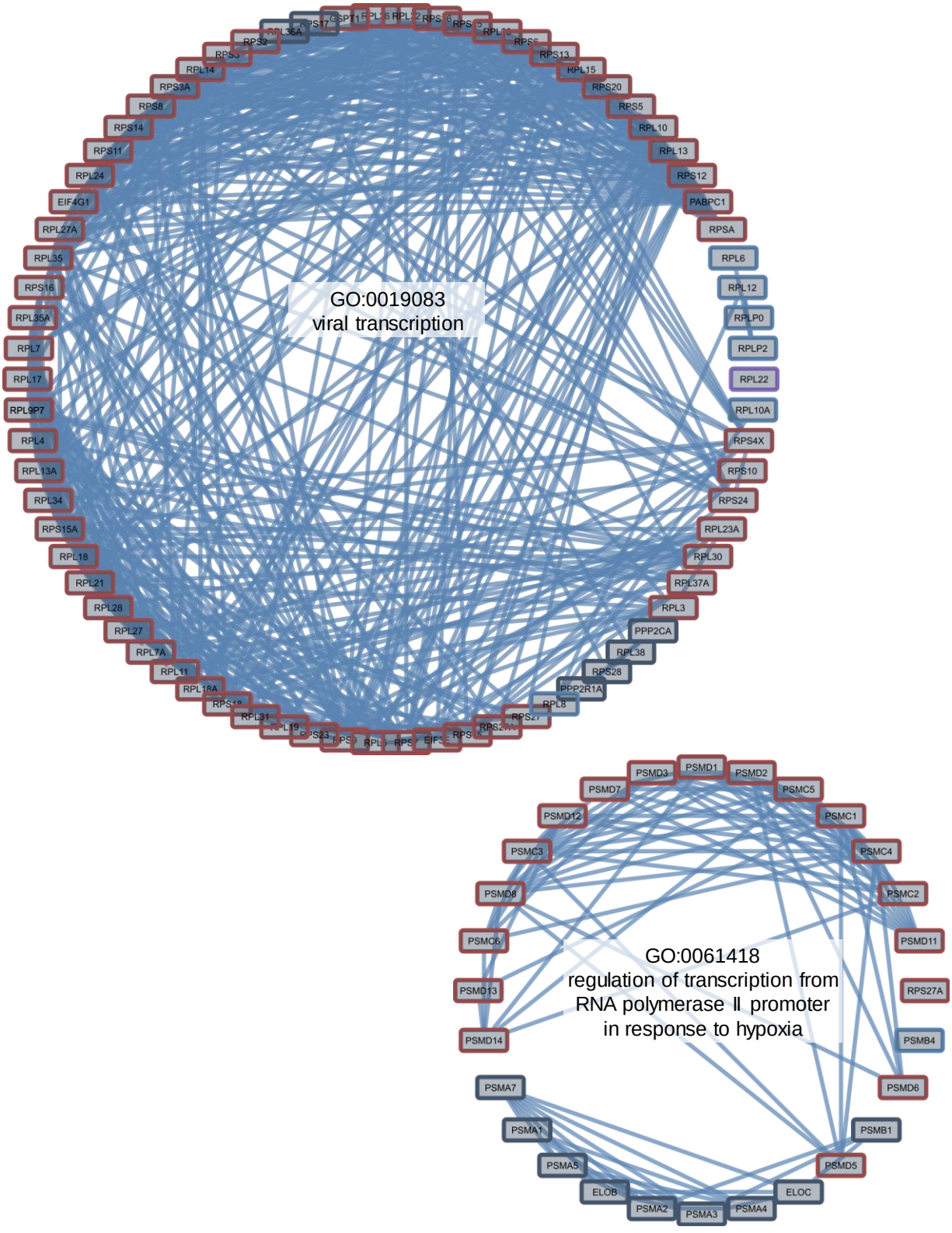
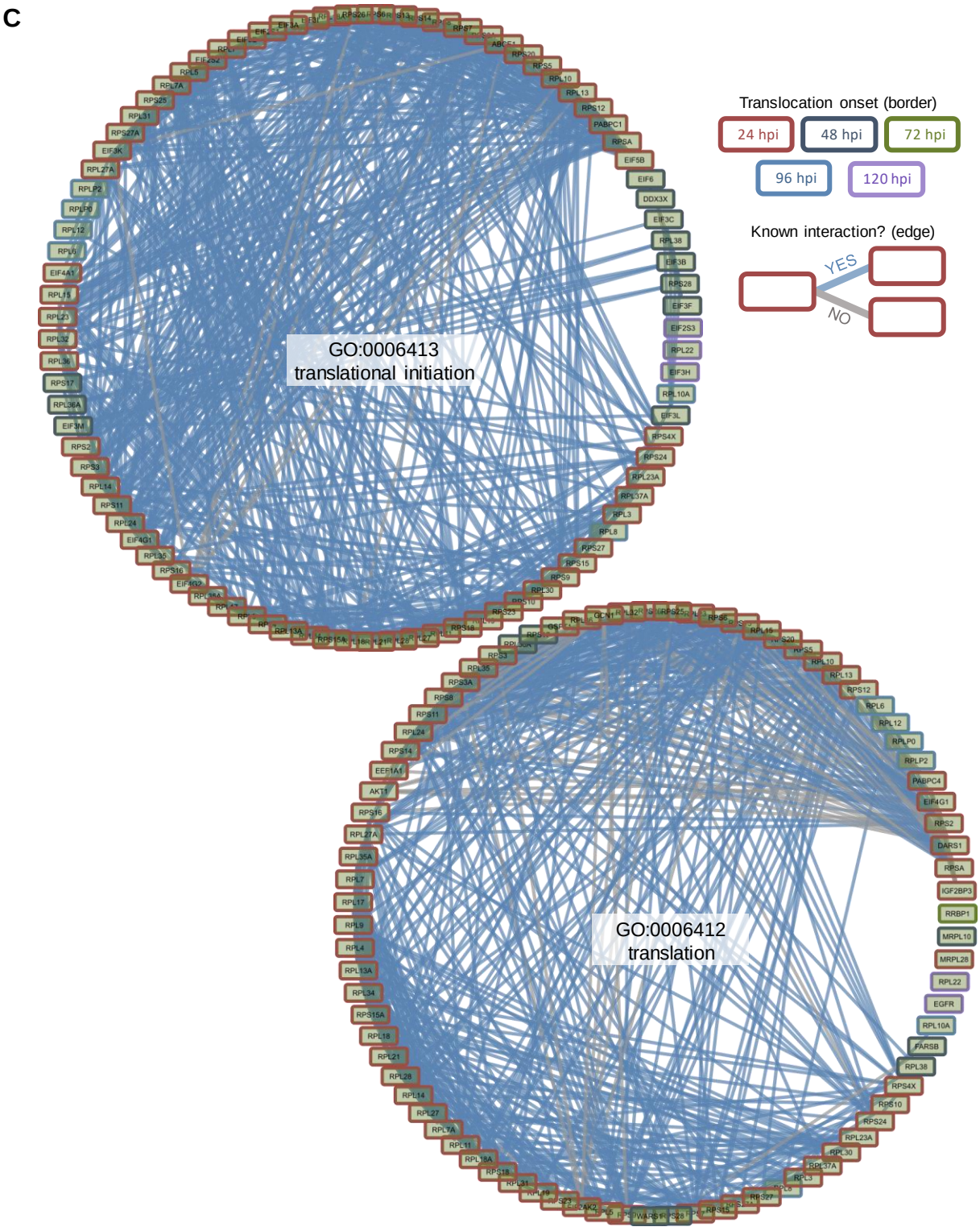


