Supplemental Figure Legends

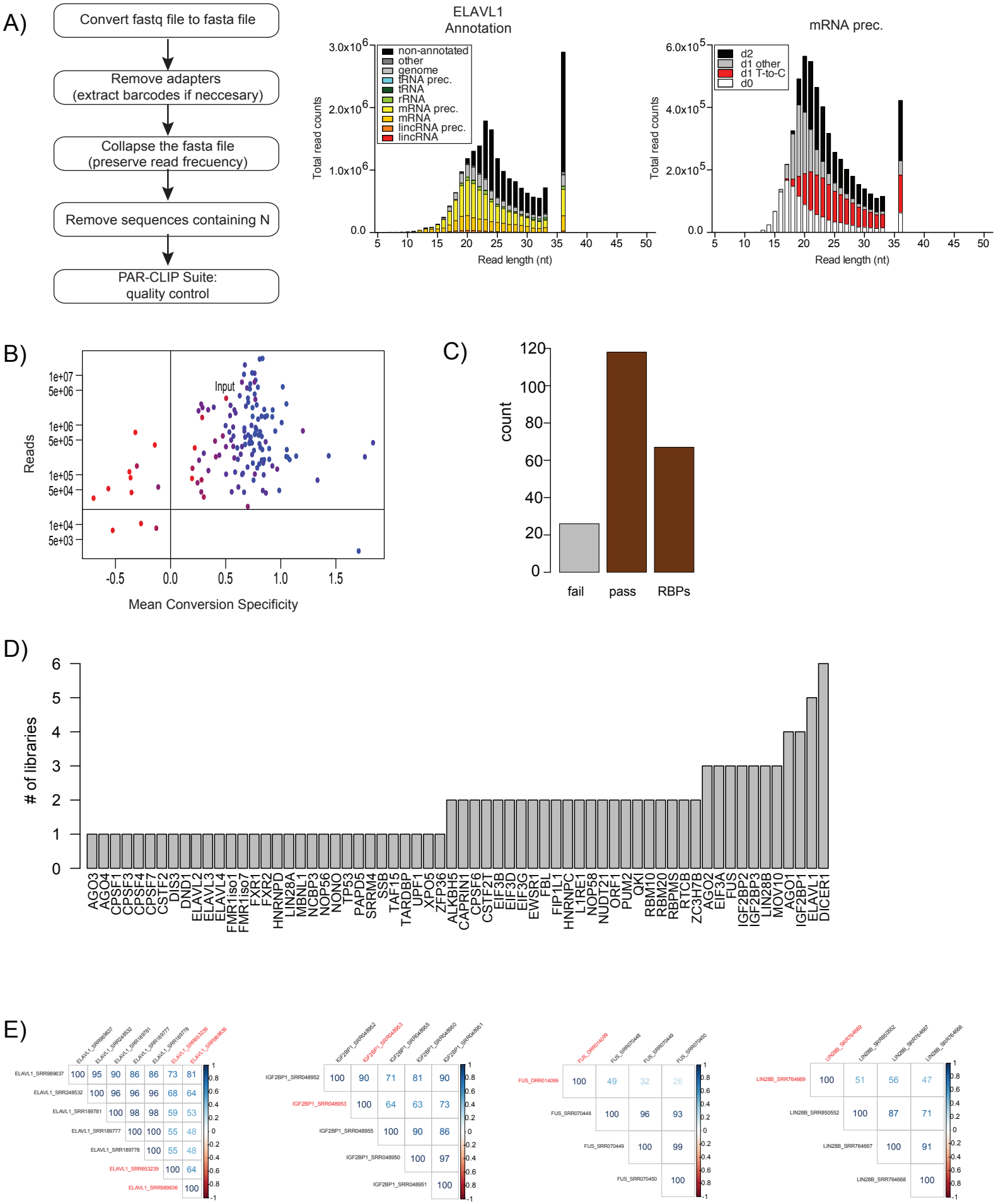
Supplemental Figure 1. QC filtering of libraries. A) Description of PAR-CLIP suite to assess library quality control per annotation category (left). Example of number of reads mapping to each RNA category with up to 2 mismatches resolved by length of adapter-extracted sequence reads for an ELAVL1 library (middle). Sequencing read composition of the most abundant RNA category fir the ELAVL1 library. Reads were assigned as d0 (white), d1 T-to-C (red), d1 other than T-to-C, (light gray), and d2 (black) (right). B) Libraries had to have > 20,000 aligned reads and a mean conversion specificity > 0, and a higher mean T-to-C fraction than the the reference library (red lower, blue higher). C) Number of libraries analyzed and their quality control status. D) Count of libraries passing QC per RBP. E) Examples of outlier library removal (libraries labeled with red text were removed) based on correlation of read 6-mer frequency for RBPs with 3 or more libraries.

