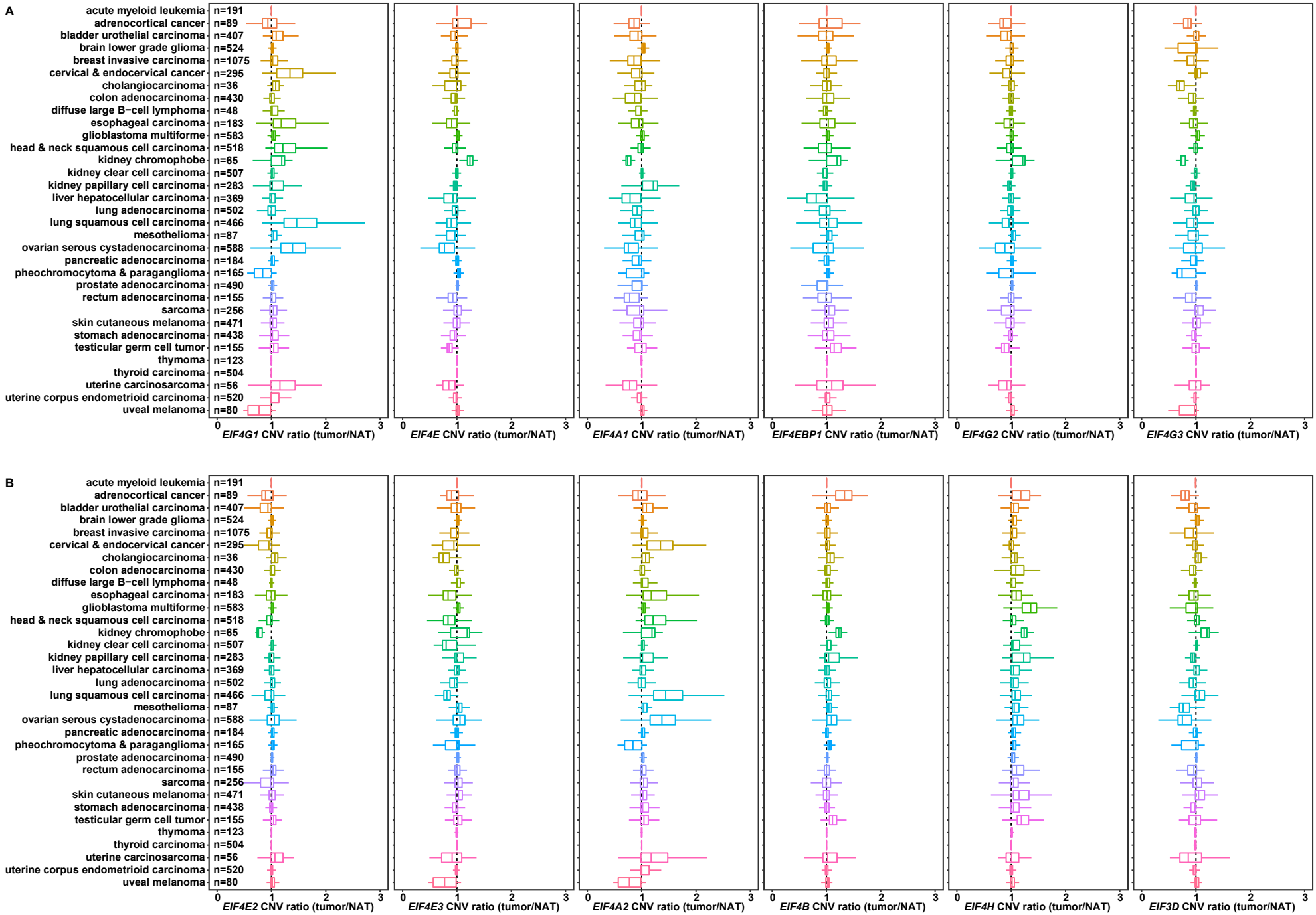


Figure S1. Copy Number Ratios of *EIF4F* Genes Between Malignant Tumors and Normal Adjacent Tissues, Related to Figure 1

(**A** and **B**) The box plots show, across 33 TCGA cancer types, ratios of tumor and normal adjacent tissue CNV values for *EIF4G1*, *EIF4E*, *EIF4A1*, *EIF4EBP1*, *EIF4G2*, *EIF4G3*, *EIF4E2*, *EIF4E3*, *EIF4A2*, *EIF4B*, *EIF4H*, and *EIF3D*. Gene-level CNV values were estimated by the GISTIC2 method using the whole genome microarray data from TCGA. CNV ratios were calculated by dividing, for each gene shown, the estimated gene-level CNV values in malignant tumors to the average CNV value in normal adjacent tissues (NATs) of the same cancer type.

