**00. Data normalization**

1] Aim: Normalize expression levels of mRNAs, proteins, and phosphopeptides

2] Required function(s): quantilenorm2

3] Input variables

[1] FPKM (# genes × # patients): FPKM values of all genes

[2] raw (# TMT sets × 1): List of peptides and corresponding proteins for each TMT set

[3] pids (# TMT sets × 1): List of proteins for each TMT set

[4] eids (# TMT sets × 1): List of gene IDs (e.g., Entrez IDs) for each TMT set

[5] ind\_used (# total TMT channels × 1): Index for used TMT channels

4] Input files

[1] scan files for TMT sets: A file contains peptide sequence, PIP value, and peptide intensities of TMT channels for each MS2 scan.

5] Output variables

[1] fc\_qnorm (# expressed genes × # patients): Normalized fold-changes of expressed genes (detected in more than 50% of the patients)

[2] IDs\_exp (# expressed genes × 1): List of expressed genes

[3] fc\_merged\_qnorm (# expressed proteins/phosphopeptides × # patients): Normalized fold-changes of expressed proteins/phosphopeptides (detected in all patients)

[4] fc\_merged\_qnorm2 (# expressed proteins/phosphopeptides × # patients): Normalized fold-changes of expressed proteins/phosphopeptides (detected in more than 50% of the patients)

[5] proteins\_union\_exp2 (# expressed proteins × 1): List of expressed proteins (detected in more than 50% of the patients)

[6] peptides\_union\_exp2 (# expressed proteins × 1): List of expressed phosphopeptides (detected in more than 50% of the patients)

[7] fc\_merged\_qnorm\_20pct (# expressed proteins/phosphopeptides × # patients): Normalized fold-changes of expressed proteins/phosphopeptides (detected in more than 20% of the patients)

[8] proteins\_union\_20pct (# expressed proteins × 1): List of expressed proteins (detected in more than 20% of the patients)

[9] peptides\_union\_20pct (# expressed proteins × 1): List of expressed phosphopeptides (detected in more than 20% of the patients)

**01. Mutation-phosphorylation correlation**

1] Aim: Identifify phosphopeptides correlated with protein altering mutations

2] Required function(s): permall\_uneq\_ranksum, stat2\_uneq\_ranksum, calc\_ranksum\_stat, main\_degID\_uneq\_pw\_pm, auto

3] Input variables

[1] fc\_merged\_qnorm (# expressed peptides × # patients): Normalized fold-changes of expressed phosphopeptides (detected in more than 50% of the patients)

[2] ind\_mut (# genes × # patients): Index for protein altering mutations in patients

[3] peptides\_union\_exp2 (# expressed peptides × 1): List of expressed phosphopeptides

4] Output variables

[1] peptides\_corr (# genes × 1): List of phosphopeptides correlated with protein altering mutations

**02. Mutation-protein and phosphopeptide association**

1] Aim: Identify association of somatic mutations with protein and phosphopeptide abundance

2] Required function(s): none

3] Input variables

[1] Mut\_type\_index (# patients × # genes): Index for protein altering mutations in patients

[2] genes\_20pct\_prot (# expressed proteins × 1): List of gene IDs corresponding to expressed proteins

[3] genes\_20pct\_phos (# expressed peptides × 1): List of gene IDs corresponding to expressed phosphopeptides

[4] freq\_genes (# genes × 1): List of gene IDs corresponding to frequently mutated genes

[5] fc\_merged\_qnorm\_20pct\_prot (# proteins × # patients): Normalized fold-changes of expressed proteins (detected in more than 20% of the patients)

[6] fc\_merged\_qnorm\_20pct\_phos (# peptides × # patients): Normalized fold-changes of expressed phosphopeptides (detected in more than 20% of the patients)

[7] proteins\_union\_20pct (# expressed proteins × 1): List of expressed proteins (detected in more than 20% of the patients)

[8] peptides\_union\_20pct (# expressed peptides × 1): List of expressed phosphopeptides (detected in more than 20% of the patients)

4] Output variables

[1] proteins\_assoc (# genes × 2): List of proteins whose abundance is positively or negatively associated with somatic mutations

[2] peptides\_assoc (# genes × 2): List of phosphopeptides whose abundance is positively or negatively associated with somatic mutations

**03. mRNA-protein correlation**

1] Aim: Identify pairs of mRNAs and proteins with significant correlations between mRNA and protein abundances

2] Required function(s): fdr\_bh

3] Input variables

[1] proteins (# expressed proteins × 1): List of expressed proteins in gene IDs (detected in more than 50% of the patients)

[2] genes (# expressed genes × 1): List of expressed genes in gene IDs (detected in more than 50% of the patients)

[3] fc\_merged\_qnorm\_prot (# proteins × # patients): Normalized fold-changes of expressed proteins (detected in more than 50% of the patients)

[4] fc\_qnorm\_gene (# genes × # patients): Normalized fold-changes of expressed genes (detected in more than 50% of the patients)

4] Output variables

[1] genes\_sig (# genes × 1): List of genes whose mRNA and protein abundance are significantly correlated

**04. Subtype identification (mRNA)**

1] Aim: Identify mRNA clusters by two-stage clustering & Identify molecular signatures for the mRNA clusters

2] Required function(s): aoNMF\_subtyping, aonmf, auto, scale, aoNMF\_subtyping\_fix, aonmf\_fix, permall\_uneq, main\_degID\_uneq