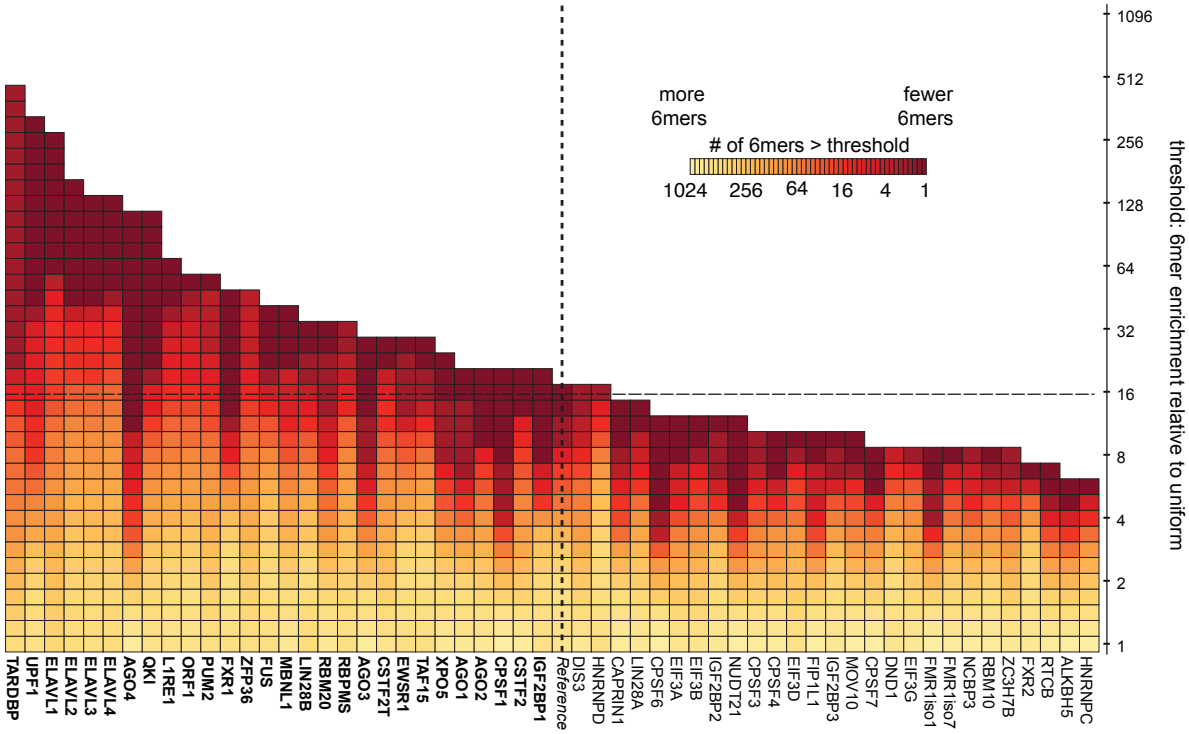
Supplemental Figure 2. Grouping RBPs by sequence specificity. Heatmap of the number of 6-mers enriched per RBP at different specificity thresholds. The color scale represents the $\log_2$ [number of 6-mers] that are enriched at a given threshold (y-axis). The thresholds are represented as $\log_2$ [6-mer frequency]. There are 4096 different 6-mers and if they were uniformly present this would represent a value of -12 =$\log_2$ [1/4096]. The horizontal dashed lines at -8, represents 16-fold enrichment over a uniform background. For reference, the vertical dashed lines indicate the behavior of the reference library.