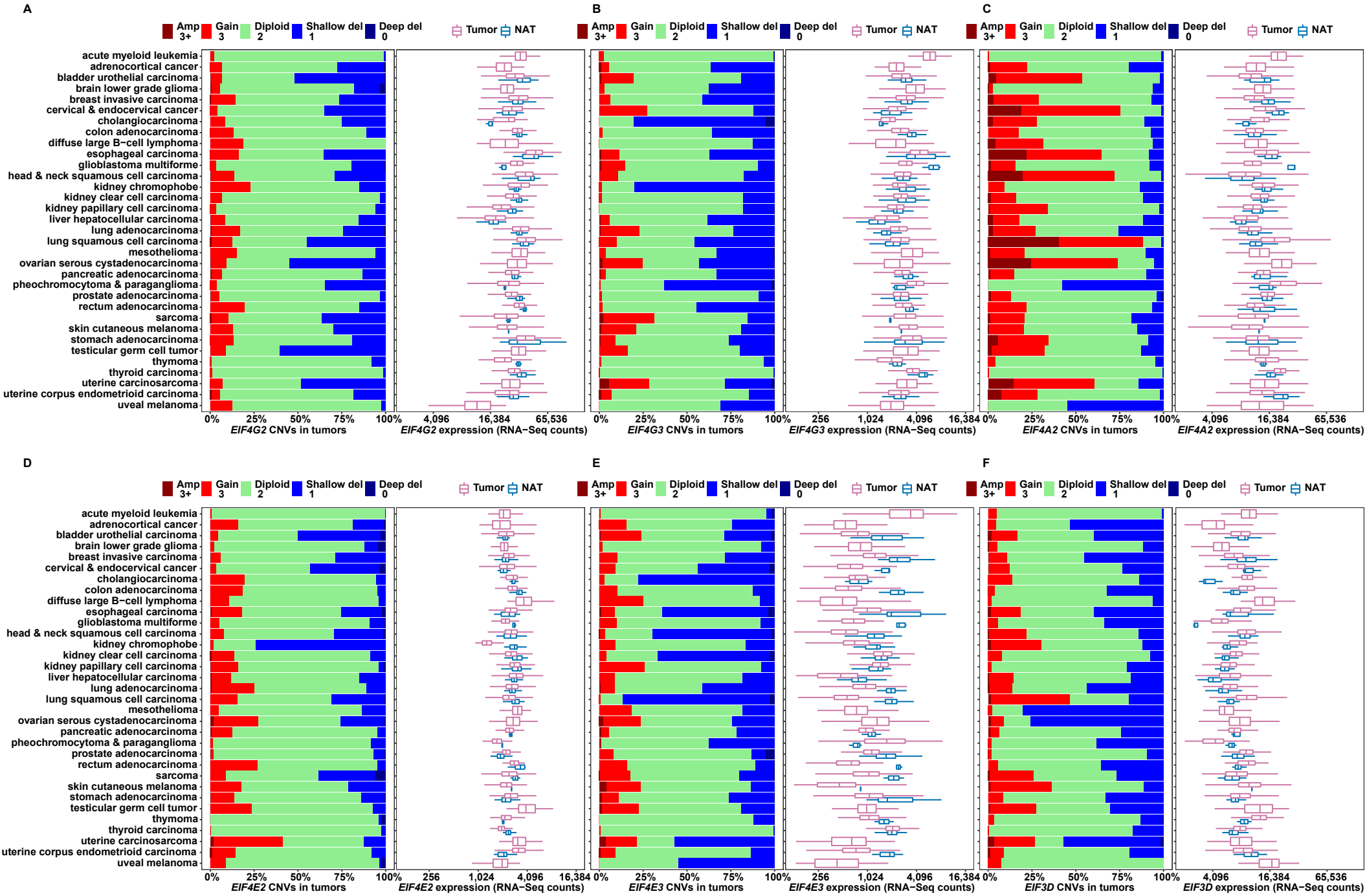


Figure S2. Copy Number Variations and RNA Expression of *EIF4F* Genes in Tumors, Related to Figure 1

(**A** to **F**) For each cancer study group, a stacked bar plot shows CNV status for a translation initiation gene (marked at the bottom of each plot), and a box and whisker plot shows corresponding mRNA expression of the same gene in tumor samples and Normal Adjacent Tissues (NATs). The X axes of box plots represent normalized gene level expression (transcripts per million) in $\log_2$ scale. TCGA uses the same bioinformatics pipeline to process and normalize RNA-Seq data from different cancer study groups, to minimize batch effects of sequencing data processing.

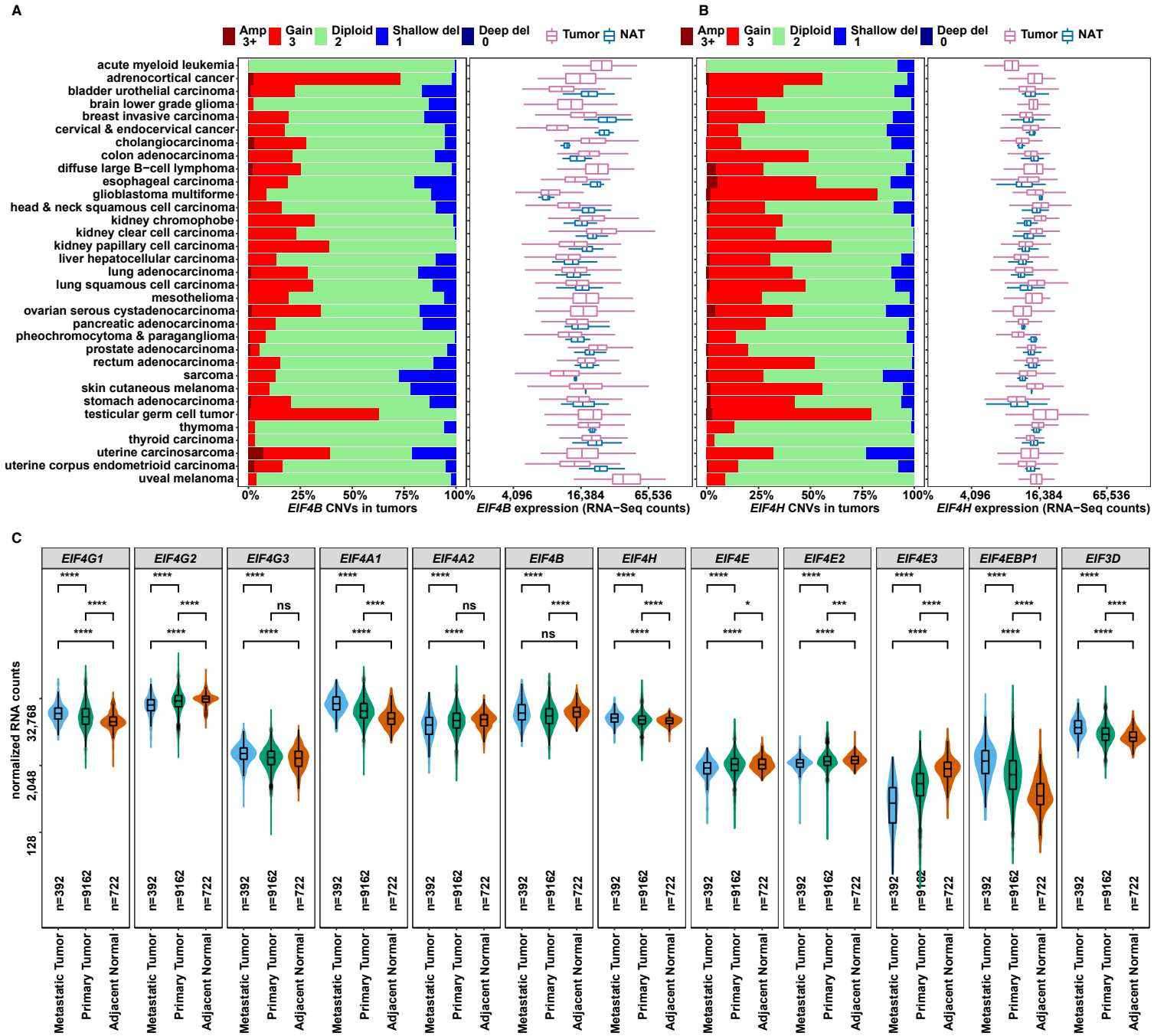


Figure S3. RNA Expression of *EIF4F* Genes in Malignant Tumors and Normal Adjacent Tissues, Related to Figure 1

(**A** and **B**) For each cancer study group, a stacked bar plot shows CNV status for a translation initiation gene (marked at the bottom of each plot), and a box and whisker plot shows corresponding mRNA expression of the same gene in tumor samples and Normal Adjacent Tissues (NATs). The X axes of box plots represent normalized gene level expression (transcripts per million) in $\log_2$ scale. TCGA uses the same bioinformatics pipeline to process and normalize RNA-Seq data from different cancer study groups, to minimize batch effects of sequencing data processing.

(**C**) The violin plots show the normalized mRNA expression (transcripts per million), in all TCGA cancer patient samples combined, of *EIF4G1*, *EIF4G2*, *EIF4G3*, *EIF4A1*, *EIF4A2*, *EIF4B*, *EIF4H*, *EIF4E*, *EIF4E2*, *EIF4E3*, *EIF4EBP1*, and *EIF3D*. Gene expression was compared among all available samples: 392 metastatic tumors, 9162 primary tumors and 722 NATs.

