

# Phosphoproteomic analysis of TMEM16A knockdown and EGF treatment on pancreatic cancer line AsPc-1

August 2, 2018

## Motivation

The calcium-activated chloride channel (TMEM16A), also known as anoctamin 1 (ANO1), is over-expressed in many cancers. The transmembrane protein is an attractive cancer target owing to its ability to activate many signaling pathways, including the epidermal growth factor receptor (EGFR) signaling, to support cell growth and migration. However, the effect of TMEM16A on cellular phenotype appears to be cell- and tumor-specific (Molecular Cancer 2017). For instance, Wu et al. found that TMEM16A over-expression promotes proliferation in ER-positive, PR-positive, and HER2-negative MCF-7 cells, but inhibits proliferation in ER-negative, PR-negative, and HER2-negative MDA-MB-435S cells (Oncotarget 2017). This suggests that further validation of TMEM16A as a viable target is needed on a per-cell or -tumor basis.

Pancreatic ductal adenocarcinoma (PDAC) is the most common type of pancreatic cancer, and TMEM16A is often over-expressed in PDAC cells. Given that TMEM16A directly interacts with EGFR, which in turn can trigger downstream growth signaling, it was hypothesized that TMEM16A over-expression drives the aggressive cell growth in PDAC. However, a study by Sauter et al. showed that TMEM16A silencing by siRNA did not reduce proliferation in PDAC cell lines (Eur J Physio 2015). In order to dissect the mechanism behind the sustained growth, we aim to probe the phosphoproteome of the PDAC cell line AsPc-1 by mass spectrometry.

## Workflow

1. David cultured three biological replicates of AsPc-1 cells.
  - Cell conditions:
    - shControl (shControl\_CTRL)
    - shControl + EGF (shControl\_EGF)
    - shTMEM16A (shTMEM\_CTRL)
    - shTMEM16A + EGF (shTMEM\_EGF)
  - Cells were washed twice in cold PBS and freshly lysed in buffer containing 8M urea, 0.1M Tris pH8.5, 40mM 2-AA, 10mM TCEP, and 1X HALT protease/phosphatase inhibitor.
  - Cell lysates were snap frozen and stored in -80°C.
2. Processed samples through the phosphoproteomics pipeline.
  - Sonicated cell lysates.
  - Digested 1mg lysate overnight with trypsin at 1:50 (w/w) enzyme-to-substrate ratio.
  - Desalted and enriched for phosphopeptides using Fe<sup>3+</sup>-IMAC columns.
  - Dried down and stored phosphopeptides in -80°C.
  - *Note: Replicate 1 processed on 02/14/2017 and replicates 2 and 3 processed on 06/01/2018.*
3. Analyzed samples on LC-MS/MS.
  - Reconstituted peptides and injected 1μg of each sample onto the instrument.
  - Data acquisition with LC-MS/MS on a 4-hour gradient.
  - Samples were queued up by replicates and sample conditions in the listed order.
    - First run: shControl\_CTRL\_bR1, shControl\_EGF\_bR1, shTMEM\_CTRL\_bR1, shTMEM\_EGF\_bR1.
    - Second run: shControl\_CTRL\_bR2, shControl\_EGF\_bR2, shTMEM\_CTRL\_bR2, shTMEM\_EGF\_bR2.

- Third run: **shControl\_CTRL\_bR3**, **shControl\_EGF\_bR3**, **shTMEM\_CTRL\_bR3**, **shTMEM\_EGF\_bR3**.
- 4. Processed MS data by MaxQuant using Phospho (STY) settings and match-between-runs.
- 5. Data analysis and visualization in R.
  - **Filtering**
    - Remove contaminants and reverse hits.
    - Retain confident phosphosite identifications.
    - Retain sites where at least two out of three replicates were quantified in one condition.
  - **Normalization**
    - Transform data by taking  $\log_2$  of phosphosite intensities.
    - Quantile normalization to remove technical variabilities.
  - **Imputation**
    - Hybrid imputation to fill in missing values with either a low value when the signal in replicates is low or a value assigned by maximum likelihood estimation when signal is medium or high.
  - **Statistical Testing**
    - Welch's two-sample t-test to compare any two conditions.