Supplemental Figure 3. Factor analysis model selection and performance. A) Plot of eigenvalues versus number of factors to determine the optimal number of factors using four methods (different colors). B) Heatmap of the median factor score coefficient value for all genes

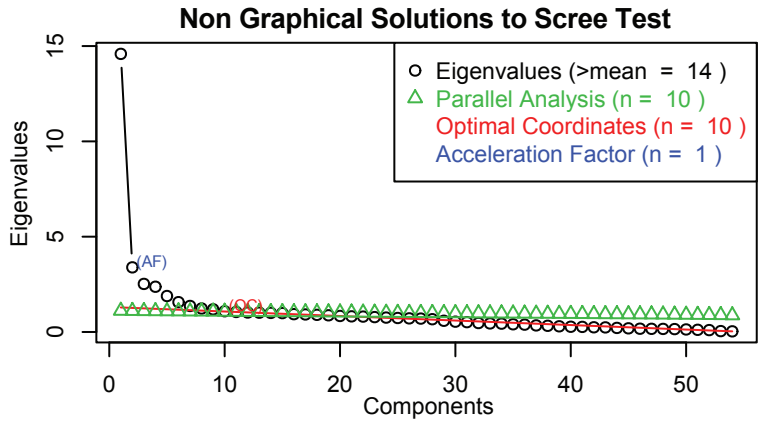
that clustered together. The number of genes assigned to a specific factor and the top two most significant enriched GO annotations for each ontology class: molecular function (MF), cellular component (CC), and biological process (BP).

Supplemental Figure 4. RNA metabolism profiles for factor-associated gene sets. A) Box-and-whisker plot for each gene set of the synthesis rates, processing rates, degradation rates, cytoplasmic versus nuclear localization (Cyt vs Nuc), polyribosomal versus cytoplasmic localization (Poly vs Cyt), and translational status from ribosome profiling data. B) Heatmap of the odds-ratio of the overlap between factor associated gene sets with RNA categories based on similar metabolic profiles from (18). C) Heatmap of the odds-ratio of the overlap between factor associated gene sets and protein localization annotation. Box-and-whisker plot for each gene set of the D) coefficient of variation across and E) median expression across 25 HEK293 cells.

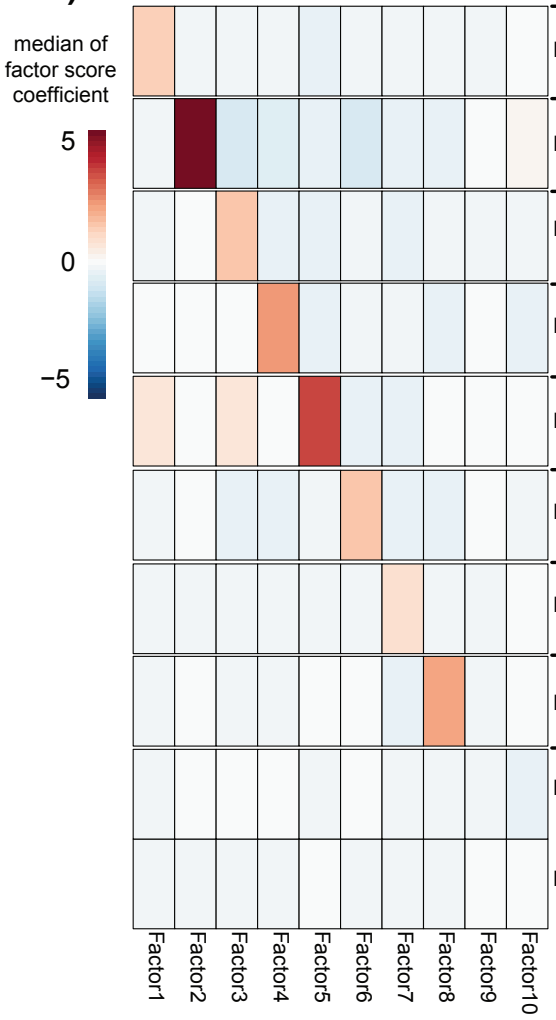

A)
Convert fastq file to fasta file
Remove adapters
(extract barcodes if neccesary)
Collapse the fasta file
(preserve read frecuency)
Remove sequences containing N
PAR-CLIP Suite:
quality control
ELAVL1
Annotation
non-annotated
other
genome
tRNA prec.
tRNA
rRNA
mRNA prec.
mRNA
lincRNA prec.
lincRNA
Total read counts
Read length (nt)
mRNA prec.
d2
d1 other
d1 T-to-C
d0
Total read counts
Read length (nt)
B)
Input
Reads
Mean Conversion Specificity
C)
count
fail
pass
RBPs
D)
# of libraries
AGO3
AGO4
CPSF1
CPSF3
CPSF4
CPSF7
CSTF2
DIS3
DND1
ELAVL2
ELAVL3
ELAVL4
FMR1iso1
FMR1iso7
FXR1
FXR2
HNRNPD
LIN28A
MBNL1
NCBP3
NOP56
NONO
TP53
PAPD5
SRRM4
SSB
TAF15
TARDBP
UPF1
XPO5
ZFP36
ALKBH5
CAPRIN1
CPSF6
CSTF2T
EIF3B
EIF3D
EIF3G
EWSR1
FBL
FIP1L1
HNRNPC
L1RE1
NOP58
NUDT21
ORF1
PUM2
QKI
RBM10
RBM20
RBPMS
RTCB
ZC3H7B
AGO2
EIF3A
FUS
IGF2BP2
IGF2BP3
LIN28B
MOV10
AGO1
IGF2BP1
ELAVL1
DICER1
E)