FIGURE S4



(continued on following page)

FIGURE S4

C *(continued from previous page)*

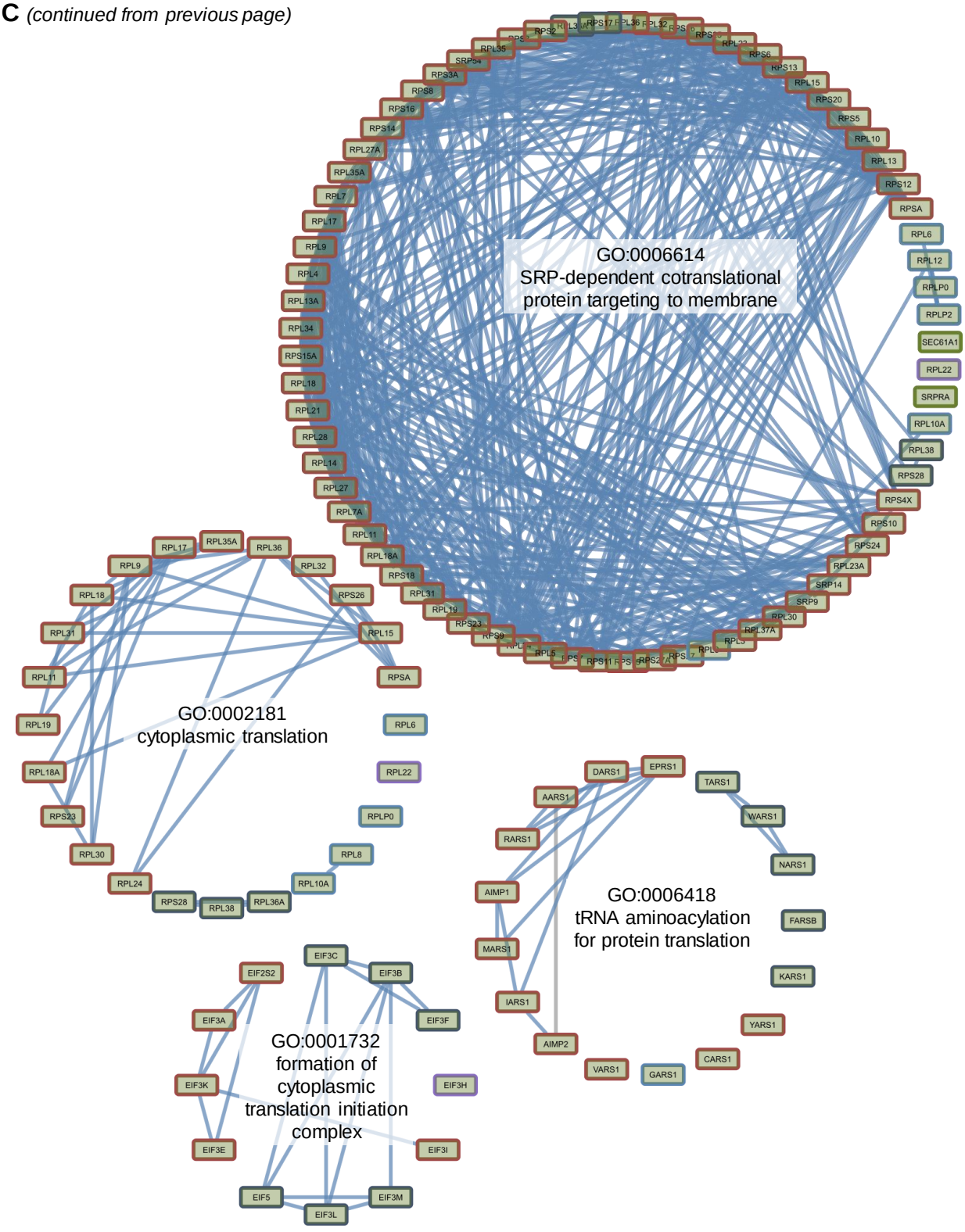


FIGURE S4

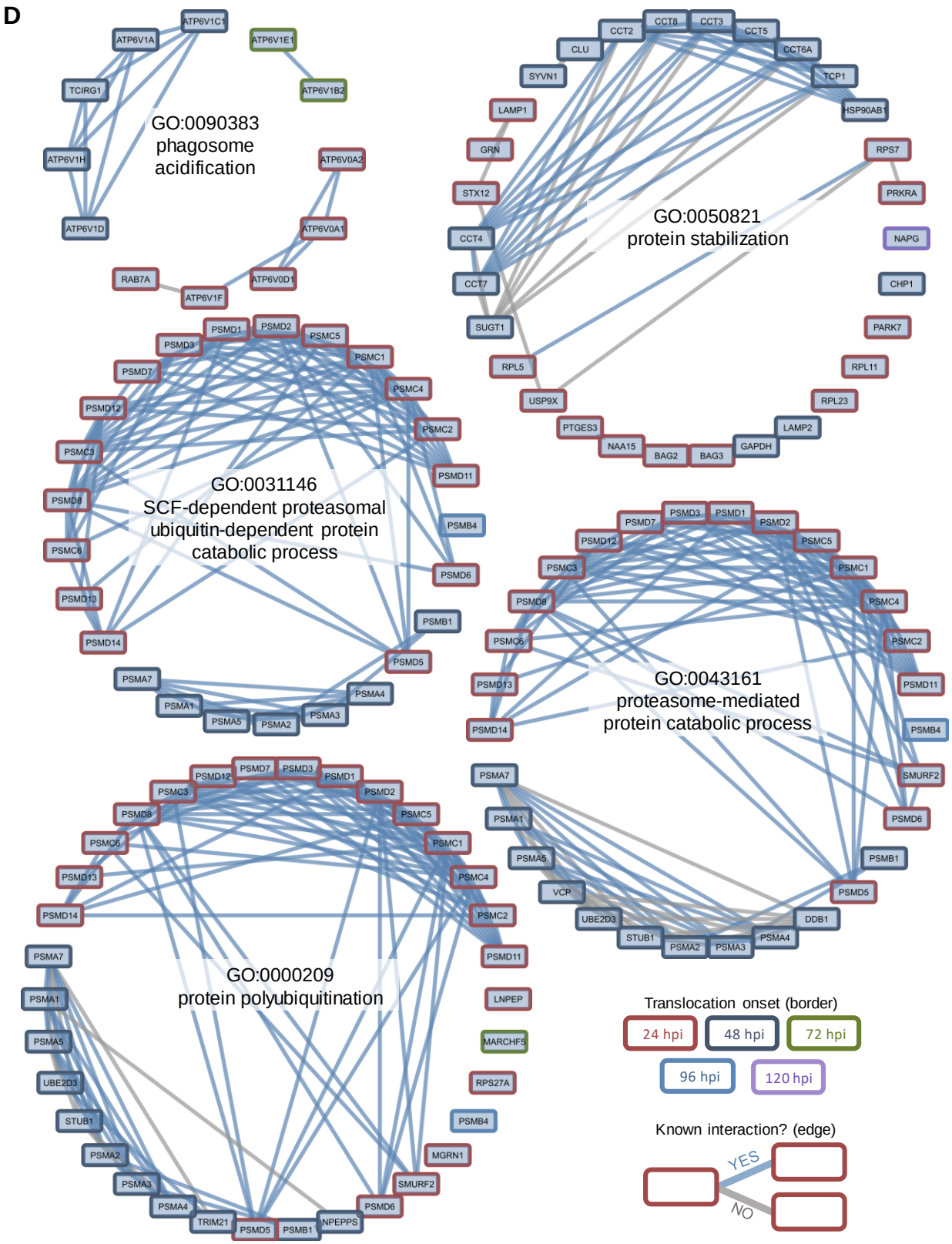
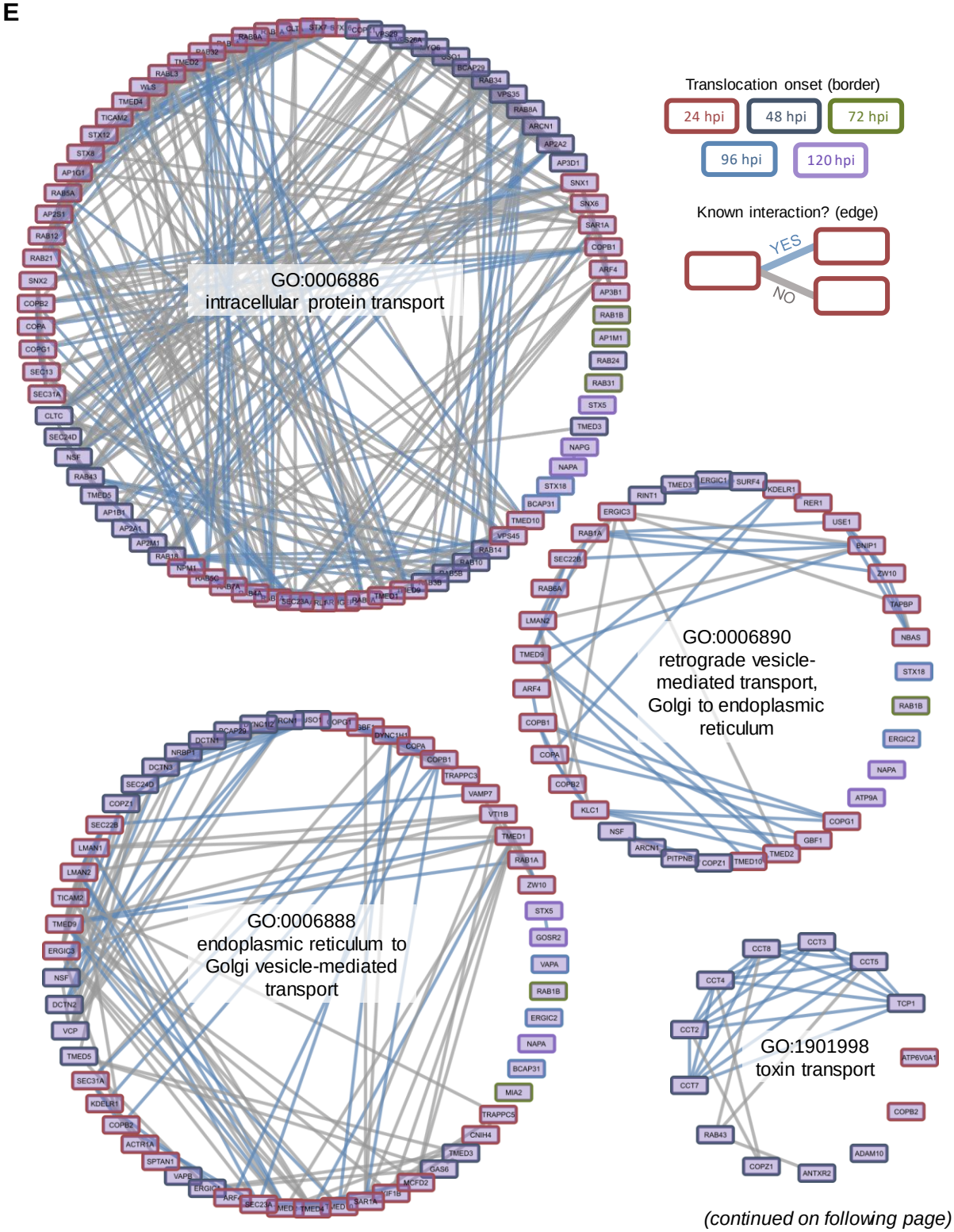

D
GO:0090383
phagosome
acidification
ATP6V1C1
ATP6V1A
ATP6V1E1
ATP6V1B2
TCIRG1
ATP6V1H
ATP6V0A2
ATP6V1D
ATP6V0A1
RAB7A
ATP6V0D1
ATP6V1F
GO:0050821
protein stabilization
CCT2
CCT8
CCT3
CCT5
CLU
CCT6A
SYVN1
TCP1
LAMP1
HSP90AB1
GRN
RPS7
STX12
PRKRA
CCT4
NAPG
CCT7
CHP1
SUGT1
PARK7
RPL5
RPL11
USP9X
RPL23
PTGES3
LAMP2
NAA15
GAPDH
BAG2
BAG3
GO:0031146
SCF-dependent proteasomal
ubiquitin-dependent protein
catabolic process
PSMD1
PSMD2
PSMC5
PSMD3
PSMC1
PSMD7
PSMC4
PSMD12
PSMC2
PSMC3
PSMD11
PSMD8
