Logo, company name

Description automatically generated

**The Hinds Laboratory Pathways Report**

**STK Analysis**

28 July, 2022

# Introduction

## Background

The Pamstation12 instrument provides a profiling of kinase activity of cell or tissue samples. The device is loaded with either serine/threonine or tyrosine microarray chips. Each chip has 4 wells for up to four samples, and 3 chips total can be loaded per run. The microarray represents 144 (STK chip) or 196 (PTK chip) reporter peptides that can be phosphorylated by serine/threonine or tyrosine kinases. Using fluorescently labeled antibodies, the PamStation12 measures the degree of the phosphorylation in real time by detecting at different exposure times. The list of peptides present in each microarray can be viewed here: [STK chip](https://pamgene.com/wp-content/uploads/2020/09/STK-144-PamChip-87102.pdf), [PTK chip](https://pamgene.com/wp-content/uploads/2020/09/PTK-196-PamChip-86402.pdf)

A picture containing diagram

Description automatically generatedChart

Description automatically generated with medium confidence

# Results

## Image Analysis

After the PamStation12 run is completed, the first step in data analysis is using BioNavigator, PamGene’s proprietary software, to analyze the images produced from each cycle. For each picture, the user must view the microarray to ensure the correct placement of the superimposed grid used to quantify peptide fluorescence. As each spot on the microarray represents a specific group of peptides used to determine the identify of kinases, it is imperative that this step be completed with detail. After the viewing and correction of grids, these images are summarized and quantified using PamGrid, which provides a numeric values using the median value of fluorescence minus the median value of the background pixels. The resulting file, and most important for further KRSA and UKA analysis, is the median signal minus background file (Median\_SigmBg).

## Reading Data

First, two BioNavigator files are uploaded: Median\_SigmBg and Signal\_Saturation. These files are read into the environment and cleaned to a tidy format for further analysis, which involves modeling and visualization. To continue with further analysis, we will define the chip type (STK or PTK).

## QC Initial Steps and Groups Assignments

The Signal\_Saturation file is used for quality control, which is used to adjust the negative values in the Median\_SigmBg file. Additionally, a new column is defined to group the existing variables. This combines the Sample Name and Barcode are combined for a unique identifier. Finally, the end point signal values at the max exposure time are extracted.

## QC Steps and Model Fitting

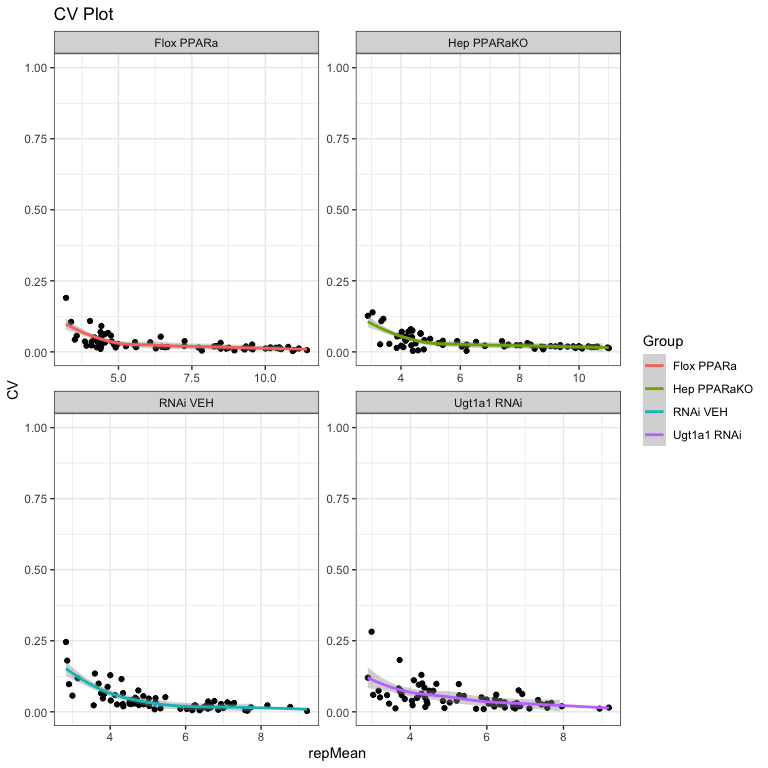
The peptides with the lowest signals are filtered out. A simple linear regression model is used to fit the multiple exposure time values into a single value. We accomplish this using the Median\_SigmBg as a function of the exposure time. The slope of the model fit and are employed for quality control step and for samples comparison. The slope is multiplied by 100 and log2 transformed. The peptides with a poor linear fit and reference peptides are filtered out.

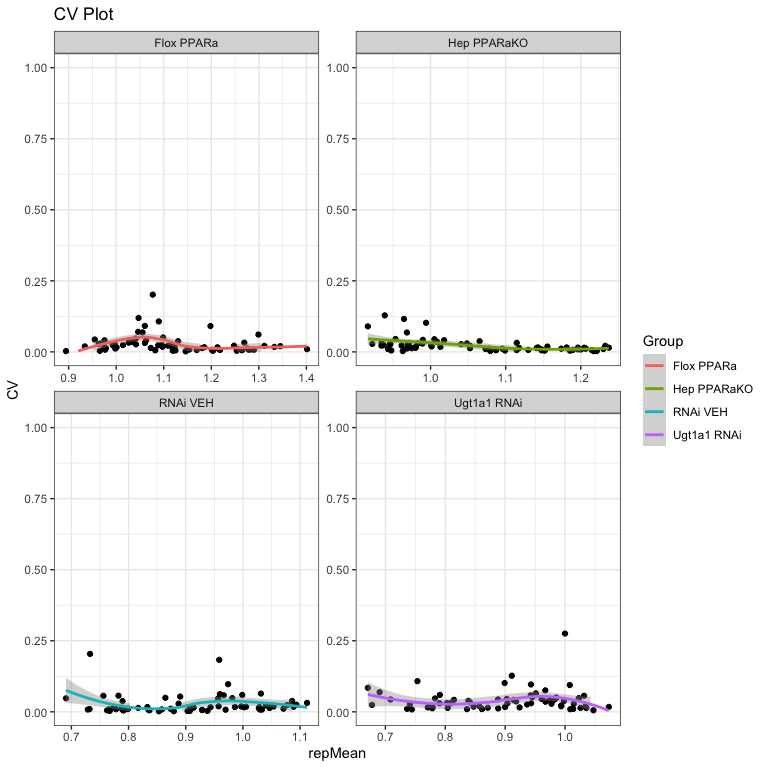
## Global Signal Intensity

The slope transformed values are used for global signal intensity determination across samples and groups. We will create CV and violin plots, and create heatmaps to visualize our fitted data.

### Global CV Plots

We will plot the coefficient of variation on both the normal and normalized fits. This will help us to identify groups with high variation that could be explained by sample outliers.