# Results

## Data Tables

The *Phospho (STY)Sites* MaxQuant output on phosphosite intensity was used in the following analyses. Please see *TMEM16A-EGF-phospho-analysis.xlsx* for the data tables. Table descriptions are outlined below.

- **raw:** Raw *Phospho (STY)Sites* output from MaxQuant.
- **shControl\_EGF-vs-shControl:** Effect of EGF on phosphorylation status. The *Log2FC* column shows the log<sub>2</sub>-transformed fold change in phosphorylation after EGF treatment.
  - Green highlight = Greater than 3-fold enrichment after EGF versus no treatment.
  - Red highlight = Greater than 3-fold depletion after EGF versus no treatment.
- **shTMEM\_EGF-vs-shTMEM:** Effect of EGF on phosphorylation status of TMEM-knockdown cells. The *Log2FC* column shows the log<sub>2</sub>-transformed fold change in phosphorylation after EGF treatment on TMEM-depleted cells.
  - Green highlight = Greater than 3-fold enrichment after EGF treatment on shTMEM cells.
  - Red highlight = Greater than 3-fold depletion after EGF treatment on shTMEM cells.
- **shTMEM-vs-shControl:** Effect of TMEM knockdown on phosphorylation status. The *Log2FC* column shows the log<sub>2</sub>-transformed fold change in phosphorylation after TMEM16A depletion.
  - Green highlight = Greater than 3-fold enrichment after TMEM16A knockdown.
  - Red highlight = Greater than 3-fold depletion after TMEM16A knockdown.
- **shTMEM\_EGF-vs-shControl:** Combined effect of TMEM knockdown and EGF treatment on phosphorylation status. The *Log2FC* column shows the log<sub>2</sub>-transformed fold change in phosphorylation after the treatment.
  - Green highlight = Greater than 3-fold enrichment after TMEM16A knockdown + EGF.
  - Red highlight = Greater than 3-fold depletion after TMEM16A knockdown + EGF.

Table 1: Summary statistics on the processed samples (after data filtering)

Samples	# sites detected	# sites quantified	% sites quantified
shControl_CTRL_bR1	5336	2783	52.2
shControl_CTRL_bR2	5336	3832	71.8
shControl_CTRL_bR3	5336	3344	62.7
shControl_EGF_bR1	5336	2817	52.8
shControl_EGF_bR2	5336	4174	78.2
shControl_EGF_bR3	5336	3467	65
shTMEM_CTRL_bR1	5336	2584	48.4
shTMEM_CTRL_bR2	5336	4185	78.4
shTMEM_CTRL_bR3	5336	3303	61.9
shTMEM_EGF_bR1	5336	2780	52.1
shTMEM_EGF_bR2	5336	3394	63.6
shTMEM_EGF_bR3	5336	3159	59.2