PSMB4
PSMC6
PSMD6
PSMD13
PSMB1
PSMD14
PSMD5
PSMA7
PSMA4
PSMA1
PSMA3
PSMA5
PSMA2
GO:0043161
proteasome-mediated
protein catabolic process
PSMD7
PSMD3
PSMD1
PSMD2
PSMD12
PSMC5
PSMC3
PSMC1
PSMD8
PSMC4
PSMC6
PSMC2
PSMD13
PSMD11
PSMD14
PSMB4
SMURF2
PSMA7
PSMD6
PSMA1
PSMB1
PSMA5
PSMD5
VCP
DDB1
UBE2D3
PSMA4
STUB1
PSMA2
PSMA3
GO:0000209
protein polyubiquitination
PSMD7
PSMD3
PSMD12
PSMD1
PSMC3
PSMD2
PSMD8
PSMC5
PSMC6
PSMC1
PSMD13
PSMC4
PSMD14
PSMC2
PSMA7
PSMD11
PSMA1
LNPEP
PSMA5
MARCHF5
UBE2D3
RPS27A
STUB1
PSMB4
PSMA2
MGRN1
PSMA3
SMURF2
PSMA4
PSMD6
TRIM21
PSMD5
PSMB1
NPEPPS
Translocation onset (border)
24 hpi
48 hpi
72 hpi
96 hpi
120 hpi
Known interaction? (edge)
YES
NO

FIGURE S4

E



(continued on following page)