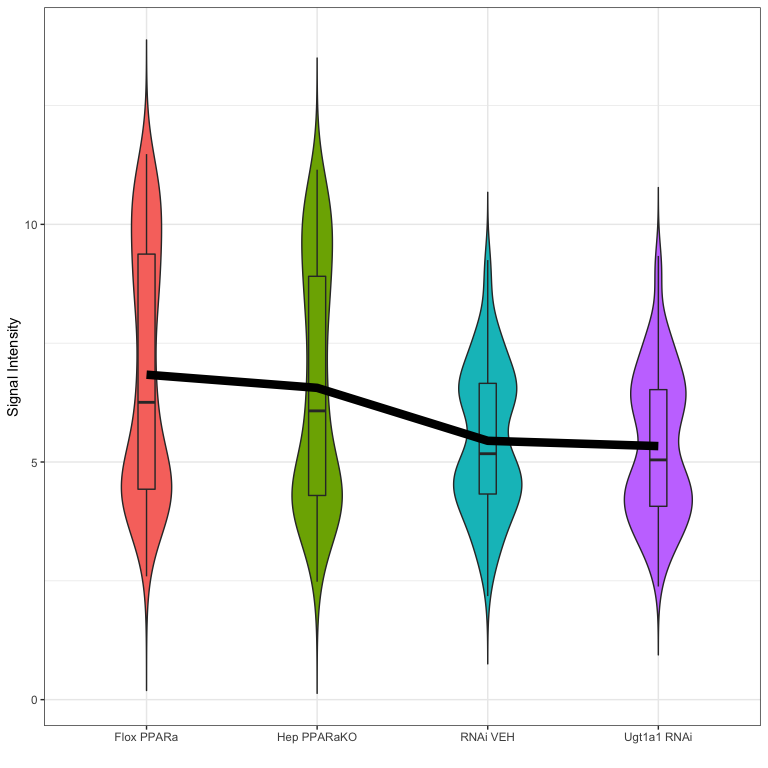
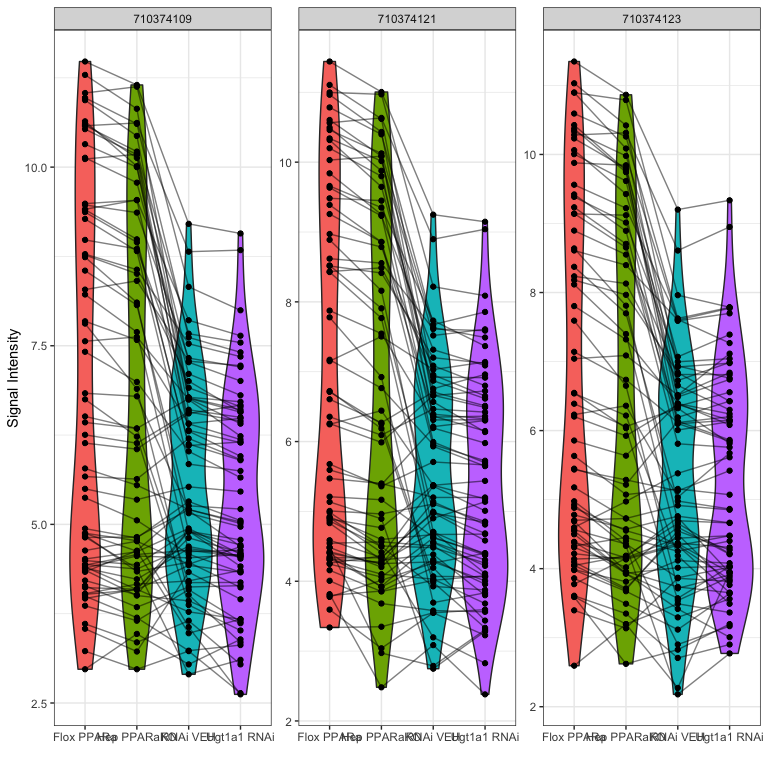




### Global Violin Plots

Violin plots will help us to examine the global signal differences between the groups/samples. Chart

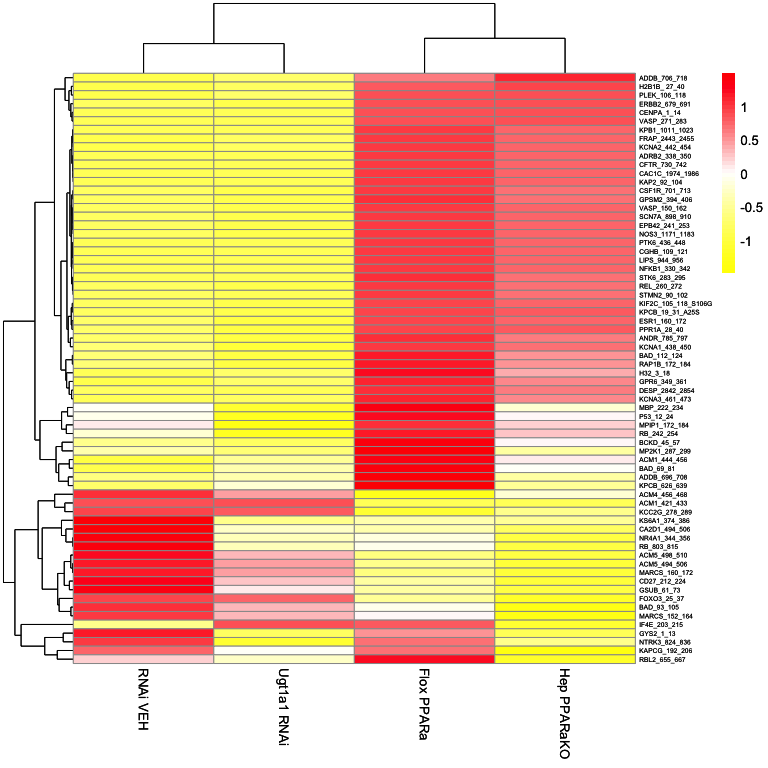
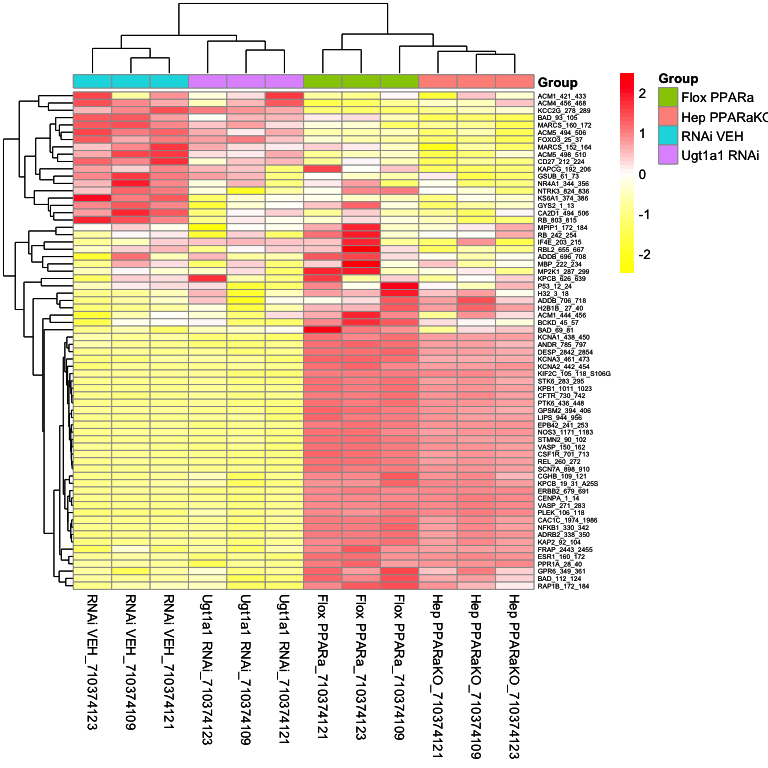
Description automatically generated