**Figure 1 - Correlation between replicates after data filtering and normalization**

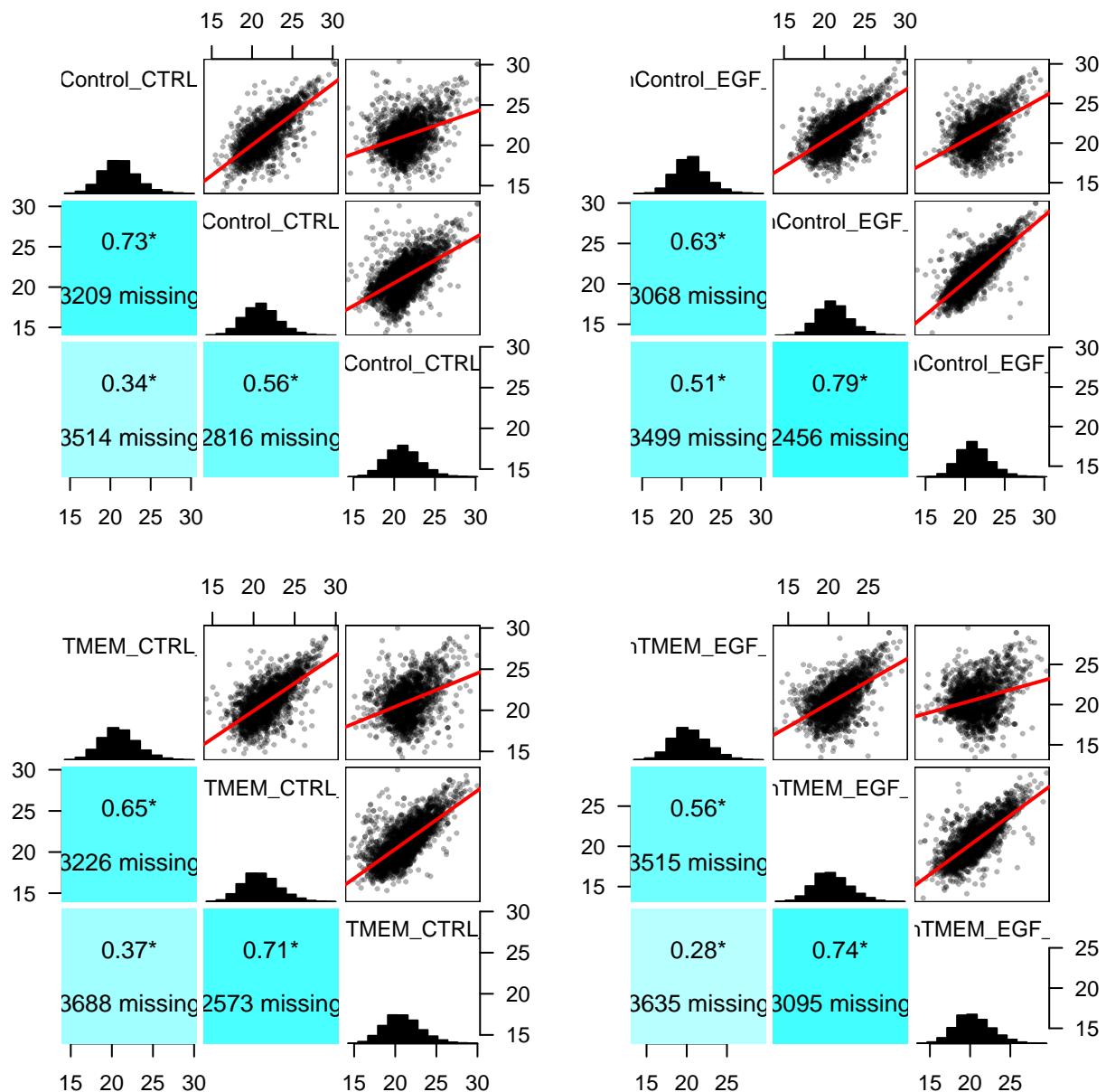
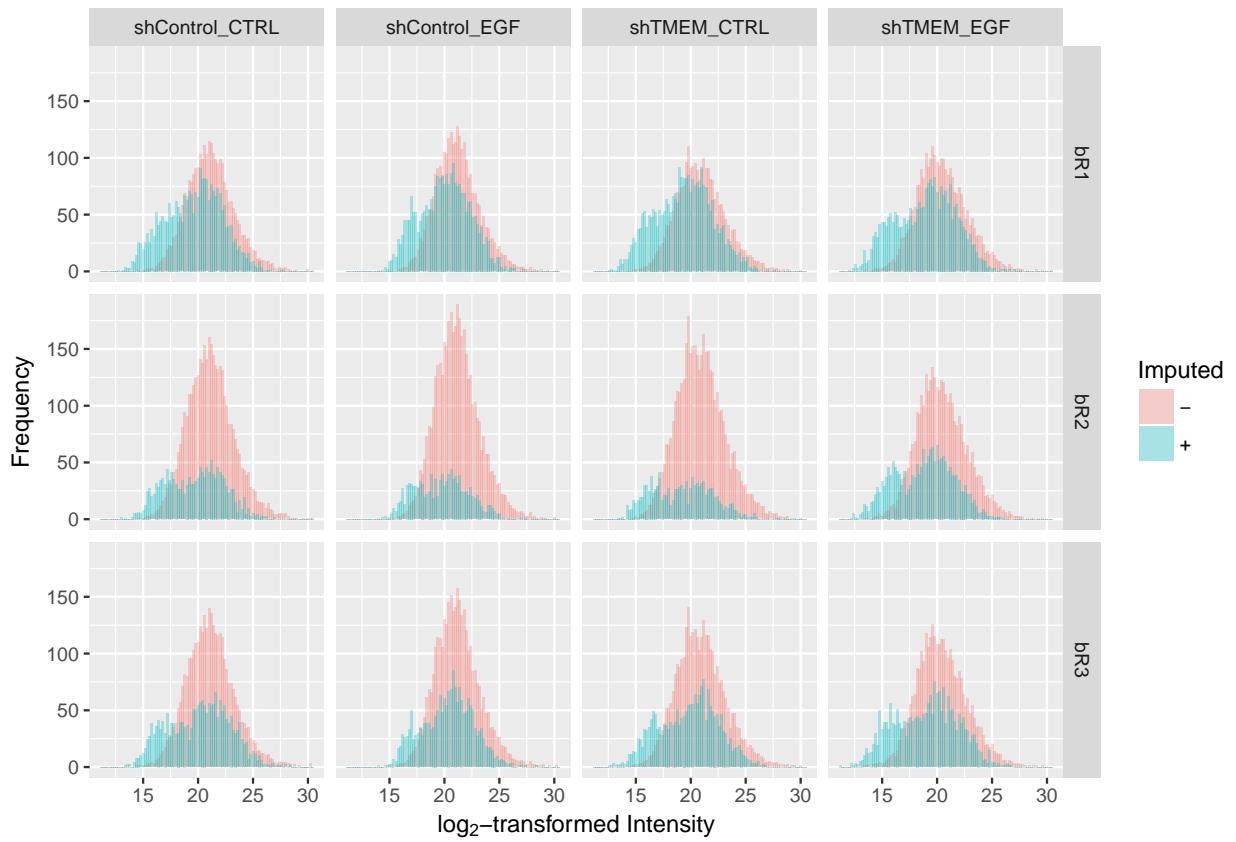