FIGURE S4

E *(continued from previous page)*

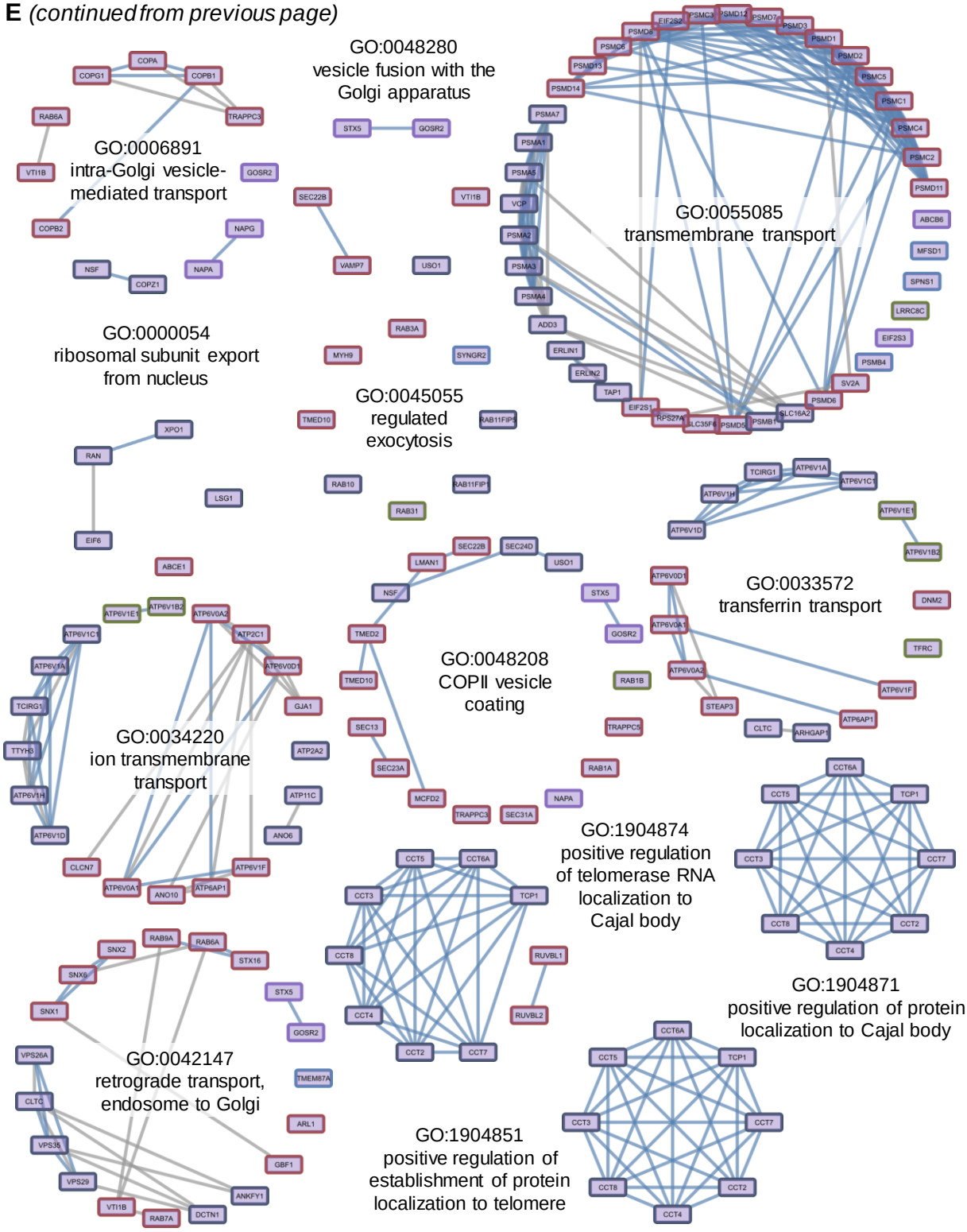
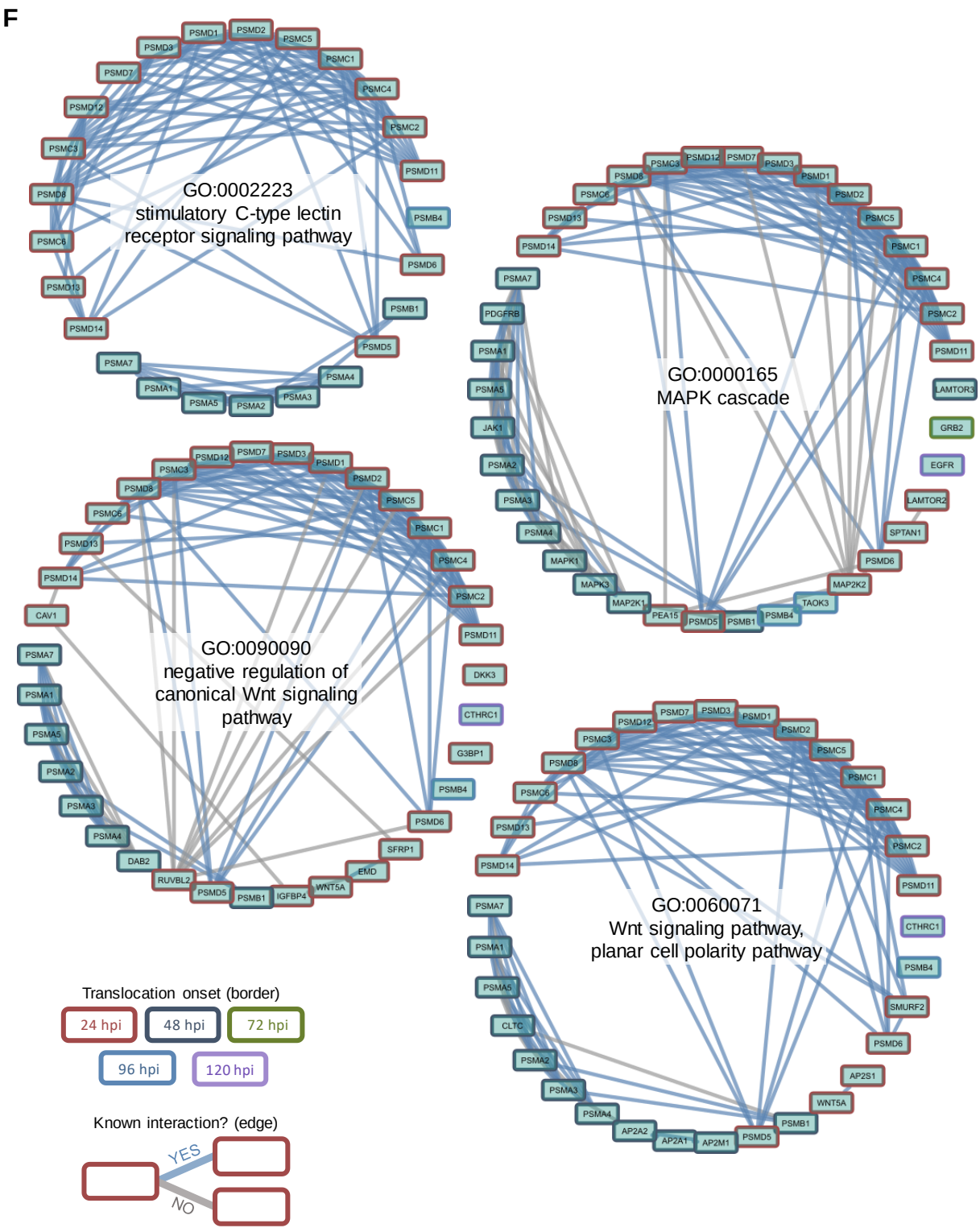