### Global Heatmaps

The global heatmap represents all peptides present on the chip, with the postive/internal controls removed. The peptides that failed to pass quality control are also removed. The heatmaps shown below are scaled by row to highlight the similar/different peptide signals across the samples. Unsupervised hierarchial clustering is used on both peptides and samples to highlight potentially similar peptide signal signatures. Chart

Description automatically generated



## Group Comparison

We can also compare the differential peptide signals between samples/groups. For this report, the two group comparsion is:

* **Hep PPARaKO vs Flox PPARa**

The *Slope\_Transformed* ratio between each group, paired by chip, is calculated to the fold change. Based on the fold change, peptides that pass a certain fold change threshold are considered significant hits. Also, quality control steps applied in each comparison to filter out peptides that do not reach specific criteria:

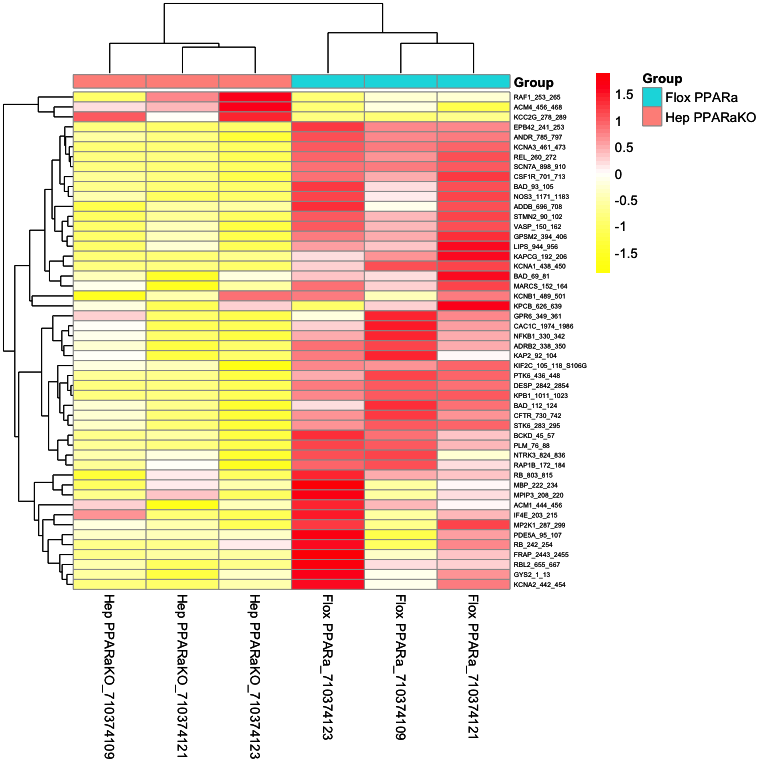
* The *Median\_SigmBg* at max exposure *200ms* must be above a certain value
* of the linear model fit must be above a threshold value

These *Filtering Parameters* (fold change threshold, QC criteria) can be modified to adjust the stringency of the analysis. The *Filtering Parameters* that are used for this analysis:

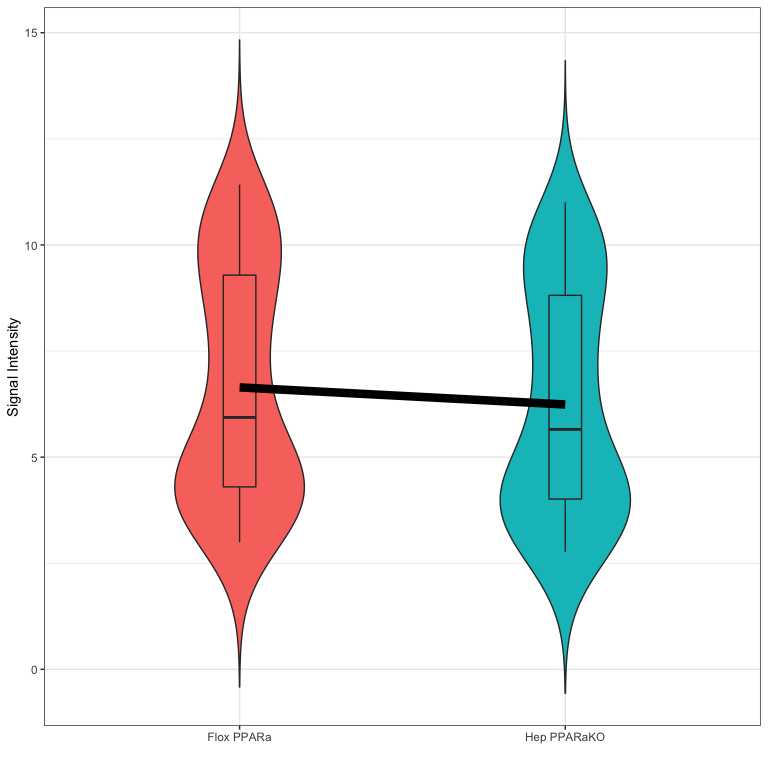
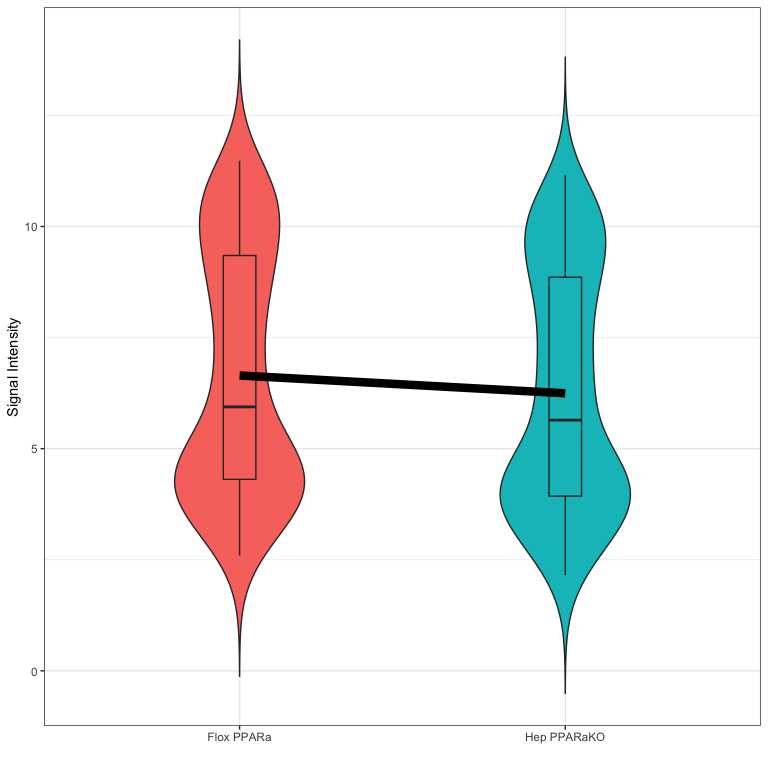
* The *Median\_SigmBg* at max exposure *200ms* must be equal or above 5
* of the linear model fit must be above or equal 0.9
* Log fold change (LFC) cutoffs at (0.2,0.3,0.4)