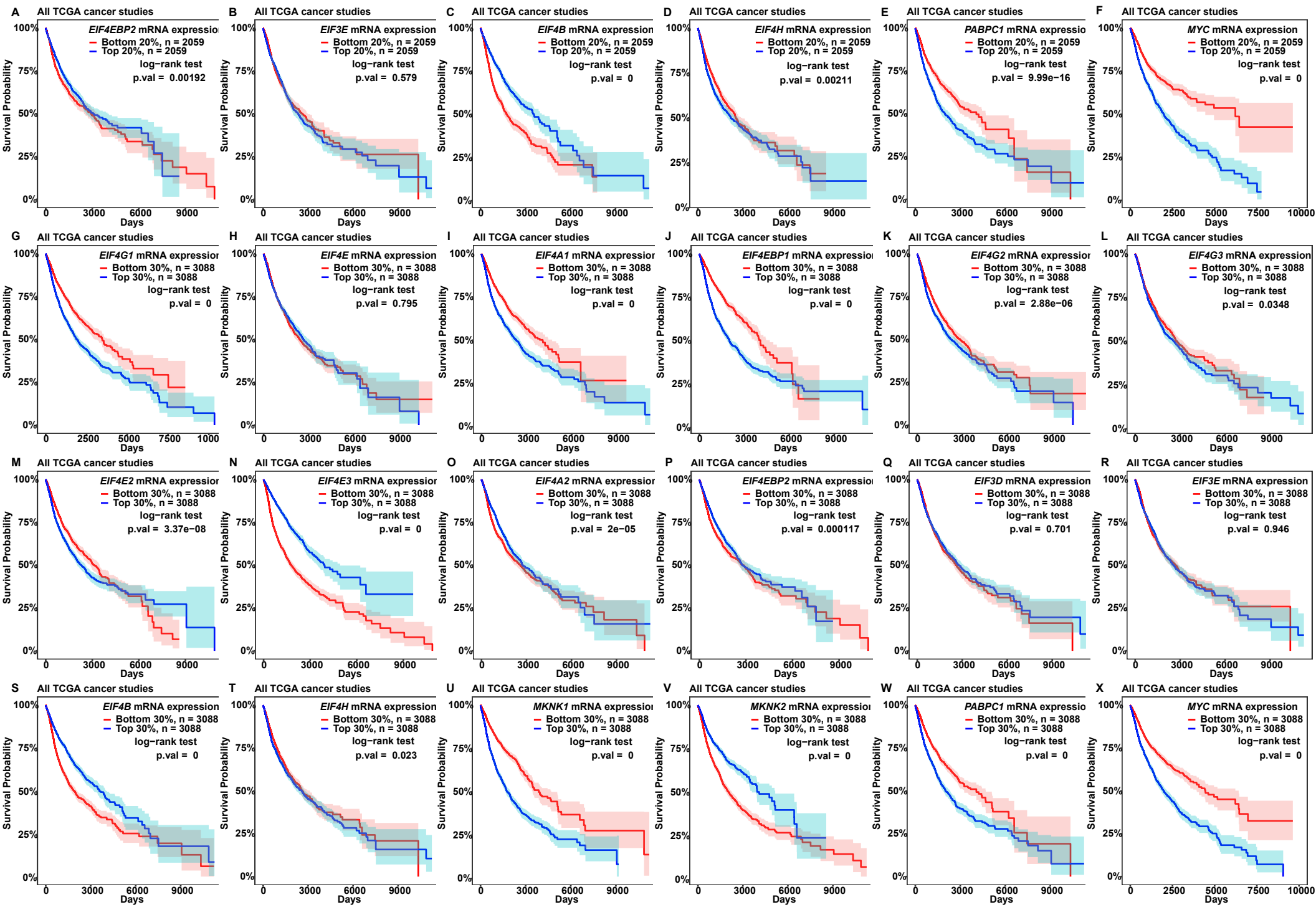
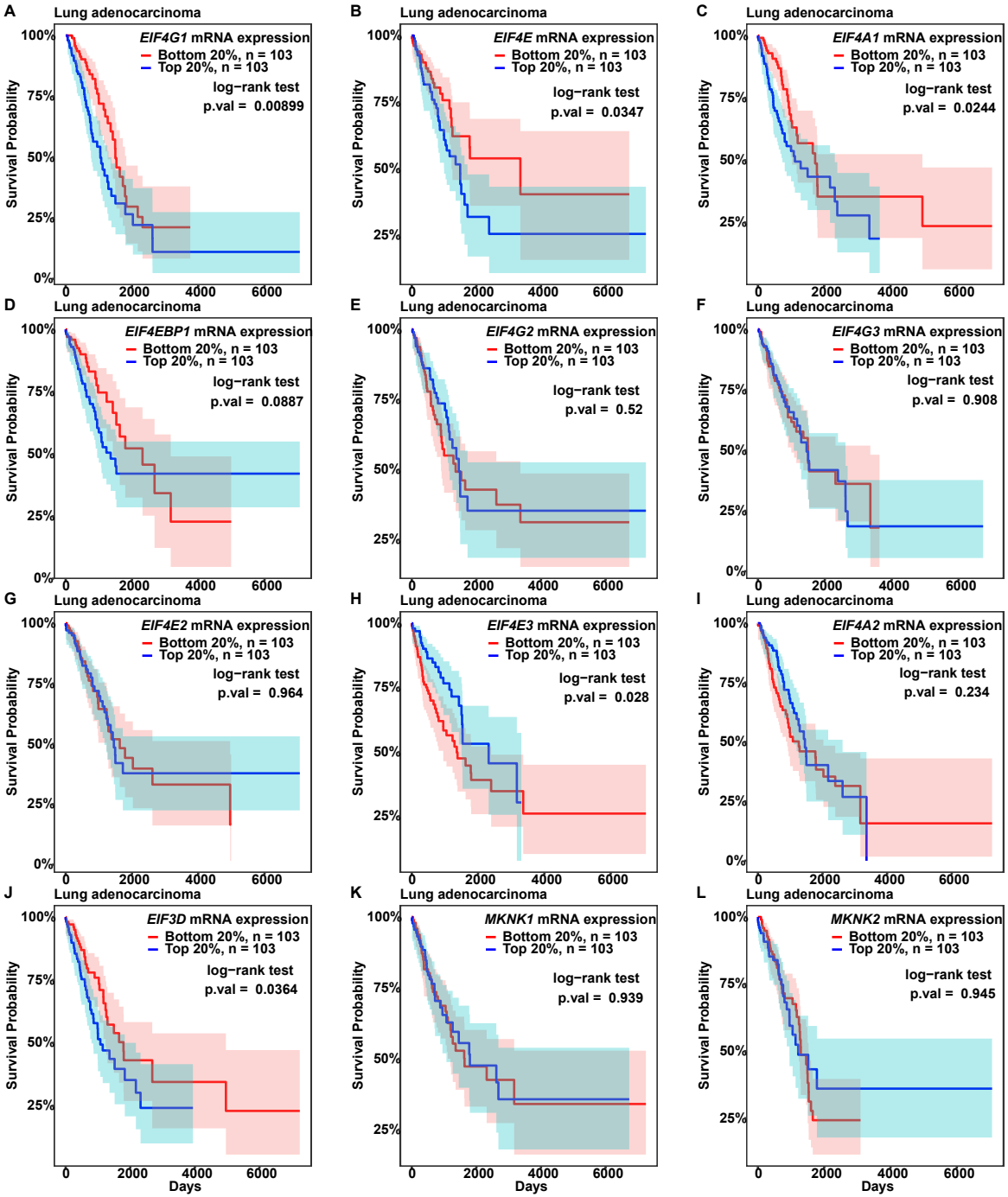
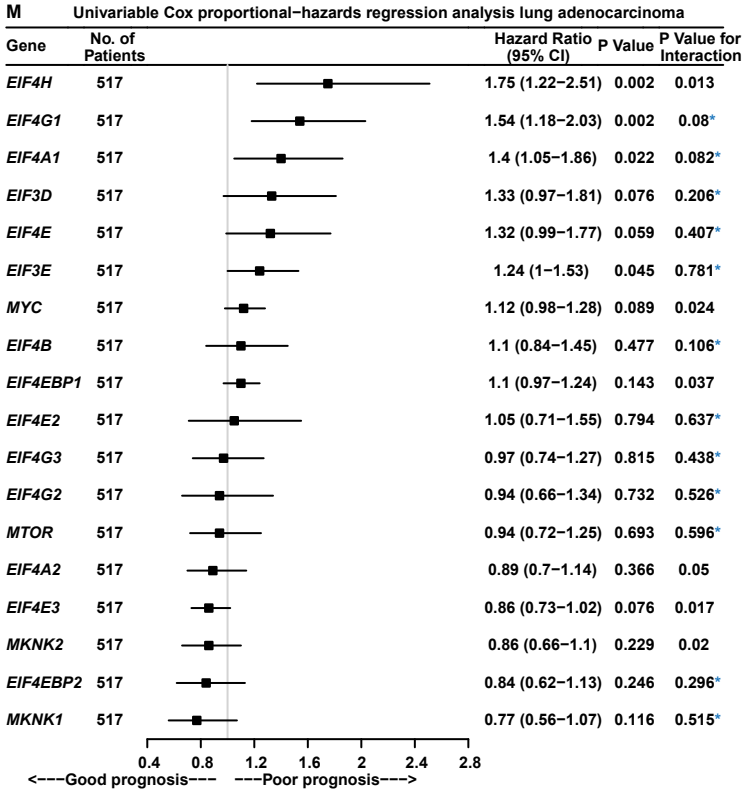