FIGURE S4



(continued on following page)

FIGURE S4

F (*continued from previous page*)

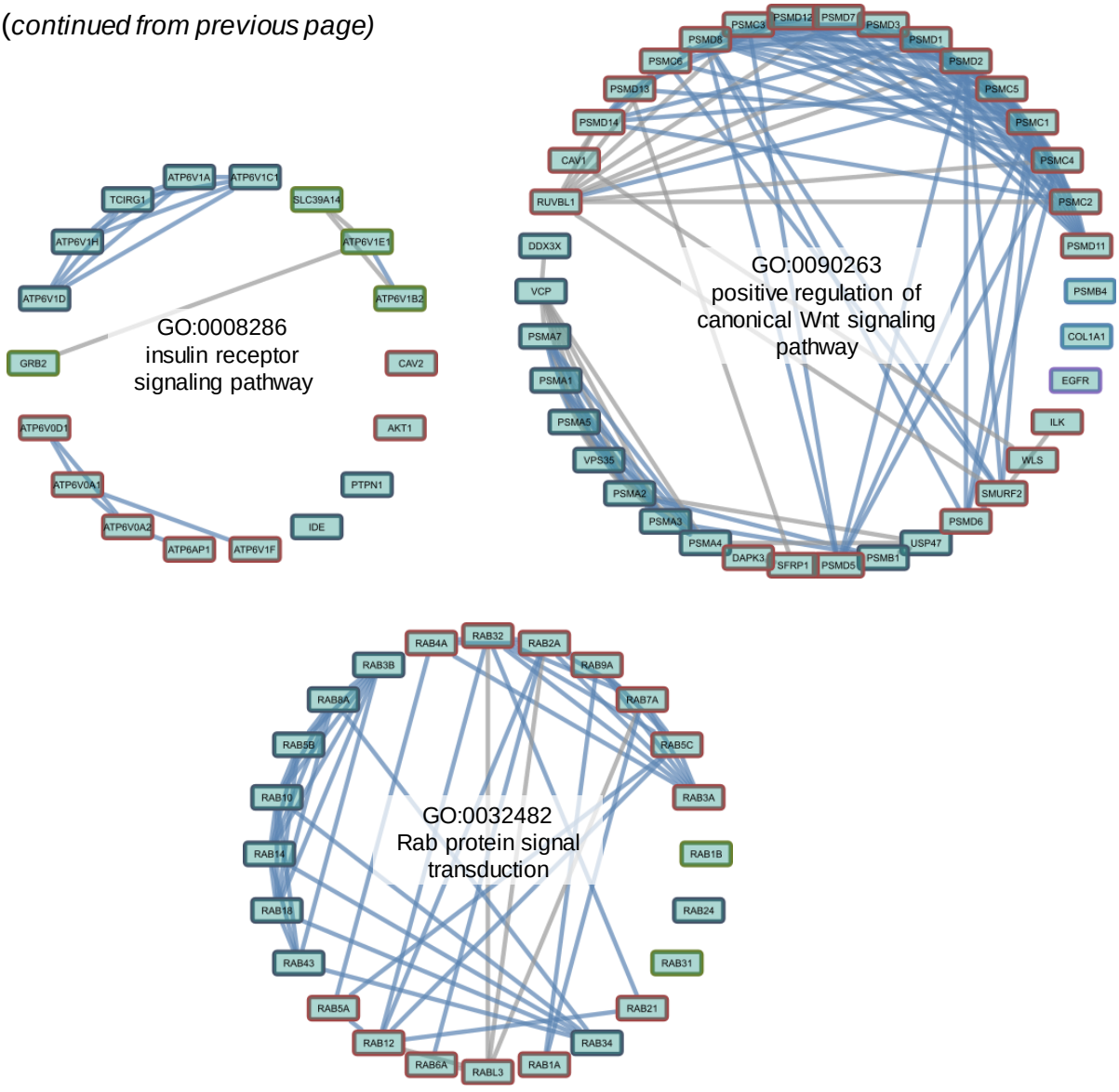


FIGURE S4

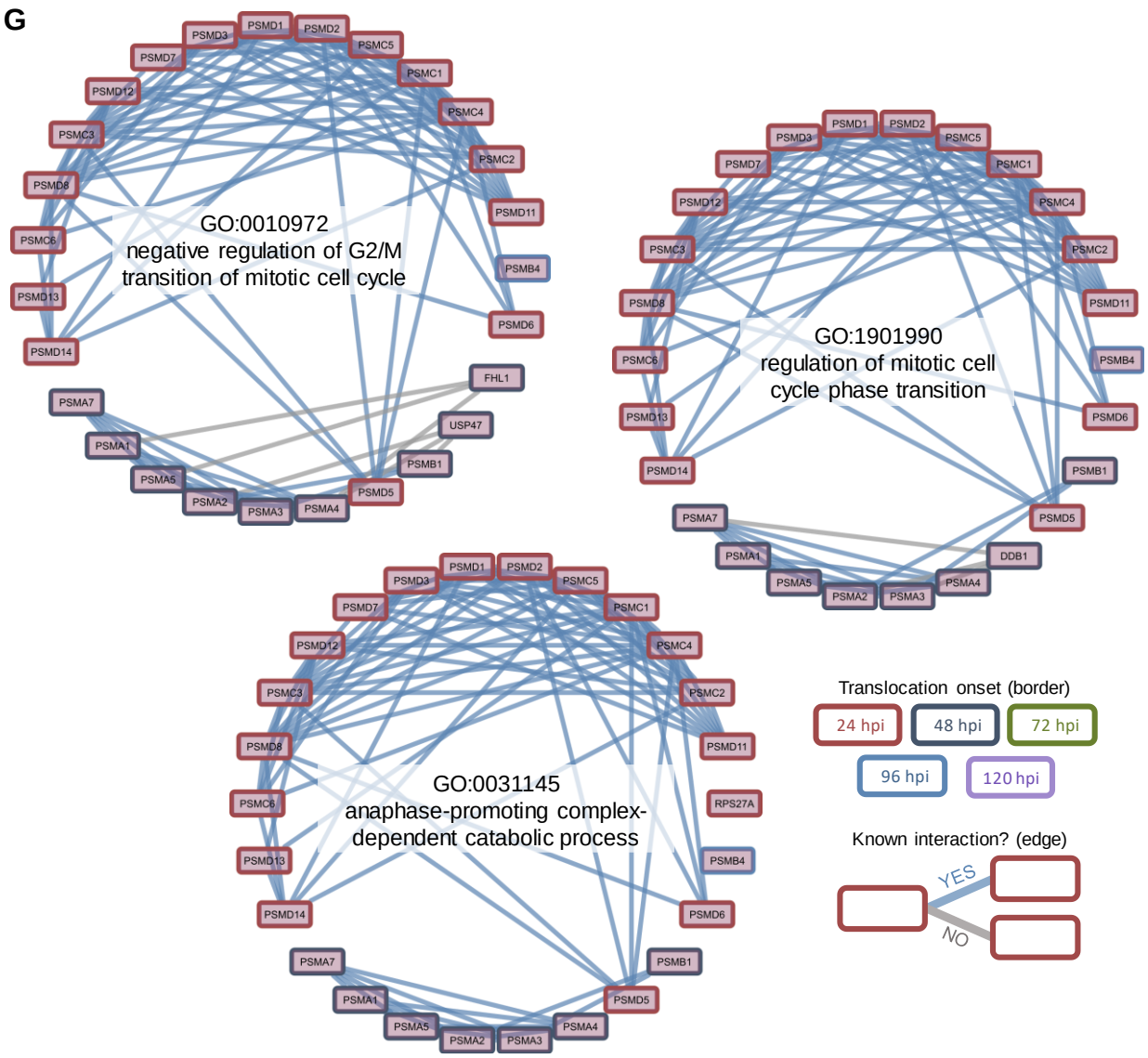

G
GO:0010972
negative regulation of G2/M
transition of mitotic cell cycle
PSMD1
PSMD2
PSMC5
PSMC1
PSMC4
PSMC2
PSMD11
PSMB4
PSMD6
FHL1
USP47
PSMB1
PSMD5
PSMA4
PSMA3
PSMA2
PSMA5
PSMA1
PSMA7
PSMD14
PSMD13
PSMC6
PSMD8
PSMC3
PSMD12
PSMD7
PSMD3
GO:1901990
regulation of mitotic cell
cycle phase transition
DDB1
GO:0031145
anaphase-promoting complex-
dependent catabolic process
RPS27A
Translocation onset (border)
24 hpi
48 hpi
72 hpi
96 hpi
120 hpi
Known interaction? (edge)
YES
NO

FIGURE S4

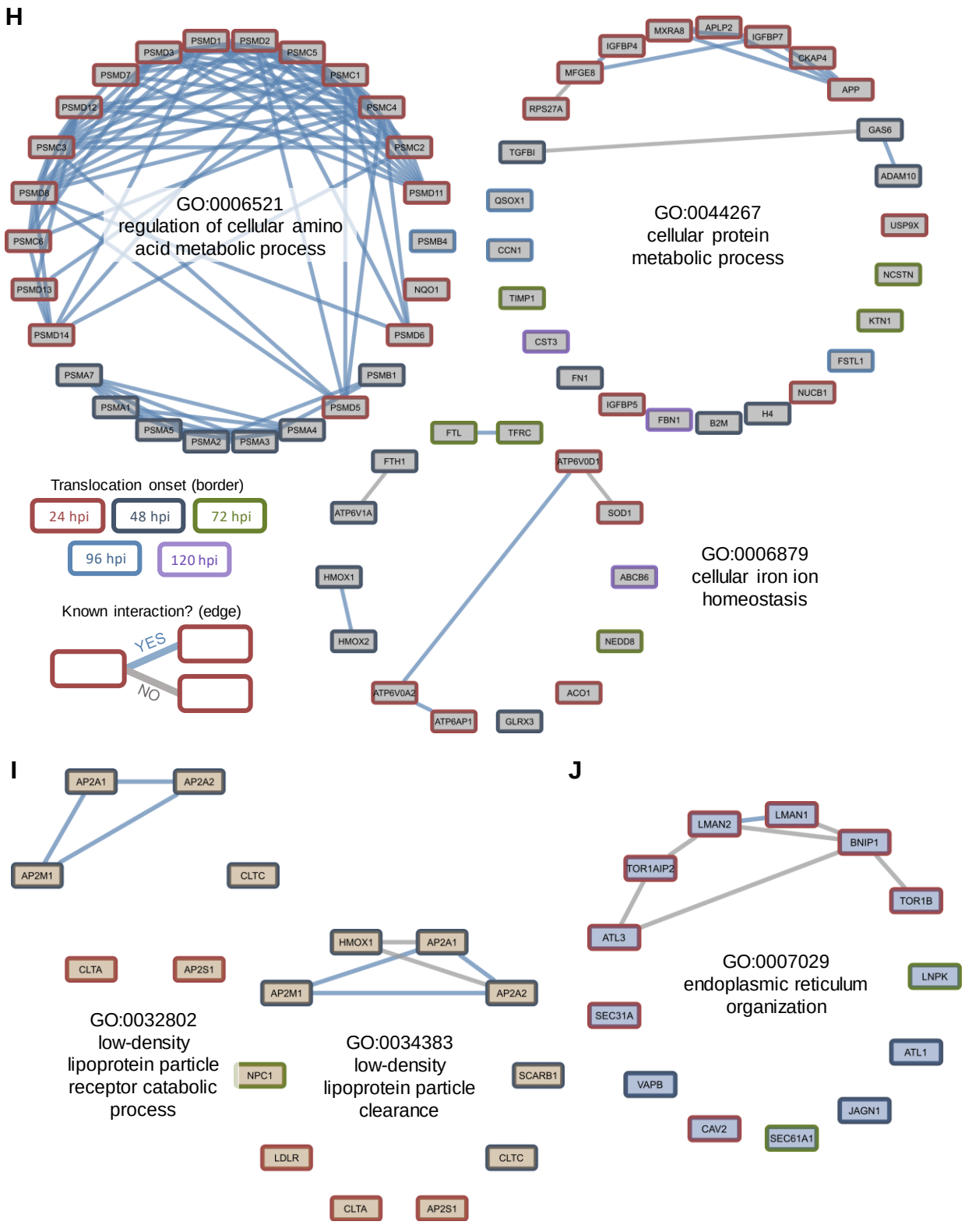

H
PSMD1
PSMD2
PSMD3
PSMC5
PSMD7
PSMC1
PSMD12
PSMC4
PSMC3
PSMC2
PSMD8
PSMD11
PSMC6
PSMB4
PSMD13
NQO1
PSMD14
PSMD6
PSMA7
PSMB1
PSMA1