### Hep PPARaKO vs Flox PPARa (Heatmap)

After applying the *Filtering Parameters* for this group comparison, only *50* peptides carried forward in the analysis (i.e. *50 hits*). Below are some figures to visualize the differences between these samples for considering these *hits*.



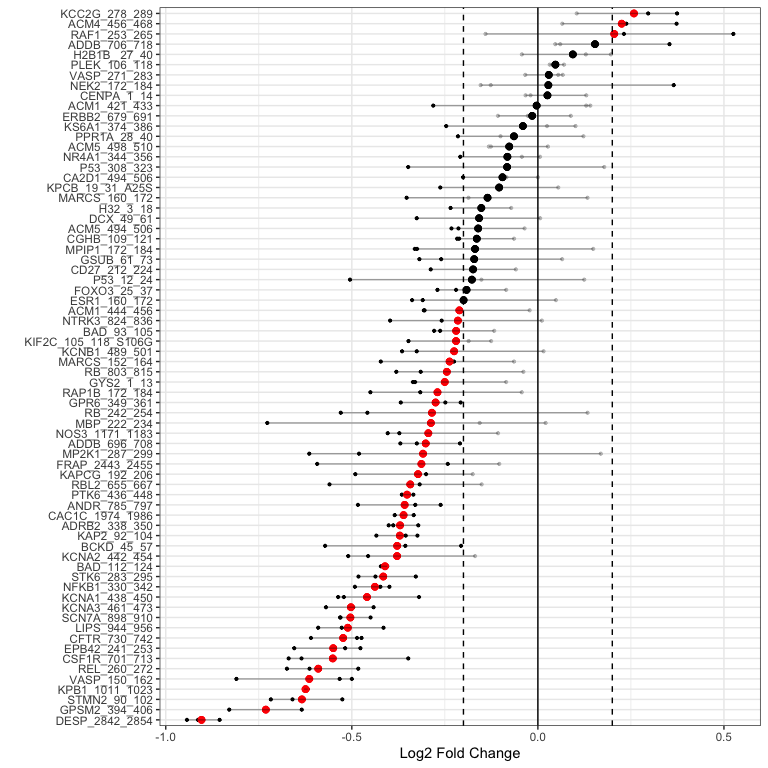
### Hep PPARaKO vs Flox PPARa (Violin Plot) A picture containing text, measuring stick, different, row Description automatically generated



### Hep PPARaKO vs Flox PPARa (Waterfall Plot)

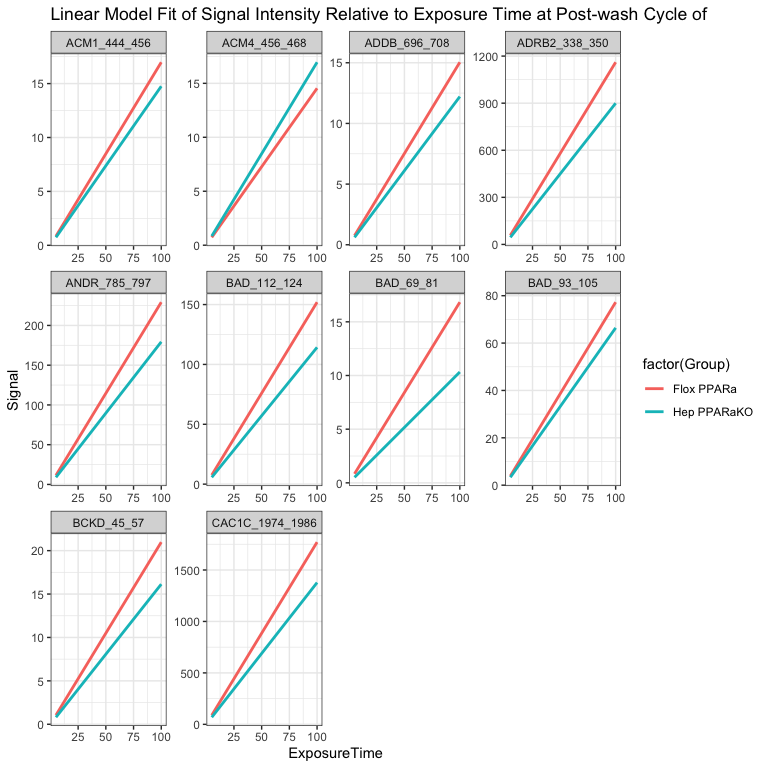
This waterfall represents the log2 fold changes between the two groups at each peptide. Chart

Description automatically generated



### Hep PPARaKO vs Flox PPARa (Curve Plot)

We generate curves plots per peptide using the last cycle data to visualize the linear fit model