3] Input variables

[1] fc\_qnorm (# expressed genes × # patients): Normalized fold-changes of expressed genes (detected in more than 50% of the patients)

[2] cellularity\_rank (# patients × 1): Averaged rank of cellularity

[3] IDs\_exp (# expressed genes × 1): List of expressed genes (detected in more than 50% of the patients)

4] Output variables

[1] clus\_mad30\_c2 (1 × # high cellularity patients): 1st stage cluster membership of patients with high cellularity

[2] clus\_mad20\_c3\_fix (1 × # patients): 2nd stage cluster membership of patients

[3] ind\_core\_mad20\_c3\_fix (# patients × 1): Index for core samples in mRNA data

[4] gene\_sig\_mad20\_c3\_fix (# gene signatures × 2): List of gene signatures for mRNA clusters

**05. Subtype identification (protein)**

1] Aim: Identify protein clusters by two-stage clustering & Identify molecular signatures for the protein clusters

2] Required function(s): aoNMF\_subtyping\_NaN, aonmf\_NaN, nanauto, cluster\_counter, silhouette\_NaN, sqeuclidean\_NaN, scale, aoNMF\_subtyping\_NaN\_fix, aonmf\_NaN\_fix, permall\_uneq, main\_degID\_uneq, signature\_extractor

3] Input variables

[1] fc\_merged\_qnorm (# expressed proteins × # patients): Normalized fold-changes of expressed proteins (detected in all patients)

[2] cellularity\_rank (# patients × 1): Averaged rank of cellularity

[3] fc\_merged\_qnorm2 (# expressed proteins × # patients): Normalized fold-changes of expressed proteins (detected in more than 50% of the patients)

[4] proteins\_union\_exp2 (# expressed proteins × 1): List of expressed proteins (detected in more than 50% of the patients)

4] Output variables

[1] clus\_mad30\_c3 (1 × # high cellularity patients): 1st stage cluster membership of patients with high cellularity

[2] clus\_mad20\_c5\_fix (1 × # patients): 2nd stage cluster membership of patients

[3] protein\_sig\_mad20\_c5\_fix (# protein signatures × 2): List of protein signatures for protein clusters

**06. Subtype identification (phospho)**

1] Aim: Identify phophopeptide clusters by two-stage clustering & Identify molecular signatures for the phophopeptide clusters

2] Required function(s): aoNMF\_subtyping\_NaN, aonmf\_NaN, nanauto, cluster\_counter, silhouette\_NaN, sqeuclidean\_NaN, scale, aoNMF\_subtyping\_NaN\_fix, aonmf\_NaN\_fix, permall\_uneq, main\_degID\_uneq, signature\_extractor

3] Input variables

[1] fc\_merged\_qnorm (# expressed phosphopeptides × # patients): Normalized fold-changes of expressed phosphopeptides (detected in all patients)

[2] cellularity\_rank (# patients × 1): Averaged rank of cellularity

[3] fc\_merged\_qnorm2 (# expressed phosphopeptides × # patients): Normalized fold-changes of expressed phosphopeptides (detected in more than 50% of the patients)

[4] peptides\_union\_exp2 (# expressed phosphopeptides × 1): List of expressed phosphopeptides (detected in more than 50% of the patients)

4] Output variables

[1] clus\_mad30\_c3 (1 × # high cellularity patients): 1st stage cluster membership of patients with high cellularity

[2] clus\_mad20\_c5\_fix (1 × # patients): 2nd stage cluster membership of patients

[3] peptide\_sig\_mad20\_c5\_fix (# genes × 2): List of phosphopeptide signatures for phosphopeptide clusters

**07. Integrated clustering**

1] Aim: Perform integrated clustering using all three types of data (mRNA, global proteome, and phosphoproteome)

2] Required function(s): none

3] Input variables

[1] clus (# patients × 3): Cluster membership from each type of data (one column for each type)

4] Output variables

[1] int\_clus (# patients × 1): Integrated clusters for all patients

**08. Application of mRNA signatures**

1] Aim: Apply our mRNA signatures to other cohorts and classify the patients in other cohorts into the mRNA clusters defined by our data

2] Required function(s): quantilenorm\_mean, quantilenorm\_given, corr\_pval\_emp\_v2

3] Input variables

[1] FPKM (# genes × # patients): FPKM values of all genes in our cohort

[2] ind\_core\_mad20\_c3\_fix (# patients × 1): Index for core samples in mRNA data

[3] IDs\_exp (# expressed genes × 1): List of expressed genes (detected in more than 50% of the patients)

[4] FPKM\_other (# genes × # patients): FPKM values of all genes in other cohort

[5] IDs\_other (# genes × 1): List of all genes in other cohort

[6] gene\_sig\_C1\_mad20\_c3\_fix (# RNA1 signatures × 1): List of RNA1 signatures

[7] gene\_sig\_C2\_mad20\_c3\_fix (# RNA2 signatures × 1): List of RNA2 signatures

[8] gene\_sig\_C3\_mad20\_c3\_fix (# RNA3 signatures × 1): List of RNA3 signatures

[9] clus\_mad20\_c3\_fix\_core (1 × # core samples): Cluster membership of core samples

4] Output variables

[1] rho\_other (1 × # patients): Three observed *ρ* values of RNA1–3

[2] pval\_other (1 × # patients): p-values for the three observed *ρ* values of RNA1–3 using the corresponding estimated distributions

[3] clus\_other (1 × # patients): Cluster membership of patients in other cohort