PSMD5
PSMA5
PSMA2
PSMA3
PSMA4
GO:0006521
regulation of cellular amino acid metabolic process
MXRA8
APLP2
IGFBP4
IGFBP7
MFGE8
CKAP4
APP
RPS27A
GAS6
TGFBI
ADAM10
QSOX1
USP9X
CCN1
GO:0044267
cellular protein metabolic process
NCSTN
TIMP1
KTN1
CST3
FSTL1
FN1
NUCB1
IGFBP5
FBN1
B2M
H4
FTL
TFRC
FTH1
ATP6V0D1
ATP6V1A
SOD1
HMOX1
ABCB6
GO:0006879
cellular iron ion homeostasis
HMOX2
NEDD8
ATP6V0A2
ACO1
ATP6AP1
GLRX3
Translocation onset (border)
24 hpi
48 hpi
72 hpi
96 hpi
120 hpi
Known interaction? (edge)
YES
NO
I
AP2A1
AP2A2
AP2M1
CLTC
CLTA
AP2S1
GO:0032802
low-density lipoprotein particle receptor catabolic process
NPC1
HMOX1
AP2A1
AP2M1
AP2A2
GO:0034383
low-density lipoprotein particle clearance
SCARB1
LDLR
CLTC
CLTA
AP2S1
J
LMAN2
LMAN1
BNIP1
TOR1AIP2
TOR1B
ATL3
GO:0007029
endoplasmic reticulum organization
LNPK
SEC31A
ATL1
VAPB
JAGN1
CAV2
SEC61A1