### Hep PPARaKO vs Flox PPARa (Upstream Kinase Analysis (UKA) - Kinase Families)

The Cognitive Disorders Research Laboratory (CDRL) carefully curated and mapped the kinases that can act and phosphorylate each peptide present on the chip. This was achieved by using multiple sources including GPS 3.0, Kinexus Phosphonet, PhosphoELM and PhosphoSite Plus. Based on that association between peptides and kinases, a random sampling analysis is performed for these hits. The basic idea of *KRSA* is: For each iteration (*2000* iterations performed in this analysis), the same number of hits are randomly selected from the total 141/or 193 peptides present on the chip. This process can be likened to “boostrapping”, which refers to repeated simulated sampling to determine significance of observed results. Predicted kinases are then mapped to this sample list of peptides and number of kinases are determined. The kinase count from the actual hits and random sampling is then compared to determine the significance. The first version of this random sampling will look at kinases grouped by their relative families.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Kinase | Observed | SamplingAvg | SD | Z |
| MAPKAPK | 25 | 17.301 | 2.703 | 2.848 |
| PKA | 29 | 21.220 | 2.781 | 2.798 |
| MELK | 16 | 9.920 | 2.283 | 2.663 |
| PAKA | 18 | 11.672 | 2.422 | 2.613 |
| SGK | 15 | 9.234 | 2.212 | 2.606 |
| NMO | 0 | 3.884 | 1.515 | -2.564 |
| AMPK | 14 | 8.544 | 2.166 | 2.519 |
| MARK | 5 | 2.151 | 1.183 | 2.408 |
| PKG | 23 | 16.995 | 2.652 | 2.264 |
| CAMK1 | 12 | 7.388 | 2.078 | 2.219 |
| PKD | 21 | 15.316 | 2.593 | 2.192 |
| MLCK | 4 | 1.766 | 1.051 | 2.126 |
| PKCI | 7 | 3.892 | 1.542 | 2.016 |
| PLK | 11 | 16.286 | 2.640 | -2.003 |
| AUR | 1 | 3.917 | 1.521 | -1.918 |
| DYRK | 14 | 19.157 | 2.842 | -1.815 |
| RIPK | 18 | 13.530 | 2.486 | 1.798 |
| PDHK | 27 | 21.966 | 2.846 | 1.768 |
| DMPK | 41 | 36.505 | 2.565 | 1.752 |
| MST | 5 | 8.855 | 2.211 | -1.743 |
| STE11 | 1 | 3.546 | 1.470 | -1.732 |
| RAD53 | 16 | 11.762 | 2.451 | 1.729 |
| GRK | 7 | 4.250 | 1.609 | 1.709 |

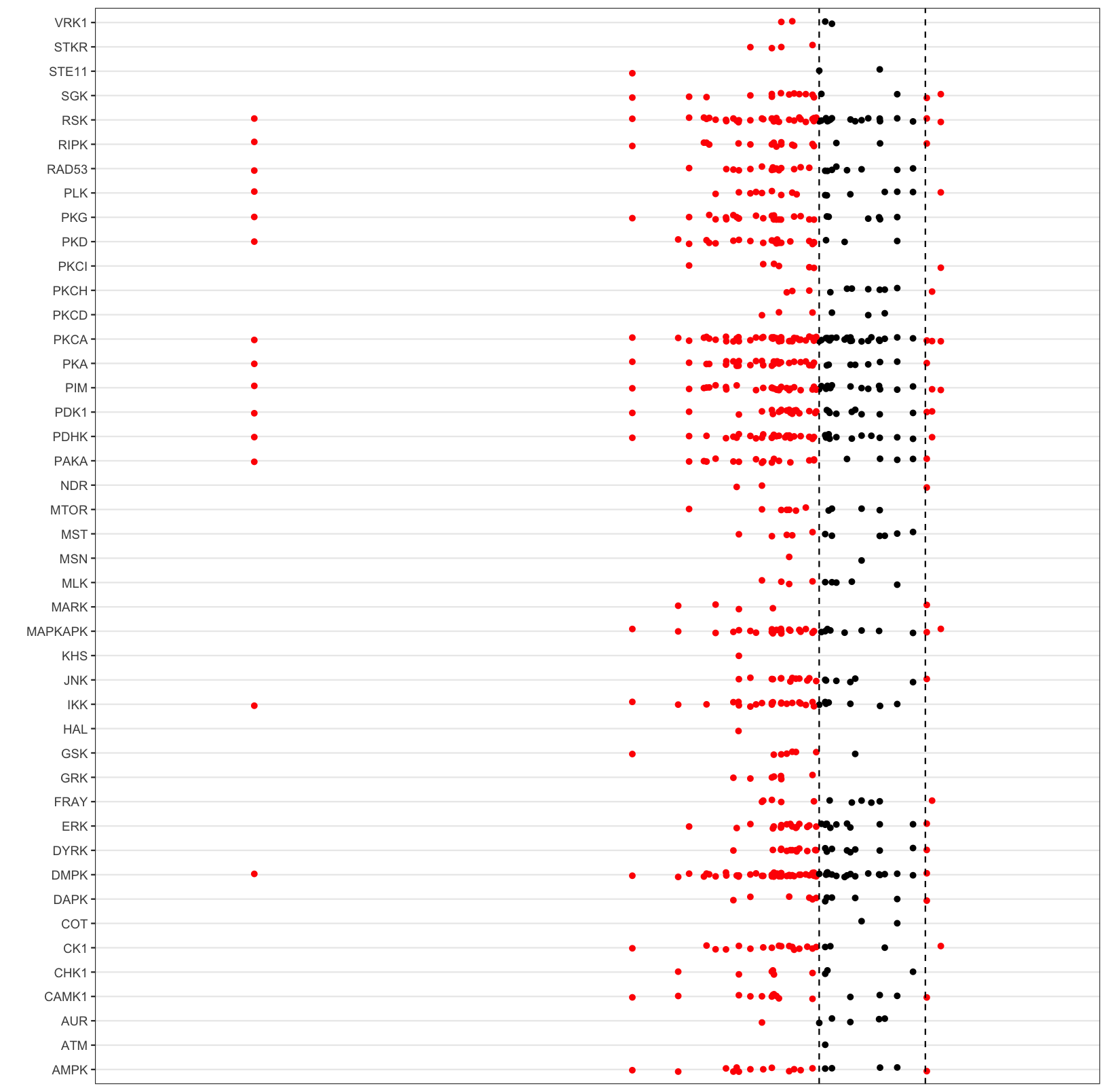
Diagram

Description automatically generated

#> Method NumberOfPeptides  
#> <chr> <int>  
#> 1 meanLFC.0.2 50  
#> 2 meanLFC.0.3 34  
#> 3 meanLFC.0.4 18  
#> 4 710374109.0.2 43  
#> 5 710374109.0.3 33  
#> 6 710374109.0.4 18  
#> 7 710374121.0.2 51  
#> 8 710374121.0.3 42  
#> 9 710374121.0.4 27  
#> 10 710374123.0.2 58  
#> 11 710374123.0.3 47  
#> 12 710374123.0.4 31

### Hep PPARaKO vs Flox PPARa (Reverse KRSA Plot)

We will use the reverse KRSA plot function, to plot the log2 fold chnage values for all peptides mapped to kinase hits. This will help us examine the activity of the kinase



(I need to add the labels for axes)