Figure S4. Survival Analyses of *EIF4F* Expression in Patients from All Cancer Types, Related to Figure 2

(**A** to **F**) Each Kaplan-Meier plot shows survival probabilities of TCGA patients with cancer according to mRNA expressions of a specific translation initiation gene (marked inside each box, at top) in their tumors. Survival probability (Y axis) is the probability of individual survival from the time origin (e.g. initial cancer diagnosis) to a specified time (X axis). We ranked all 10,295 TCGA patients with cancer based on the indicated individual gene expressions from their tumor biopsies, and selected two groups of patients with the top or bottom 20% of gene expression. Differences in survival probabilities between the two selected groups were assessed with the log-rank test. The shaded areas around each curve depict a 95% confidence region for that curve.

(**G** to **X**) Each Kaplan-Meier plot shows survival probabilities of TCGA patients with cancer according to mRNA expressions of a specific translation initiation gene (marked inside each box, at top) in their tumors. We ranked all 10,295 TCGA patients with cancer based on the indicated individual gene expressions from their tumor biopsies, and selected two groups of patients with the top or bottom 30% of gene expression. Differences in survival probabilities between the two selected groups were assessed with the log-rank test. The shaded areas around each curve depict a 95% confidence region for that curve.





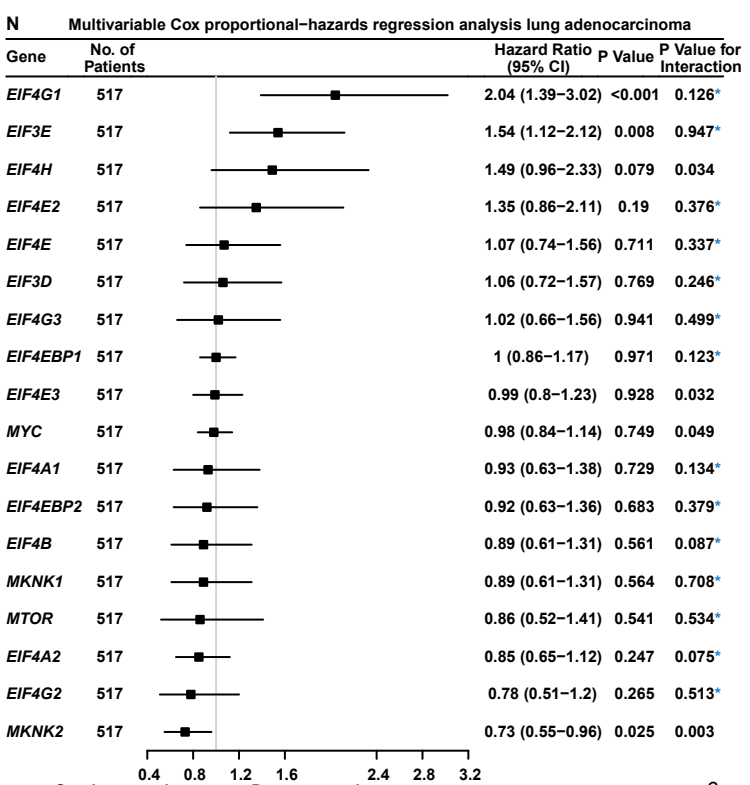


Figure S5. Survival Analyses of *EIF4F* Expression in Patients with Lung Adenocarcinoma, Related to Figure 2

(**A** to **L**) Each Kaplan-Meier plot shows survival probabilities of patients with lung adenocarcinoma according to mRNA expressions of a specific translation initiation gene (marked inside each box, at top) in their tumors. We ranked 517 TCGA patients with lung adenocarcinoma based on the indicated individual gene expressions from their tumor biopsies, and selected two groups of patients with the top or bottom 20% of gene expression. Differences in survival probabilities between the two selected groups were assessed with the log-rank test. The shaded areas around each curve depict a 95% confidence region for that curve.

(**M** and **N**) Univariable (**M**) and multivariable (**N**) Cox proportional-hazards regression models for expression of translation initiation genes in 517 patients with lung adenocarcinoma from TCGA. P value indicates the statistical significance of association between gene expression and survival (i.e. a significant fit in the Cox-PH model). “P value for interaction” is calculated using the Schoenfeld residual method; a value less than 0.05 indicates statistically significant interaction between gene expression and time – a violation of the Cox proportional-hazards model. Values without violation are marked with a blue asterisk. “CI” means “confidence interval”.