FIGURE S4

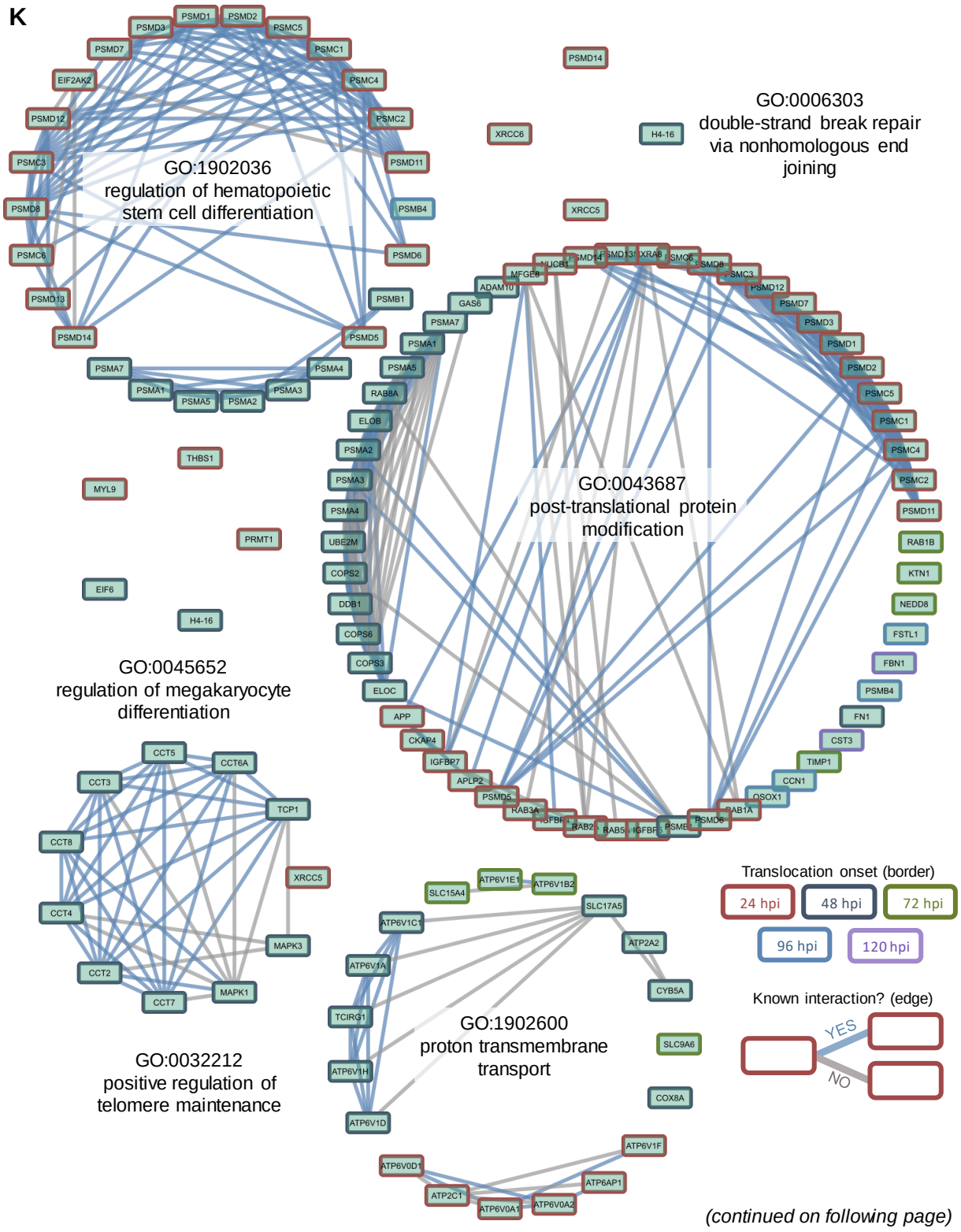

K
GO:1902036
regulation of hematopoietic
stem cell differentiation
GO:0006303
double-strand break repair
via nonhomologous end
joining
GO:0043687
post-translational protein
modification
GO:0045652
regulation of megakaryocyte
differentiation
GO:0032212
positive regulation of
telomere maintenance
GO:1902600
proton transmembrane
transport
Translocation onset (border)
24 hpi
48 hpi
72 hpi
96 hpi
120 hpi
Known interaction? (edge)
YES
NO
(continued on following page)

FIGURE S4

K (*continued from previous page*)

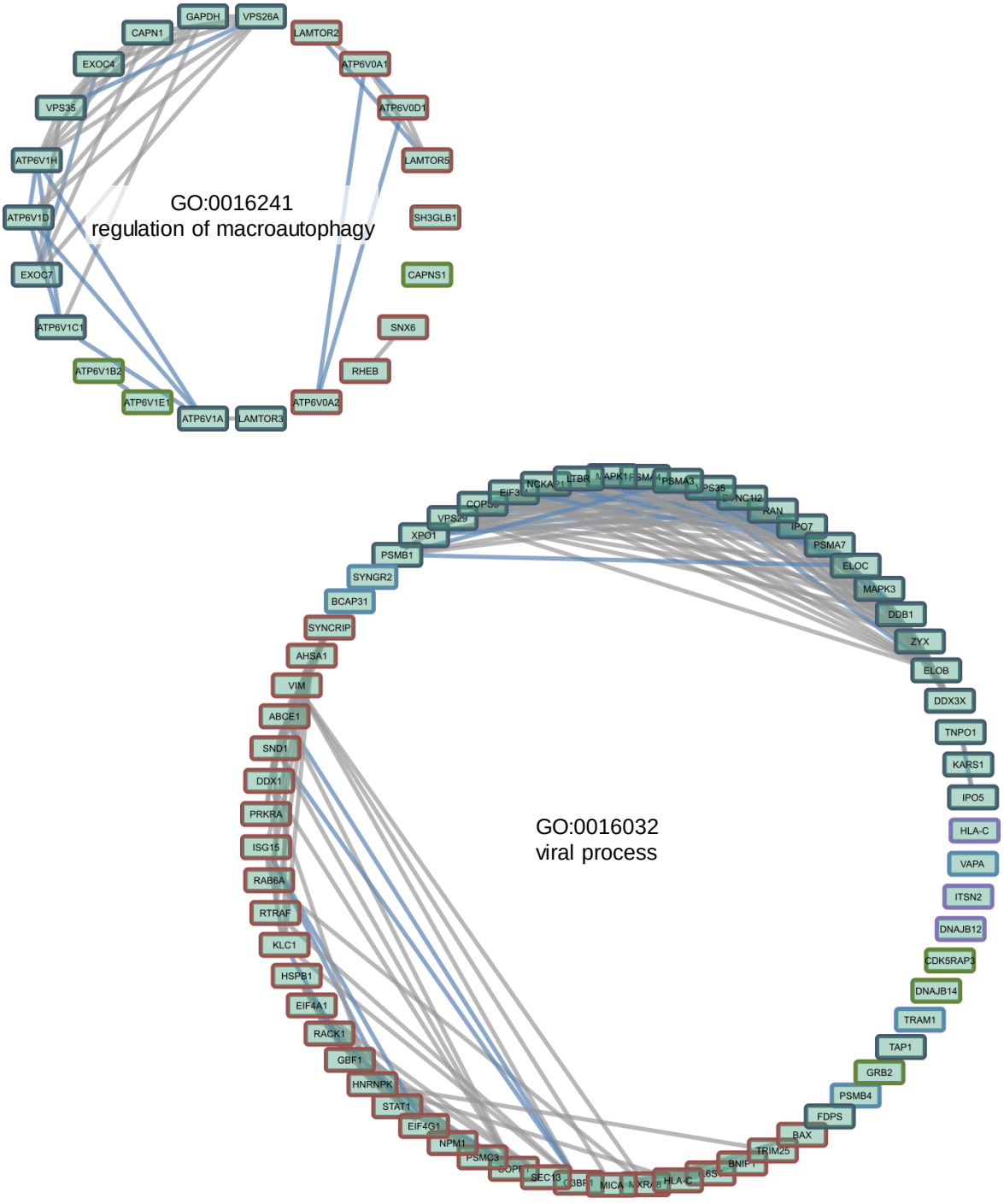
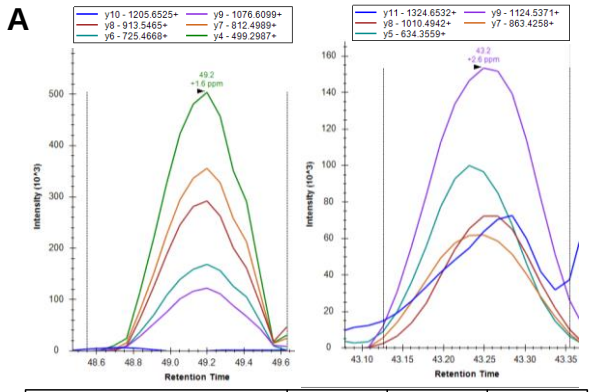
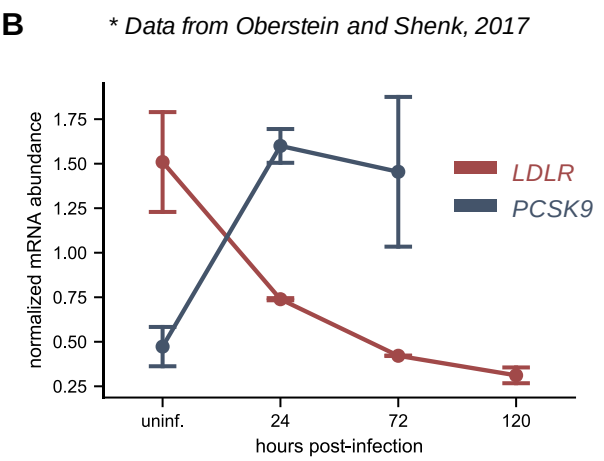