### Hep PPARaKO vs Flox PPARa (Coverage Plot)

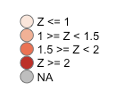
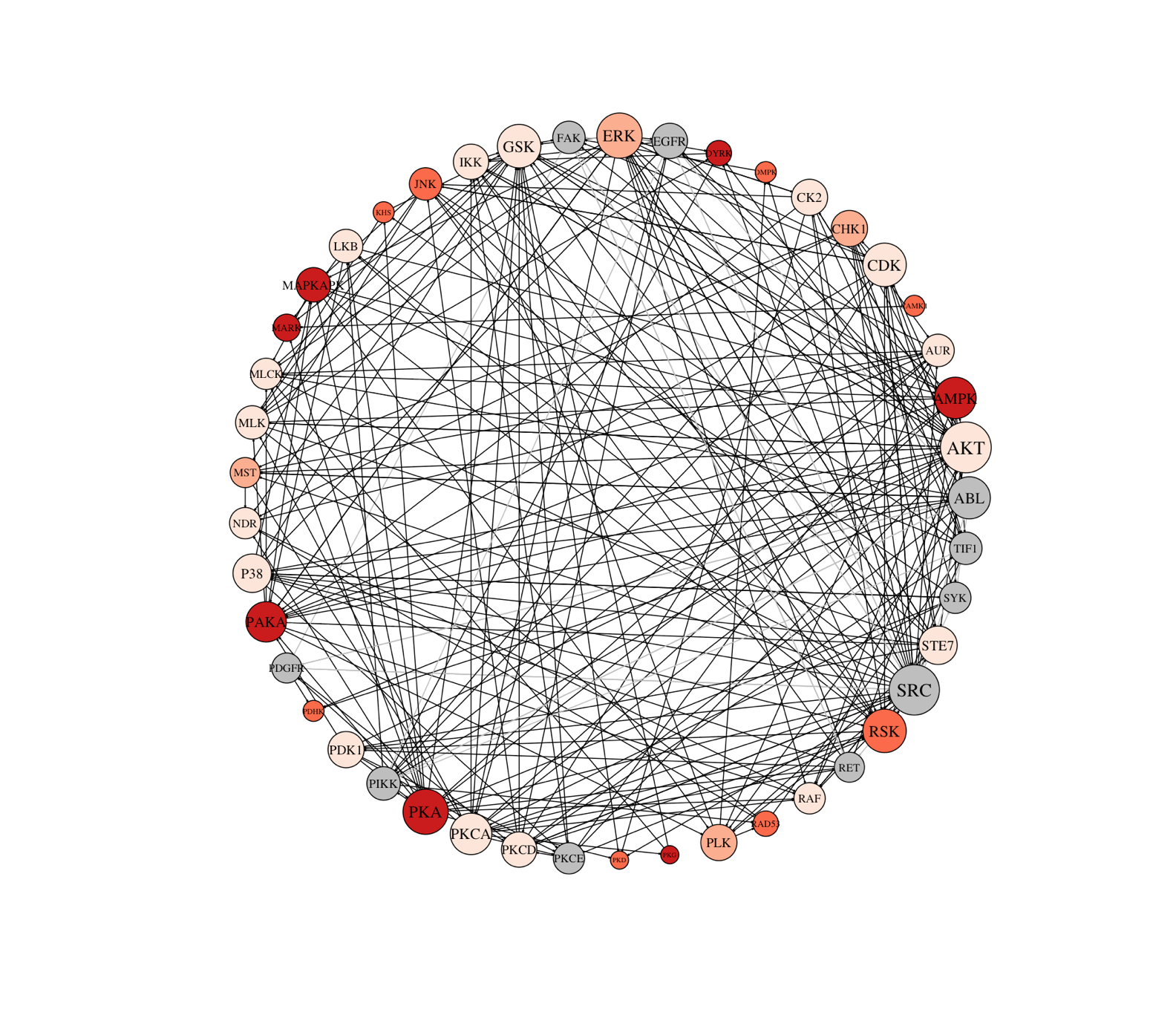
To view the coverage of kinases across the full list of peptides on the chip, we will use the coverage plot function

Chart, bar chart

Description automatically generated

### Hep PPARaKO vs Flox PPARa (Ball Model Network)

We will view the ball model network function, to generate a model representing the protein-protein interactions between kinases

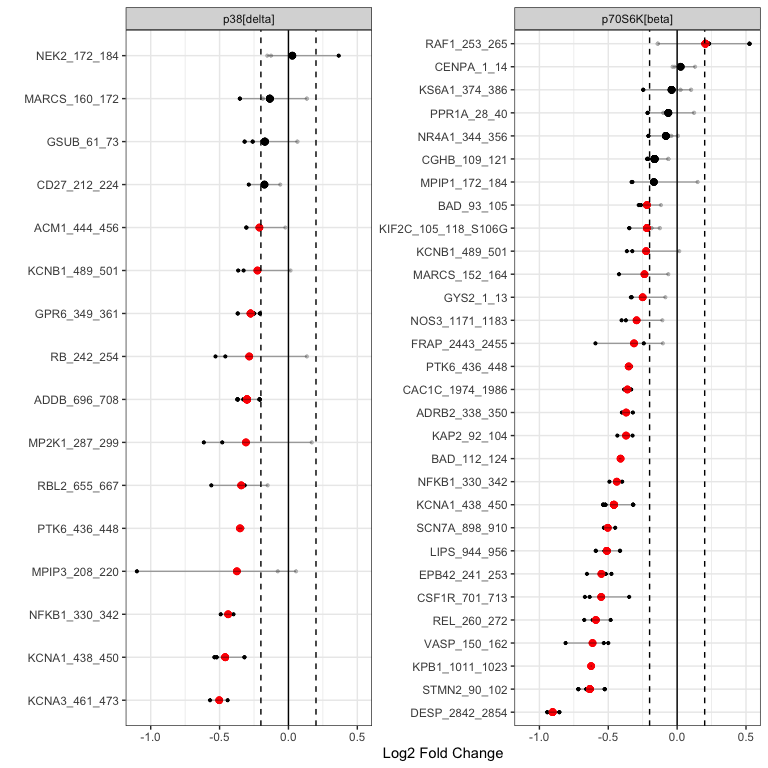


### Hep PPARaKO vs Flox PPARa (Upstream Kinase Analysis (UKA) - Individual Kinases)

We also apply the same bootstrapping/random sampling analysis to the individual kinases identified as over/under active in the UKA kinase family analysis. Instead of using mapping files that group kinases by family, we use PamGene’s carefully curated individual kinase list, which details all of the known peptides in which each kinase phosphorylates. This page marks the beginning of individual kinase analysis.

### Hep PPARaKO vs Flox PPARa (Waterfall Plot)

This waterfall represents the log2 fold changes between the two groups at each peptide for each selected kinase.