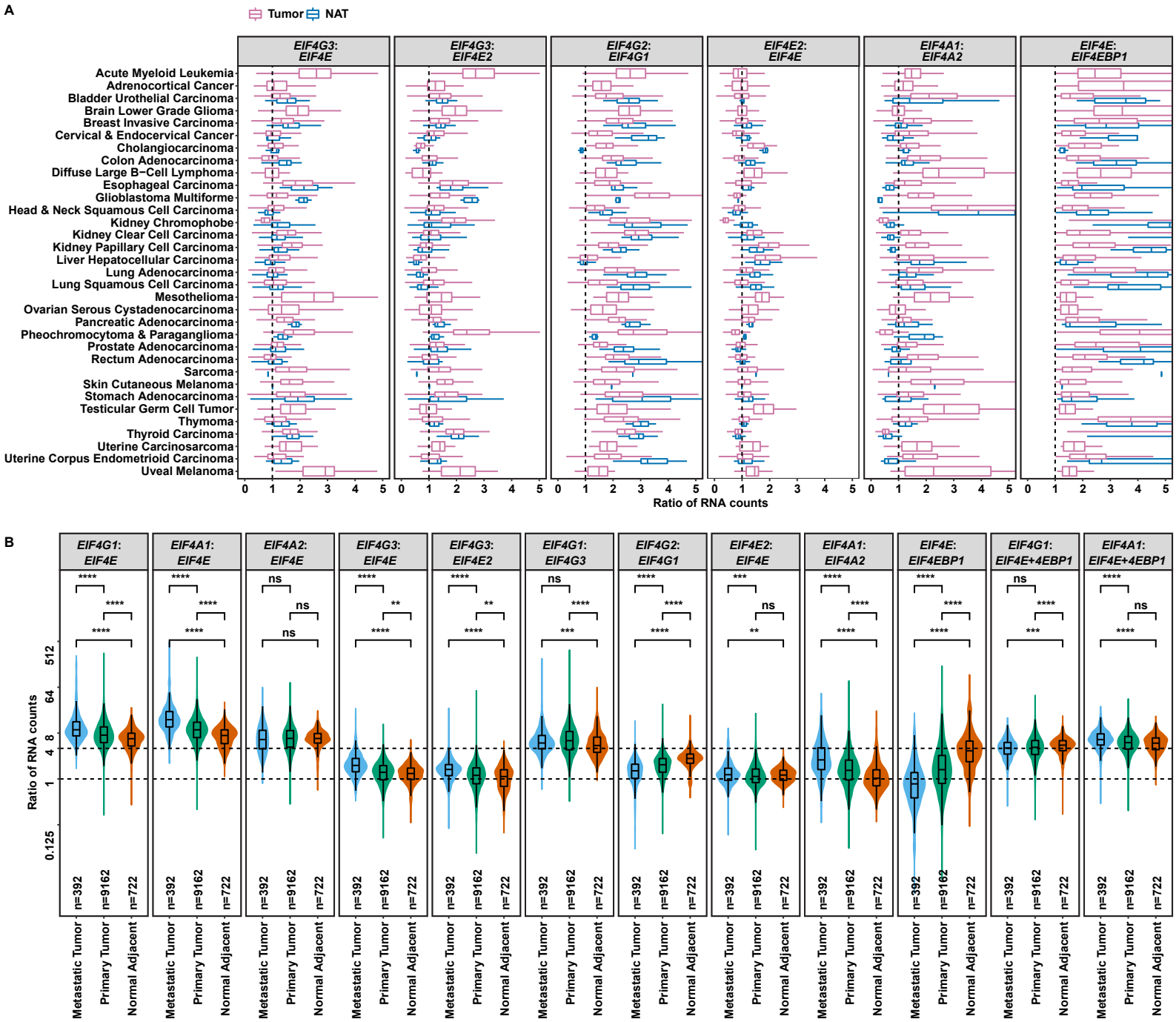


Figure S6. Relative Expressions of *EIF4F* Genes in Malignant Tumors and NATs, Related to Figure 3

(**A**) Gene expression ratios for tumors and NATs in different cancer types. Each box plot depicts ratios of RNA counts (X-axis, linear scale) between two genes (marked at top), computed from each sample of a particular cancer type (left), for tumors and NATs considered separately.

(**B**) The violin plots compare RNA ratios (indicated at top of each plot). Ratios were calculated within each sample type (metastatic tumor, primary tumor, or NAT) using normalized gene level expression data across all 33 TCGA cancer study groups combined. All available samples were considered: 392 metastatic tumors, 9162 primary tumors and 722 NATs. Dashed lines mark the ratios of 1:1 and 4:1 in all panels.

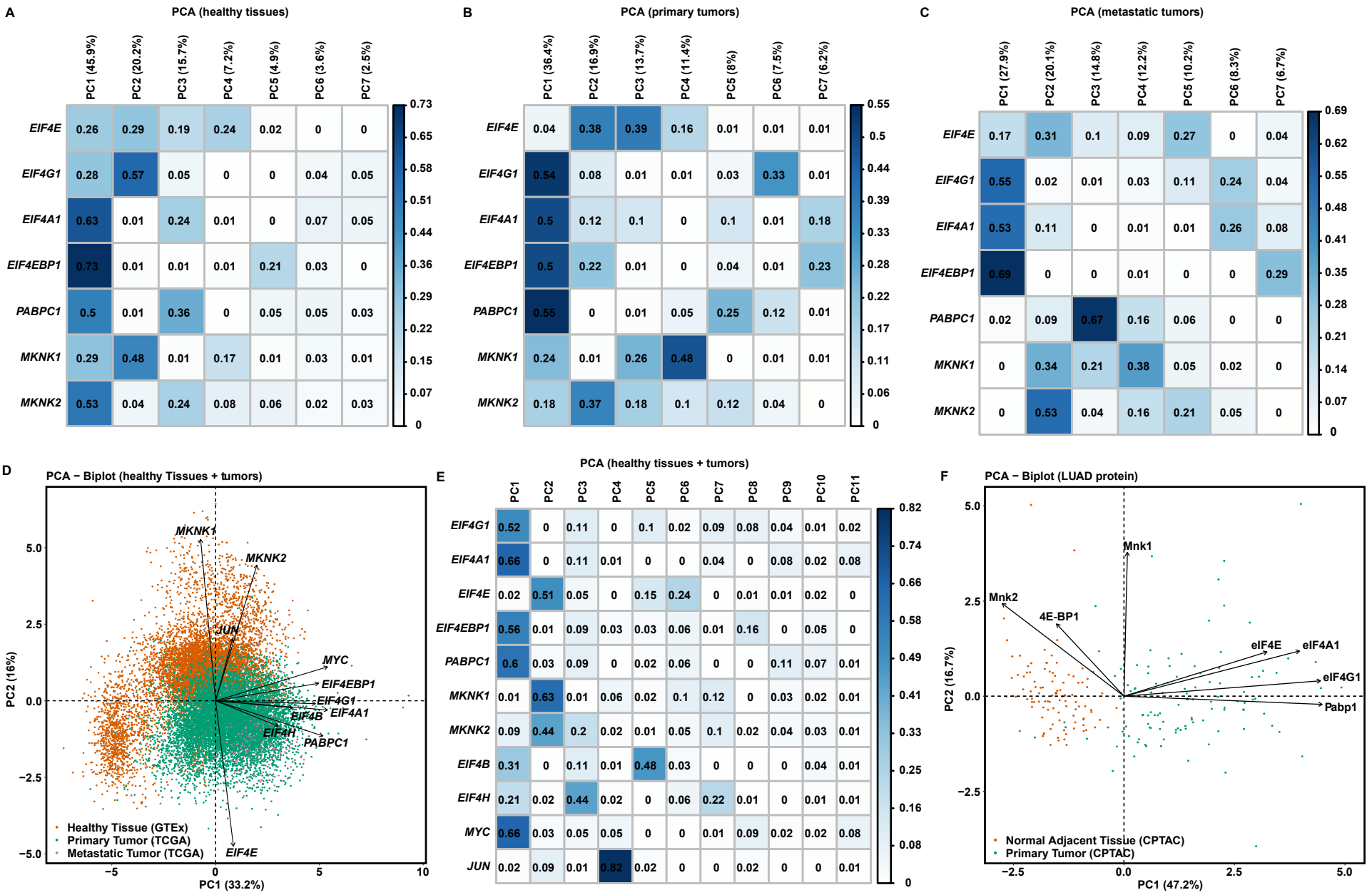


Figure S7. Malignant Tumor Types are Less Distinguished from Each Other by *EIF4F* Expressions, than Healthy Tissue Types, Related to Figure 4

(**A**) The matrix plot shows the cos2 value for the contribution of each gene to each PC, from the PCA of 7,388 tissue samples of various healthy tissue types in **Figure 4A**. The sum of values in a given row across all PCs is equal to one (+/- epsilon introduced by rounding each value to two significant figures).

(**B**) The matrix plot shows the cos2 value for the contribution of each gene to each PC, from the PCA of 9,162 primary tumors of 33 cancer types in **Figure 4B**. The sum of values in a given row across all PCs is equal to one (+/- epsilon introduced by rounding each value to two significant figures).