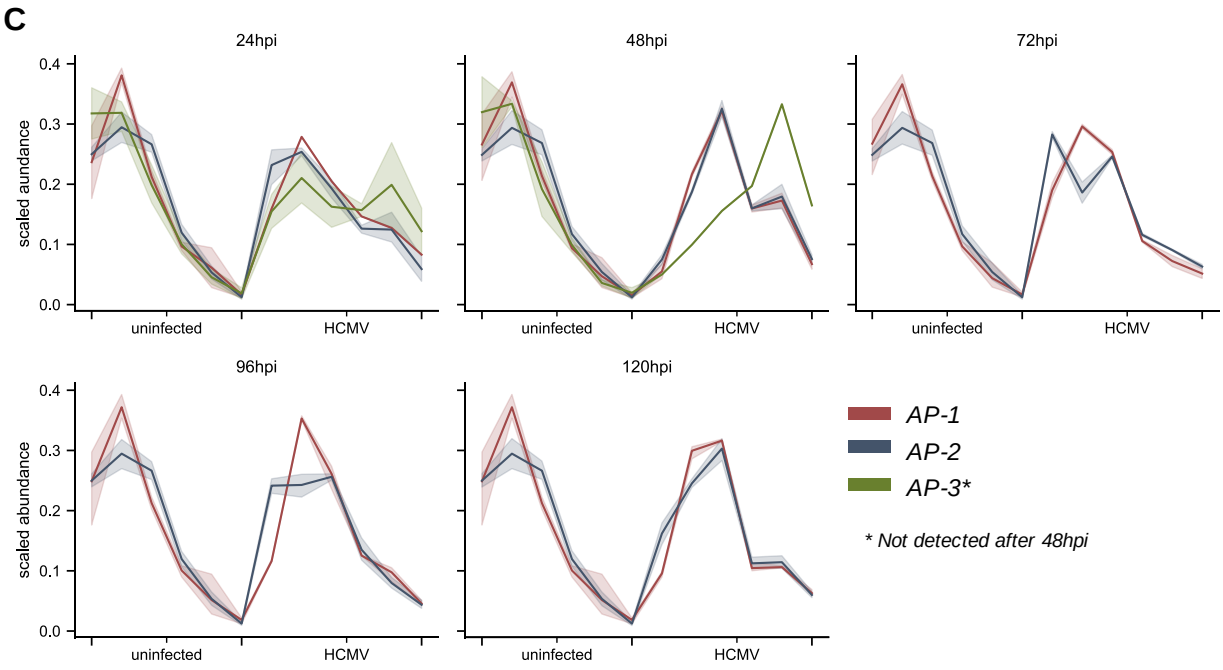
Figure S4: Enriched gene ontology associations for translocating proteins. Translocating proteins enriched for processes involved in **A)** immune signaling, **B)** transcription, **C)** translation, **D)** proteolysis, **E)** trafficking, **F)** signal transduction, **G)** cell cycle regulation, **H)** metabolism, **I)** cholesterol metabolism, **J)** subcellular organization, and **K)** “other” categories.

FIGURE S5



Unique peptide sequence	Uninf.	48 hpi	96 hpi
NINSINFDNPVYQK	1.93	0.68	0.12
SEYTSLIPNLR	1.55	0.86	0.17





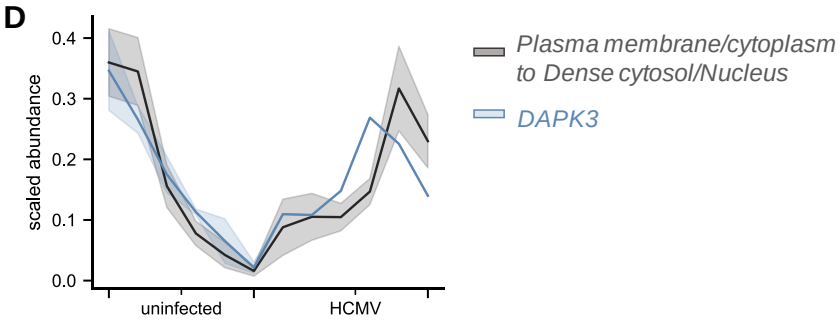


FIGURE S5

Figure S5: Investigating LDLR abundance, translocation profiles of adaptor protein complexes and DAPK3 induced upon HCMV infection. A) Selected chromatograms from targeted parallel reaction monitoring of LDLR levels during infection. A representative chromatogram is shown for each unique peptide that was monitored, and normalized abundance levels are reported in the table below. **B)** LDLR and PCSK9 mRNA levels following HCMV infection from Oberstein and Shenk, 2017. Error bars represent the standard deviation from two biological replicates. **C)** Distinct spatiotemporal translocation profiles of clathrin adaptor protein complexes AP-1, AP-2, and AP-3. Solid lines and shaded ranges represent the mean and standard deviation, respectively, of changes in abundance value for protein subunits from a given complex. **D)** Translocation profile of DAPK3 relative to synthetic translocation profiles for plasma membrane/cytoplasm to dense cytosol/nucleus movements.