### Chart Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated

Table

Description automatically generated

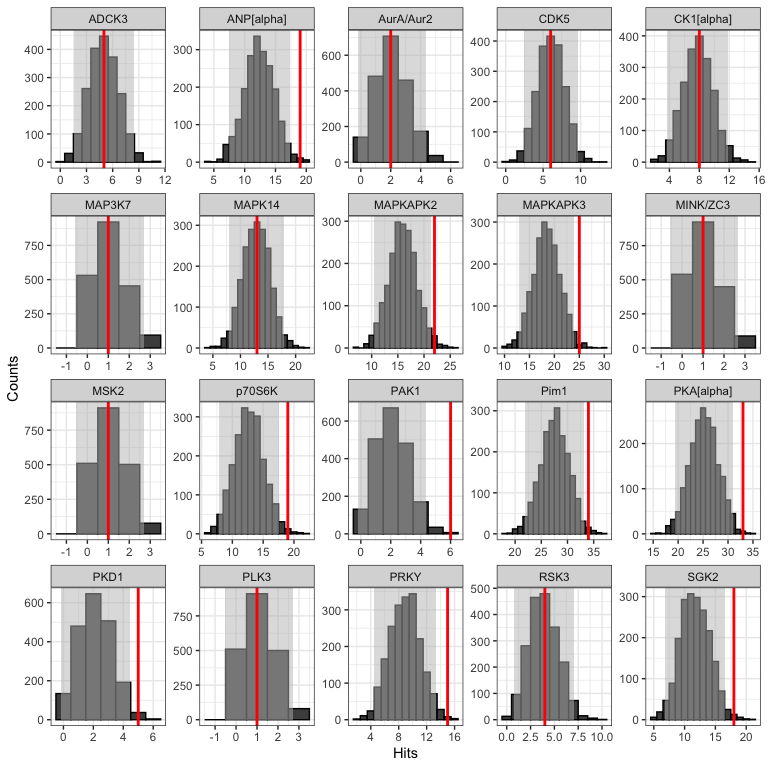
Chart

Description automatically generated

### Hep PPARaKO vs Flox PPARa (Upstream Kinase Analysis - Individual Kinases)

The following analysis applies the same UKA methodology but on individual kinases. The same kinases shown in the waterfall plots will be included as peacock plots in the following random sampling analysis.

| Kinase | Observed | SamplingAvg | SD | Z |
| --- | --- | --- | --- | --- |
| PAK1 | 6 | 2.094 | 1.141 | 3.425 |
| PRKY | 15 | 8.896 | 2.236 | 2.730 |
| PKA[alpha] | 33 | 25.186 | 2.899 | 2.695 |
| ANP[alpha] | 19 | 12.367 | 2.481 | 2.673 |
| SGK2 | 18 | 11.722 | 2.431 | 2.583 |
| p70S6K | 19 | 12.739 | 2.437 | 2.569 |
| PKD1 | 5 | 2.135 | 1.152 | 2.488 |
| MAPKAPK3 | 25 | 18.450 | 2.768 | 2.367 |
| Pim1 | 34 | 27.536 | 2.777 | 2.328 |
| MAPKAPK2 | 22 | 15.943 | 2.709 | 2.236 |
| PRKX | 26 | 19.797 | 2.806 | 2.211 |
| CaMK4 | 10 | 6.120 | 1.831 | 2.119 |
| CDKL2 | 4 | 1.775 | 1.062 | 2.095 |
| GSK3[alpha] | 1 | 4.280 | 1.571 | -2.088 |
| Pim3 | 34 | 28.053 | 2.854 | 2.084 |
| DCAMKL1 | 4 | 1.764 | 1.077 | 2.075 |
| p70S6K[beta] | 24 | 18.488 | 2.807 | 1.963 |
| PKA[gamma] | 2 | 0.724 | 0.663 | 1.926 |
| CHK2 | 16 | 11.370 | 2.421 | 1.912 |
| CHK1 | 8 | 4.954 | 1.667 | 1.828 |
| AMPK[alpha]1 | 8 | 4.956 | 1.751 | 1.738 |
| ARAF | 4 | 2.147 | 1.143 | 1.622 |
| VRK1 | 0 | 1.456 | 0.947 | -1.537 |
| PKG2 | 24 | 19.768 | 2.771 | 1.527 |
| HIPK1 | 0 | 1.451 | 0.956 | -1.519 |

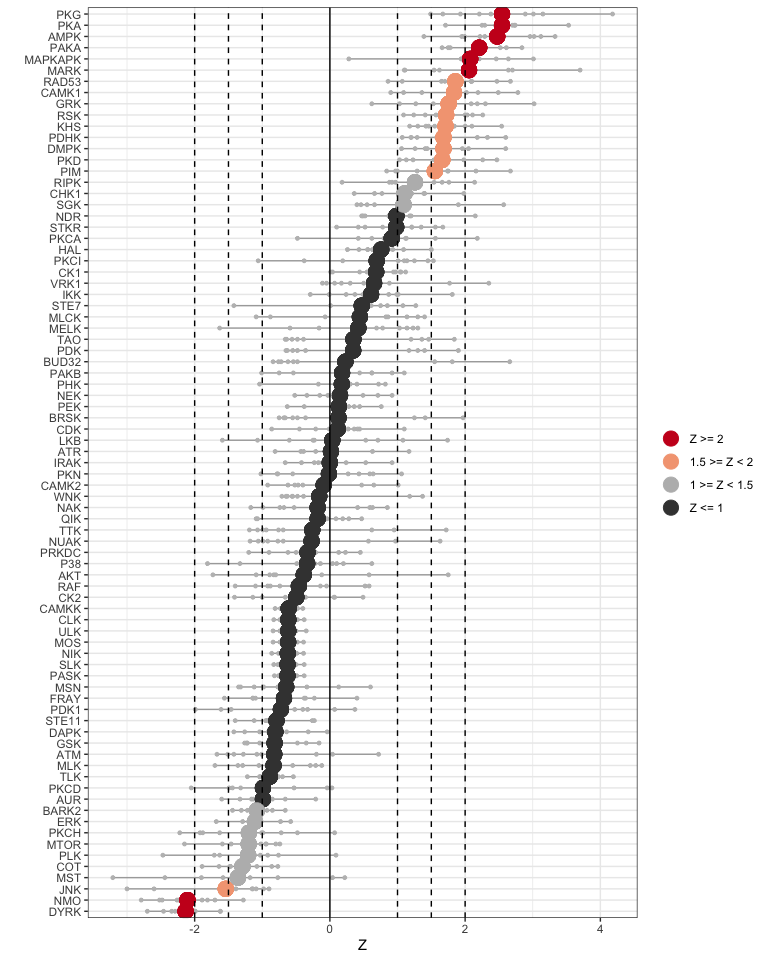


#> Method NumberOfPeptides  
#> <chr> <int>  
#> 1 meanLFC.0.2 50  
#> 2 meanLFC.0.3 34  
#> 3 meanLFC.0.4 18  
#> 4 710374109.0.2 43  
#> 5 710374109.0.3 33  
#> 6 710374109.0.4 18  
#> 7 710374121.0.2 51  
#> 8 710374121.0.3 42  
#> 9 710374121.0.4 27  
#> 10 710374123.0.2 58  
#> 11 710374123.0.3 47  
#> 12 710374123.0.4 31