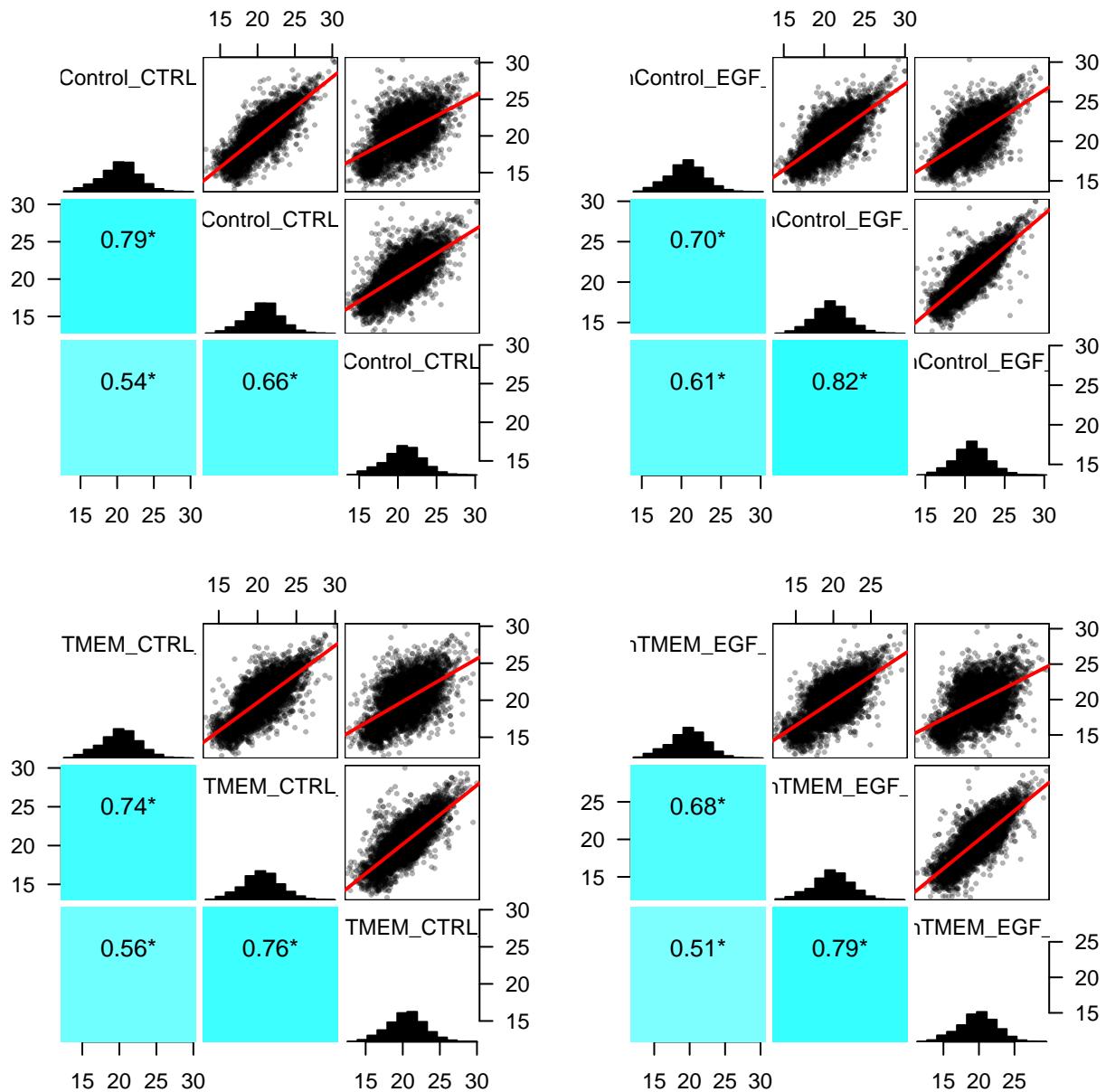
Figure 1 shows the relationship of phosphosite intensities between replicates. Each point in the scatter plot is a phosphosite quantified in both replicates. The line of best fit is drawn in red. The distribution of the normalized,  $\log_2$ -transformed phosphosite intensities are depicted as histograms along the diagonal. The Pearson correlation coefficients and the number of missing values present in the replicates are displayed in the bottom-left corner. These results suggest that a considerable amount of the variability exists between replicates, which is not surprising given the loosely conserved nature of phosphorylation marks and the stochasticity of acquisition method on the mass spectrometer. Additionally, batch effects could play a role since replicate 1 was processed more than a year apart from replicate 2 and 3.