(**C**) The matrix plot shows the cos2 value for the contribution of each gene to each PC, from the PCA of 392 metastatic tumors of 33 cancer types in **Figure 4C**. The sum of values in a given row across all PCs is equal to one (+/- epsilon introduced by rounding each value to two significant figures).

(**D**) PCA of normalized RNA-Seq-derived counts of indicated genes from 9,162 primary and 392 metastatic tumors from TCGA, and 7,388 healthy tissue samples from GTEx. RNA-Seq data from TCGA and GTEx have been processed and normalized with the same bioinformatics pipeline to minimize batch effects of sequencing experiments. Sample types were colored after analysis, for visualization.

(**E**) The matrix plot shows the cos2 value for the contribution of each gene to each PC, from the PCA of TCGA tumor samples and GTEx normal tissue samples in **Figure S7D**. The sum of values in a given row across all PCs is equal to one (+/- epsilon introduced by rounding each value to two significant figures).

(**F**) PCA was performed for 109 lung adenocarcinomas and 102 paired normal lung tissues from the CPTAC LUAD study on the standardized protein expressions of eIF4G1, eIF4A1, eIF4E, 4E-BP1, Pabp1, Mnk1, and Mnk2.

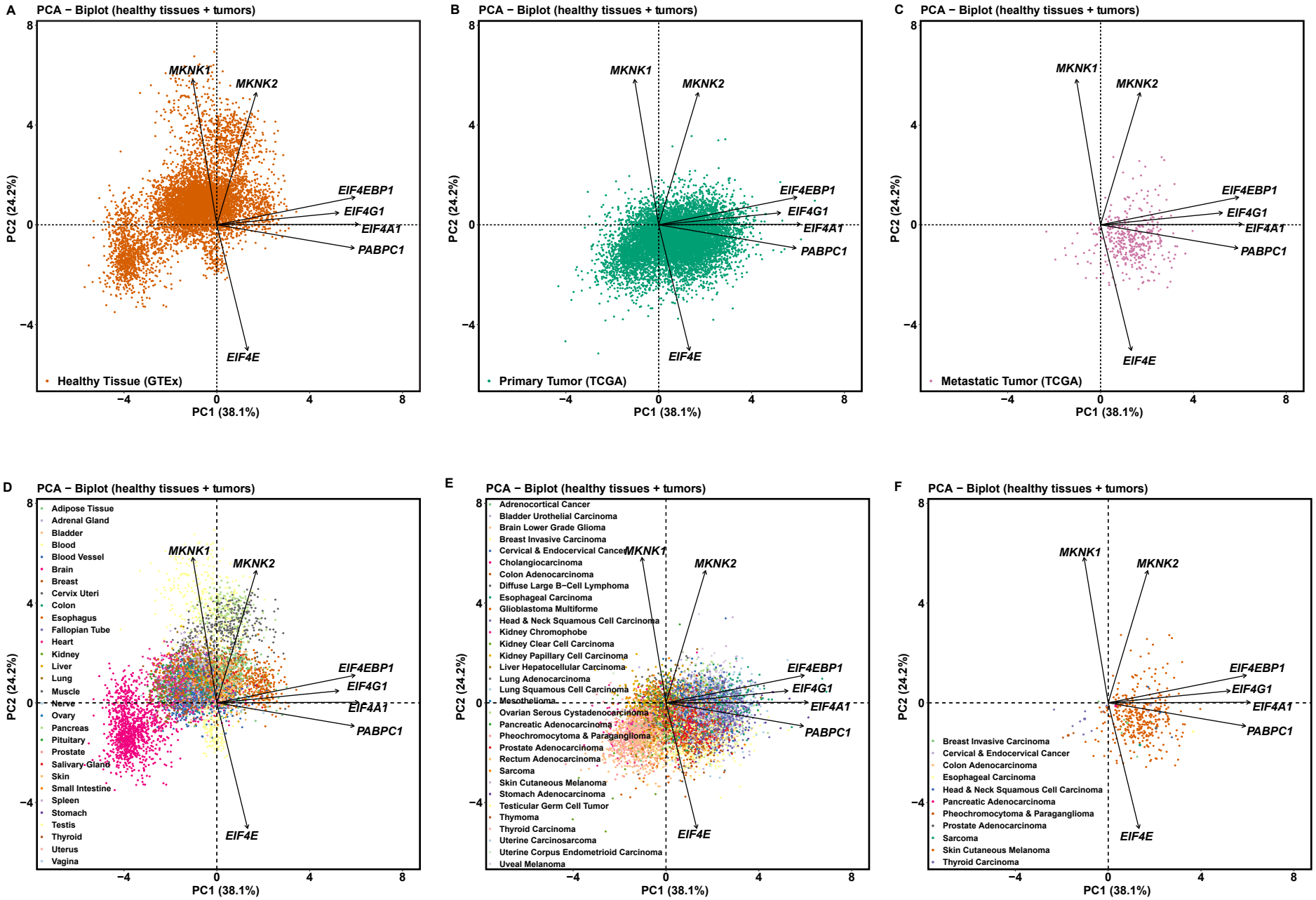


Figure S8. *EIF4F* Expressions Generally Distinguish Tumors from Healthy Tissues, Related to Figure 4

(**A** and **D**) Healthy tissue selection from **Figure 4D**. The tissue types in (**D**) were not used as variables to construct PCs, but were colored afterwards for visualization.

(**B** and **E**) Primary tumor selection from **Figure 4D**. The cancer types in (**E**) were not used as variables to construct PCs, but were colored afterwards for visualization.

(**C** and **F**) Metastatic tumor selection from **Figure 4D**. The cancer types in (**F**) not used as variables to construct PCs, but were colored afterwards for visualization.

