**Phospho-Antibody Array Data Analysis (PANDA) – A Primer**

Li Xiong\(^1\), Pu Shi\(^1\), Wei Zhou\(^2\)

\(^1\) Department of Mathematics and Computer Science, Emory University, Atlanta, Georgia, 30322
\(^2\) Winship Cancer Institute, Emory University School of Medicine, Atlanta, GA 30322

PANDA is a web-based program that identifies phosphorylated antibodies in phosphor-specific microarrays and statistically quantifies the extent of phosphorylation for these antibodies. The current version of the software performs the following analysis.

**Step 1. Preprocessing.** The background signals are first removed from all measurements. For each antibody, its respective negative control value is removed from each measurement.

**Step 2. Replicate Analysis.** Each antibody comes with 6 replicates for more reliable and consistent results. A Grubb's test or ESD method (extreme studentized deviate) is first used to detect potential outliers that are statistically inconsistent with other replicates (possibly due to an experiment error). Specifically, the test calculates a ratio \(Z\) as the difference between the potential outlier value \(y_i\) and the mean of the replicates \(y_{\text{mean}}\) divided by the standard deviation SD.

\[
Z = \left| \frac{y_{\text{mean}} - y_i}{SD} \right|
\]

The ratio is then compared to a critical value determined by the number of replicates, 6, and the desired significance level, \(\alpha = 0.05\). The adjusted mean is then computed after the outliers are removed.

**Step 3. Phosphorylation Analysis.** A ratio computation was used to measure the extent of protein phosphorylation. For each antibody (that has phosphorylated and matching unphosphorylated values in both control data and experiment data, denoted by \(\text{phospho}\) and \(\text{unphospho}\)), PANDA computes the following phosphorylation ratio:

\[
\text{phosphorylation ratio} = \frac{\text{phospho}_{\text{experiment}}}{\text{unphospho}_{\text{experiment}}} \div \frac{\text{phospho}_{\text{control}}}{\text{unphospho}_{\text{control}}}
\]

A 95% confidence interval was used to quantify the precision of the phosphorylation ratio based on the replicate analysis.