more
6mers
fewer
6mers
# of 6mers > threshold
1024 256 64 16 4 1
1096
512
256
128
64
32
16
8
4
2
1
threshold: 6mer enrichment relative to uniform
TARDBP
UPF1
ELAVL1
ELAVL2
ELAVL3
ELAVL4
AGO4
QKI
L1RE1
ORF1
PUM2
FXR1
ZFP36
FUS
MBNL1
LIN28B
RBM20
RBPMS
AGO3
CSTF2T
EWSR1
TAF15
XPO5
AGO1
AGO2
CPSF1
CSTF2
IGF2BP1
Reference
DIS3
HNRNPD
CAPRIN1
LIN28A
CPSF6
EIF3A
EIF3B
IGF2BP2
NUDT21
CPSF3
CPSF4
EIF3D
FIP1L1
IGF2BP3
MOV10
CPSF7
DND1
EIF3G
FMR1iso1
FMR1iso7
NCBP3
RBM10
ZC3H7B
FXR2
RTCB
ALKBH5
HNRNPC

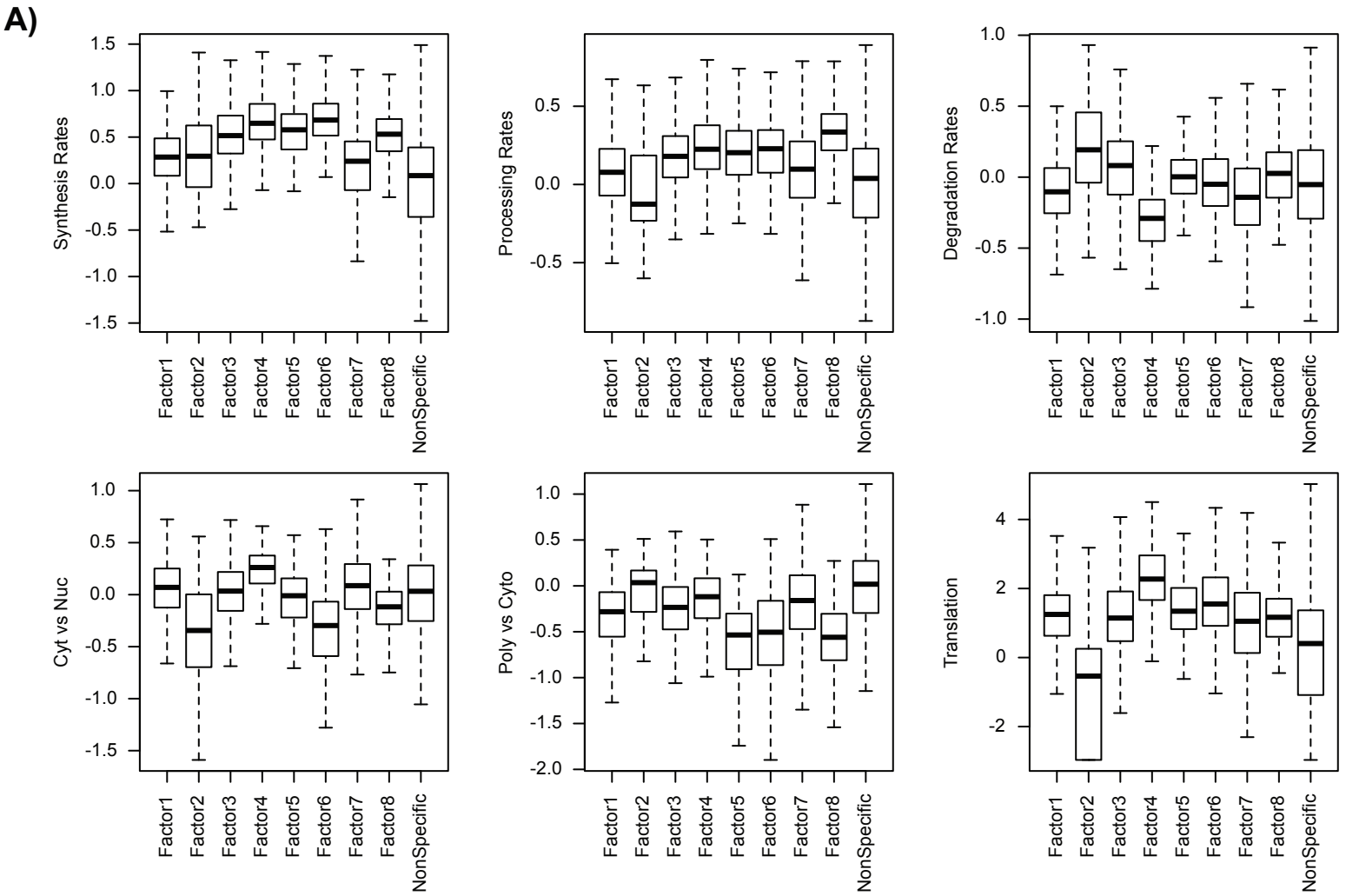
A)