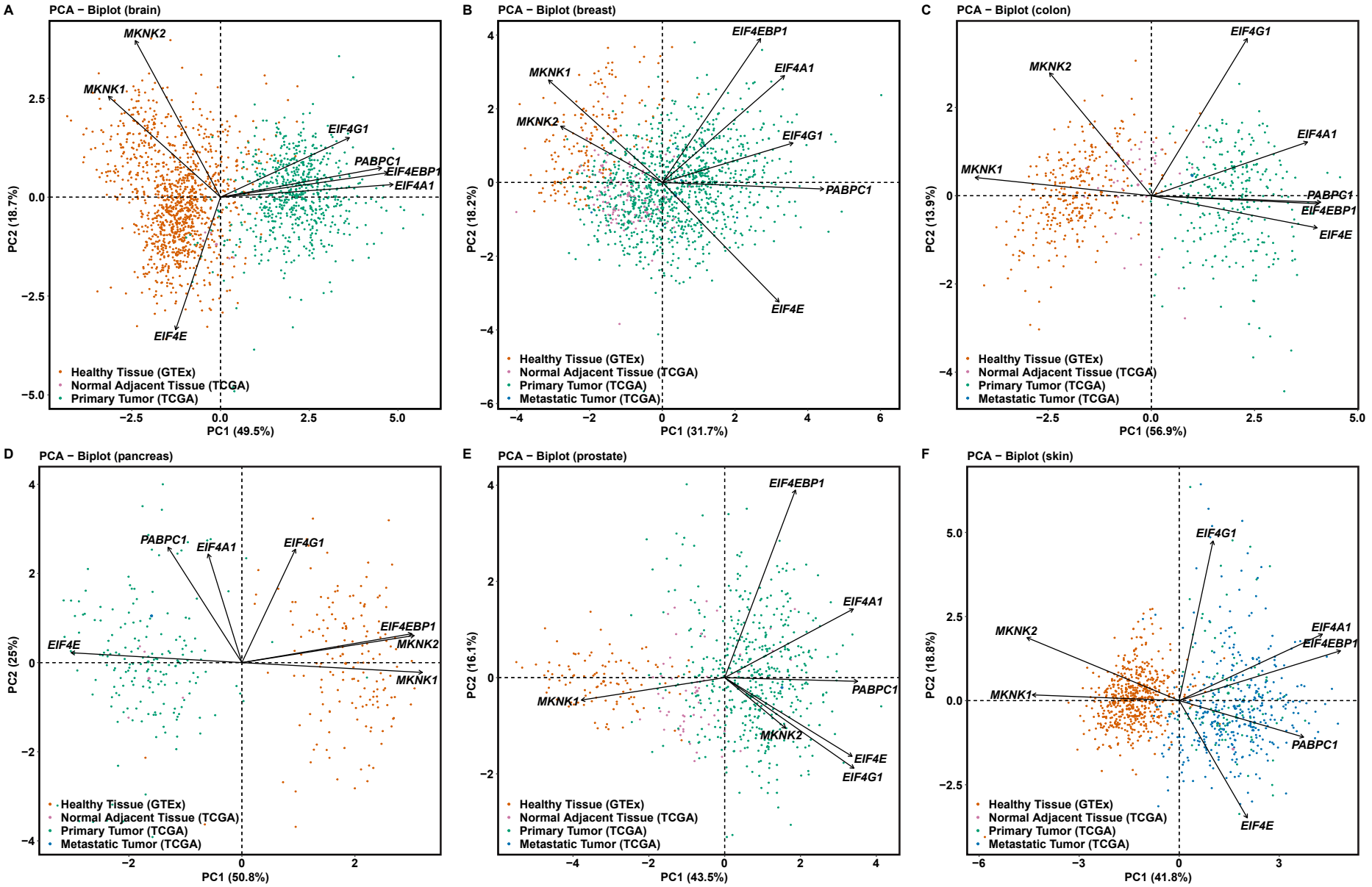


Figure S9. *EIF4F* Expressions Distinguish Tumors from Healthy Tissues in Individual Cancer Types, Related to Figure 4

(**A**) PCA of normalized RNA-Seq-derived counts of indicated genes from 662 primary brain tumor samples (509 lower grade gliomas and 158 glioblastoma multiformes) and 5 normal adjacent brain tissues from TCGA, and 1,136 healthy brain tissues (including cerebellum, caudate, cortex, nucleus accumbens, cortex, hippocampus and hypothalamus) from GTEx. RNA-Seq data from TCGA and GTEx have been processed and normalized with the same bioinformatics pipeline to minimize batch effects of sequencing experiments. Sample types were colored after analysis, for visualization.

(**B**) PCA of normalized RNA-Seq-derived counts of indicated genes from 1,092 primary breast tumors, 7 metastatic breast tumors and 113 normal adjacent breast tissues from TCGA breast invasive carcinoma group, and 179 healthy breast mammary tissues from GTEx.

(**C**) PCA of normalized RNA-Seq-derived counts of indicated genes from 288 primary colon tumors, 1 metastatic colon tumor and 41 normal adjacent colon tissues from TCGA colon adenocarcinoma group, and 304 healthy colon tissues (including transverse and sigmoid colon tissues) from GTEx.

(**D**) PCA of normalized RNA-Seq-derived counts of indicated genes from 178 primary pancreatic tumors, 1 metastatic pancreatic tumor and 4 normal adjacent pancreas tissues from TCGA pancreas adenocarcinoma group, and 165 healthy pancreas tissues from GTEx.

(**E**) PCA of normalized RNA-Seq-derived counts of indicated genes from 495 primary prostate tumors, 1 metastatic prostate tumor and 52 normal adjacent prostate tissues from TCGA prostate adenocarcinoma group, and 100 healthy prostate tissues from GTEx.

(**F**) PCA of normalized RNA-Seq-derived counts of indicated genes from 102 primary skin tumors, 366 metastatic skin tumor and 1 normal adjacent skin tissue from TCGA skin cutaneous melanoma group, and 556 healthy skin tissues (including sun exposed and not sun exposed skin tissues) from GTEx.

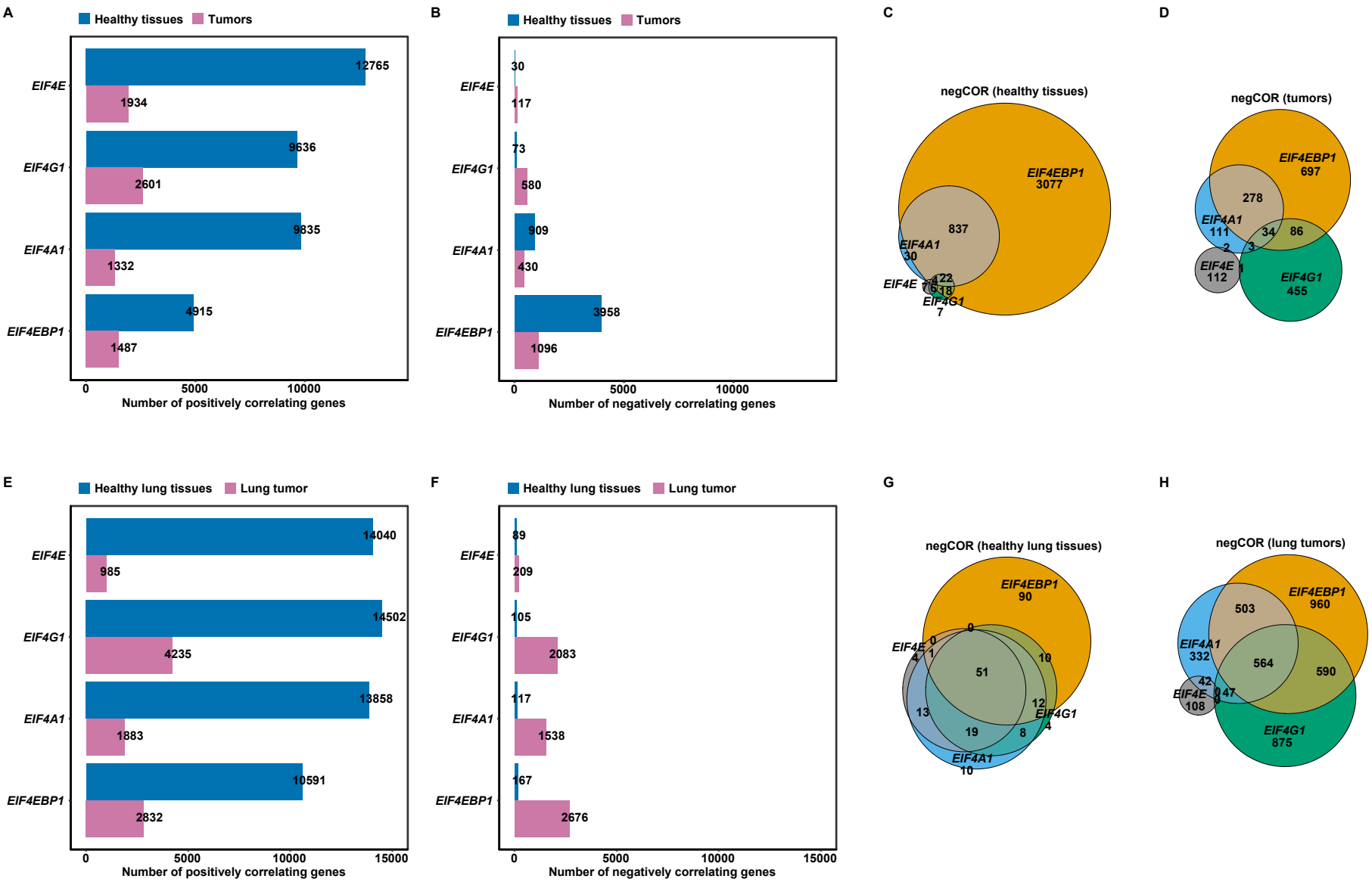


Figure S10. *EIF4F* Subunits Have More Correlations with Other Genes in Healthy Tissues than in Tumors, Related to Figure 5