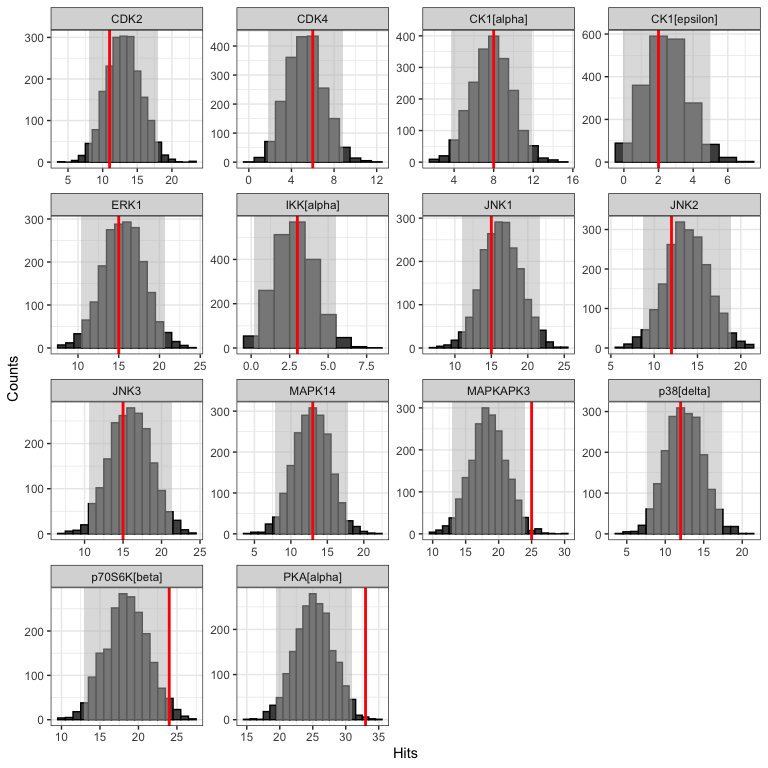
### Hep PPARaKO vs Flox PPARa (Z Scores Plot)

We will plot the individual and averaged Z scores using both the across and within chip analyses. Chart, line chart

Description automatically generated

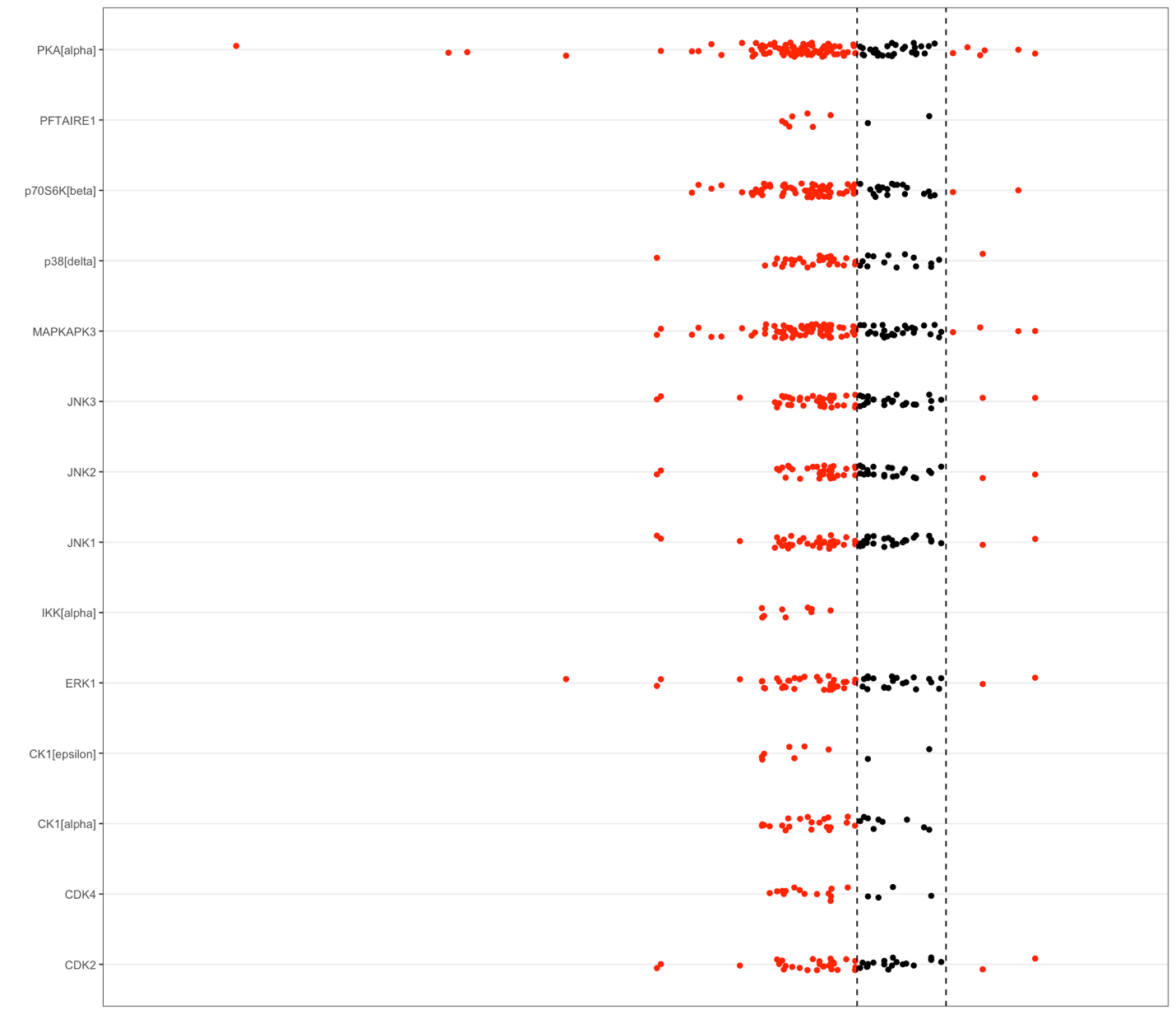


### Hep PPARaKO vs Flox PPARa individual kinases - Peacock plots



### Hep PPARaKO vs Flox PPARa (Reverse KRSA Plot)

We will use the reverse KRSA plot function, to plot the log2 fold change values for all peptides mapped to kinase hits. This will help us examine the activity of the kinase



#### Session Info

#> furrr future gt forcats stringr dplyr purrr readr   
#> "0.3.0" "1.25.0" "0.6.0" "0.5.1" "1.4.0" "1.0.9" "0.3.4" "2.1.2"   
#> tidyr tibble ggplot2 tidyverse knitr KRSA   
#> "1.2.0" "3.1.7" "3.3.6" "1.3.1" "1.39" "0.10.3"