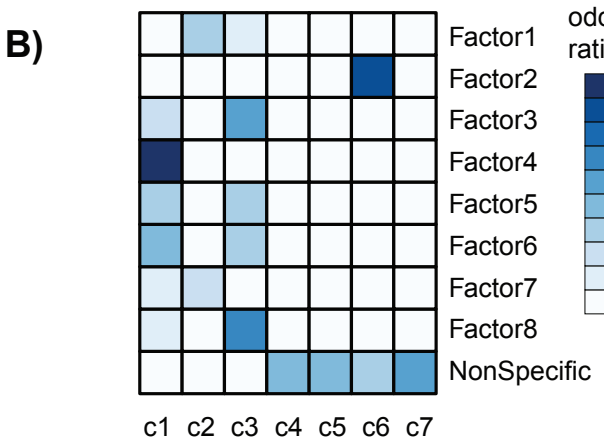


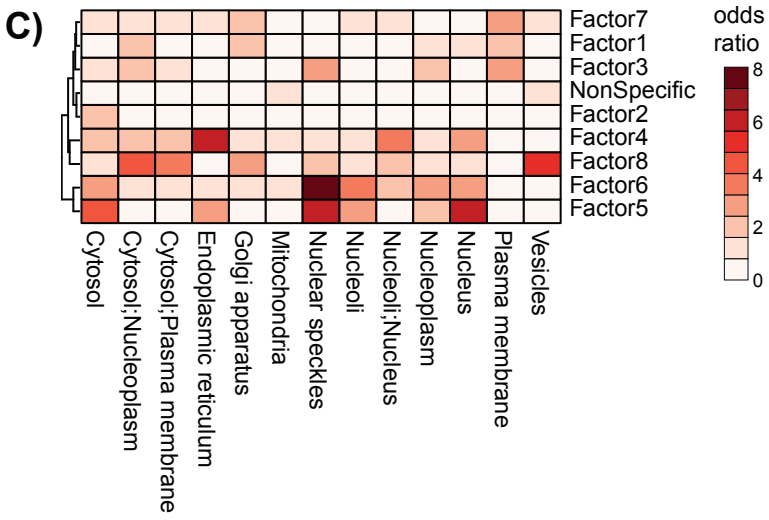
B)