(**A** and **B**) Pearson's correlation coefficients between *EIF4F* (*EIF4E*, *EIF4G1*, *EIF4A1*, or *EIF4EBP1*) and each of 58,582 other genes were calculated separately across 10,323 TCGA tumor samples from different cancer types, or across 7,414 GTEx healthy samples from different tissue types, using the Toil recomputed RNA-Seq datasets. Genes with significant positive (R > 0.3) or negative (R < -0.3) correlations were selected for analysis. The bar plots show the numbers of posCORs (**A**) and negCORs (**B**) identified for each *EIF4F* gene in tumors or in healthy tissues.

(**C** and **D**) The Venn diagrams show overlapping negCOR counts for *EIF4F* in healthy tissue samples (**C**), or in tumors (**D**).

(**E** and **F**) Pearson's correlation coefficients between *EIF4F* (*EIF4E*, *EIF4G1*, *EIF4A1*, or *EIF4EBP1*) and each of 58,582 other genes were calculated separately across 1,122 lung tumors of LUSC and LUAD from TCGA, or in 288 healthy lung tissues from GTEx, using the Toil recomputed RNA-Seq datasets. Genes with significant positive (R > 0.3) or negative (R < -0.3) correlations were selected for analysis. The bar plots show the numbers of posCORs (**E**) and negCORs (**F**) identified for each *EIF4F* gene in tumors or in healthy tissues.

(**G** and **H**) The Venn diagrams show overlapping negCOR counts for *EIF4F* genes identified in healthy lungs (**G**), or in lung tumors (**H**).

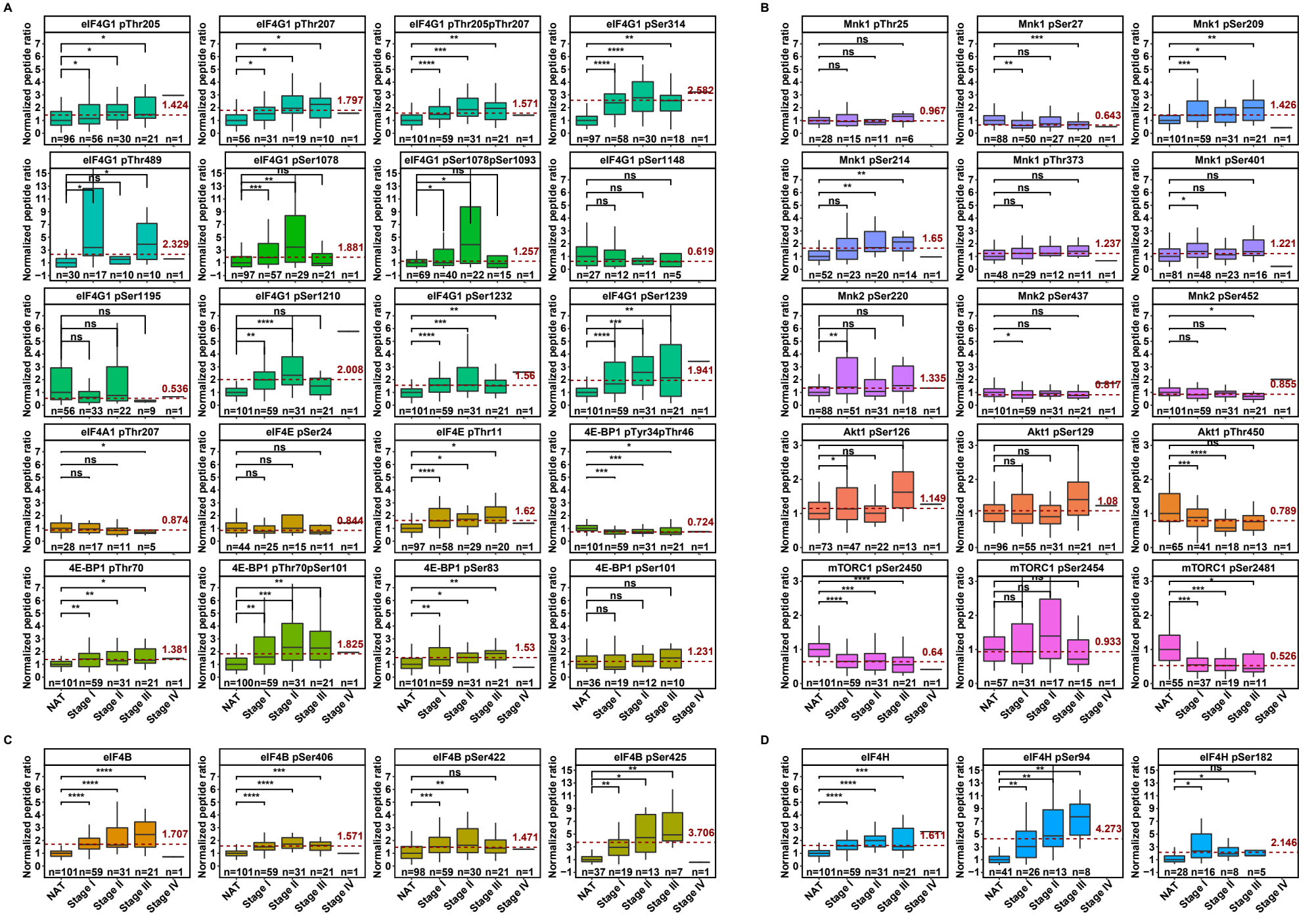


Figure S11. Phosphorylation of eIF4F Subunits in Lung Adenocarcinoma, Related to Figure 6.

(**A**) The box plots show, for eIF4F subunits, ratios to paired NATs (Y axis) of mean peptide abundance (phospho-peptide with phosphorylation at indicated serine or threonine residues). Mean expression from NATs was normalized at 1. Phospho-peptides, containing one or more fully localized phosphorylation modifications, are mapped to reference protein sites for eIF4G1(NP_886553.3), eIF4A1(NP_001407.1), eIF4E(NP_001959.1), and 4E-BP1 (NP_004086.1). Tumor stages (for each column) and sample counts (n, for each plot) are marked on X axes. The dashed red line marks average abundance in all tumor stages combined, relative to NATs. The two-tailed Student's t tests were performed. ns, not significant; *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001

(**B**) The box plots compare abundances of Mnk1/2 and Akt1 phosphorylations. Phospho-peptides, containing one or more fully localized phosphorylation modifications, are mapped to reference protein sites for Mnk1(NP_003675.3), Mnk2(NP_951009.1), Akt1(NP_001014431.1), and mTORC1 (NP_004949.1).

(**C** and **D**) The box plots compare abundances of eIF4B and eIF4H, and their phosphorylations. Phospho-peptides, containing one or more fully localized phosphorylation modifications, are mapped to reference protein sites for eIF4B (NP_001287750.1) and eIF4H (NP_071496.1).