Equation for color representations of relative peptide abundance in heatmap

In a heatmap, color representations of relative peptide abundance were generated by the following equation:

$$\bar{x} = \sqrt{\max(x_{i,j}) \times \min(x_{i,j})}$$

$$f(x_{i,j}) = 255 \times \frac{\left| Ln \dfrac{x_{i,j}}{\bar{x}} \right|}{Ln \sqrt{\max\left(\left(\dfrac{\max(x_{i,j})}{\min(x_{i,j})}\right)_j\right)}}$$

$$color\ code = \begin{cases} \mathrm{RGB}(0,0,f(x_{i,j})) & x_{i,j} < \bar{x} \\ \mathrm{RGB}(0,0,0) & x_{i,j} = \bar{x} \\ \mathrm{RGB}(f(x_{i,j}),f(x_{i,j}),0) & x_{i,j} > \bar{x} \end{cases}$$

$x_{i,j}$: SIC peak area of each timepoint (i) for each peptide (j)

$\max(x_{i,j})$: max peak area of peptide j

$\min(x_{i,j})$: min peak area of peptide j

$\mathrm{RGB}(r,g,b)$: function to generate a color based on three parameters. r, g and b representing the intensity of red, green and blue, ranging from 0 to 255.

Methods for Clustering

Peptides with missing data or whose max fold change was less than the max fold change of Lyn Y^{194} (4.34) were discarded. The peak areas of the remaining peptides were normalized to have a mean of 0 and a standard deviation of 1 in order to capture the temporal profile of each peptide. The normalized peak areas were clustered using a fuzzy k-means clustering algorithm called MFuzz (2). MFuzz assigns a membership value to each peptide for every cluster, allowing partial membership of a peptide to more than one group. Peptides with less than 70% membership to the best cluster were removed.

Since the number of clusters, k, was specified as a parameter to the algorithm, the optimal number of clusters was determined by running the clustering algorithm 100 times with k = 3, .., 9. The best run was chosen for each k as the smallest sum of all proteins to their assigned cluster center. We found that 7 clusters representing 64 proteins best represent the data. Two of the seven clusters have 3 or fewer proteins: one contains Acetoacetyl-coenzyme a synthetase Y^{524} and Actin-associated tyrosine-phosphorylated protein Y^{323}. The other cluster contains Glycyl tRNA synthetase Y^{443}, Ribosomal protein L10a Y^{11}, and Ribosomal protein S10 Y^{12}. As the temporal profiles of these two clusters are not informative, only the remaining five clusters are shown in Figure 8.

Corresponding sequence motifs around the phosphorylation sites were found according to Minimotif Miner (MnM) database.