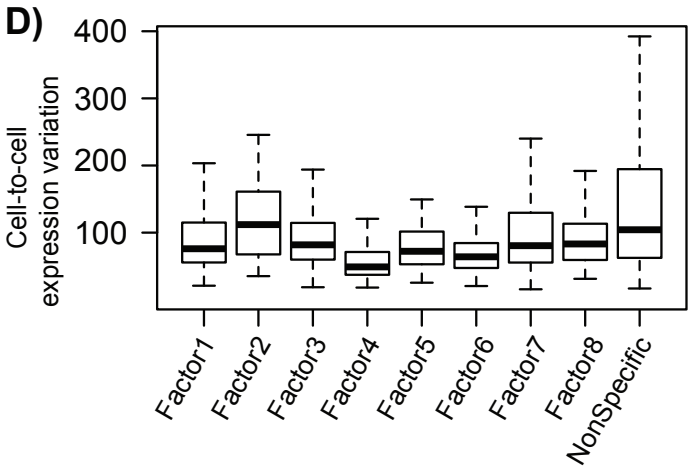


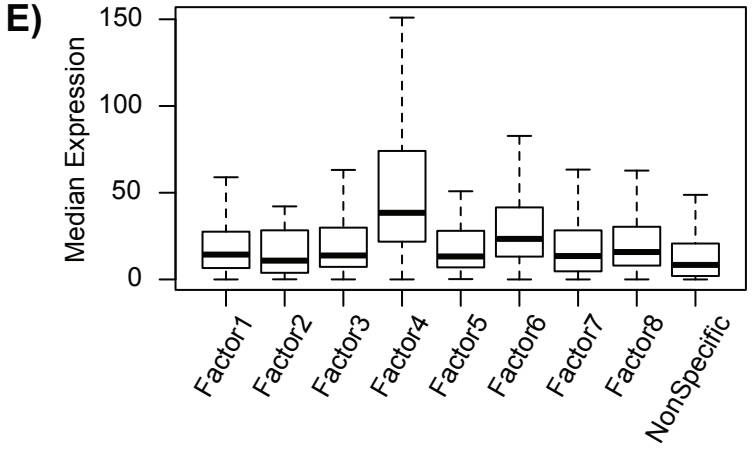
genes assigned to:	Top 2 significant GO enrichments for each Ontology class
F1, n=604	MF: cAMP response element binding, AU-rich element binding CC: endosomal membrane, ubiquitin ligase complex BP: neg. reg. of cellular response, neg. reg. of transmembrane receptor
F2, n=34	long-noncoding RNAs: NEAT1, JPX, MALAT1, PVT primary miRNAs: MIR99AHG, MIR181A1HG
F3, n=408	MF: ubiquitin-like transferase activity, SMAD binding CC: ubiquitin ligase complex, cytoplasmic ribonucleoprotein granule BP: neg. reg. of translation - ncRNA, gene silencing by miRNA
F4, n=213	MF: GTP binding, guanyl ribonucleotide binding, phosphoprotein binding CC: pigment granule, cell-substrate junction, nuclear envelope BP: reg. of establishment of protein localization, post-transcriptional reg. of gene expression
F5, n=100	MF: CC: BP: regulation of intracellular transport
F6, n=376	MF: structural constituent of nuclear pore, nucleocytoplasmic transporter activity CC: microtubule organizing center part, spindle pole BP: establishment of RNA localization, chromosome segregation
F7, n=1200	MF: phosphatidic acid binding, protein kinase A binding CC: coated vesicle membrane, endocytic vesicle BP: ER to Golgi vesicle-mediated transport, protein localization to endosome
F8, n=245	MF: transcription factor binding, helicase activity CC: plasma membrane region, cell leading edge BP: histone modification, peptidyl-lysine modification
NS, n=2403	MF: transmembrane transporter activity, endopeptidase activity CC: mitochondrial matrix, ribosomal subunit BP: translational termination, mitochondrial respiratory chain complex
NS, n=7716	


A)
Synthesis Rates
Processing Rates
Degradation Rates
Cyt vs Nuc
Poly vs Cyto
Translation
Factor1
Factor2
Factor3
Factor4
Factor5
Factor6
Factor7
Factor8
NonSpecific


B)
odds ratio
Factor1
Factor2
Factor3
Factor4
Factor5
Factor6
Factor7
Factor8
NonSpecific
c1 c2 c3 c4 c5 c6 c7


C)
odds ratio
Factor7
Factor1
Factor3
NonSpecific
Factor2
Factor4
Factor8
Factor6
Factor5
Cytosol
Cytosol;Nucleoplasm
Cytosol;Plasma membrane
Endoplasmic reticulum
Golgi apparatus
Mitochondria
Nuclear speckles
Nucleoli
Nucleoli;Nucleus
Nucleoplasm
Nucleus
Plasma membrane
Vesicles


D)
Cell-to-cell expression variation
Factor1
Factor2
Factor3
Factor4
Factor5
Factor6
Factor7
Factor8
NonSpecific


E)
Median Expression
Factor1
Factor2
Factor3
Factor4
Factor5
Factor6
Factor7
Factor8
NonSpecific