**Figure 2 - Histogram of phosphosite intensities by sample**



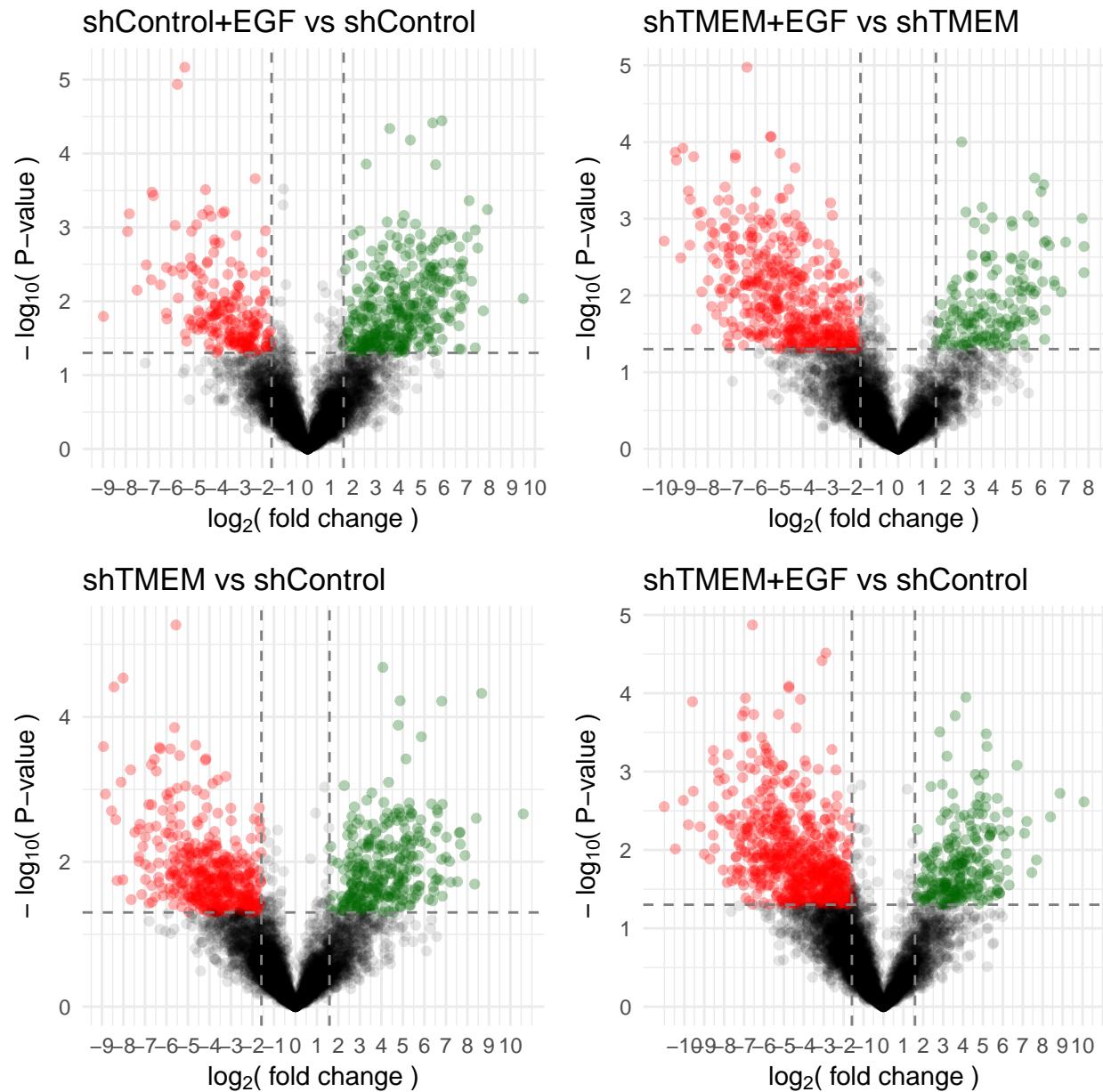
Distribution of the measured phosphosite intensities and the imputed intensities are superimposed. Clearly, replicate 1 suffers more heavily from missing values than the other replicates. Nonetheless, the hybrid imputation method appears to have compensated for this loss by filling in more substitute values along the expected unimodal distribution.

**Figure 3 - Correlation between replicates after data imputation**



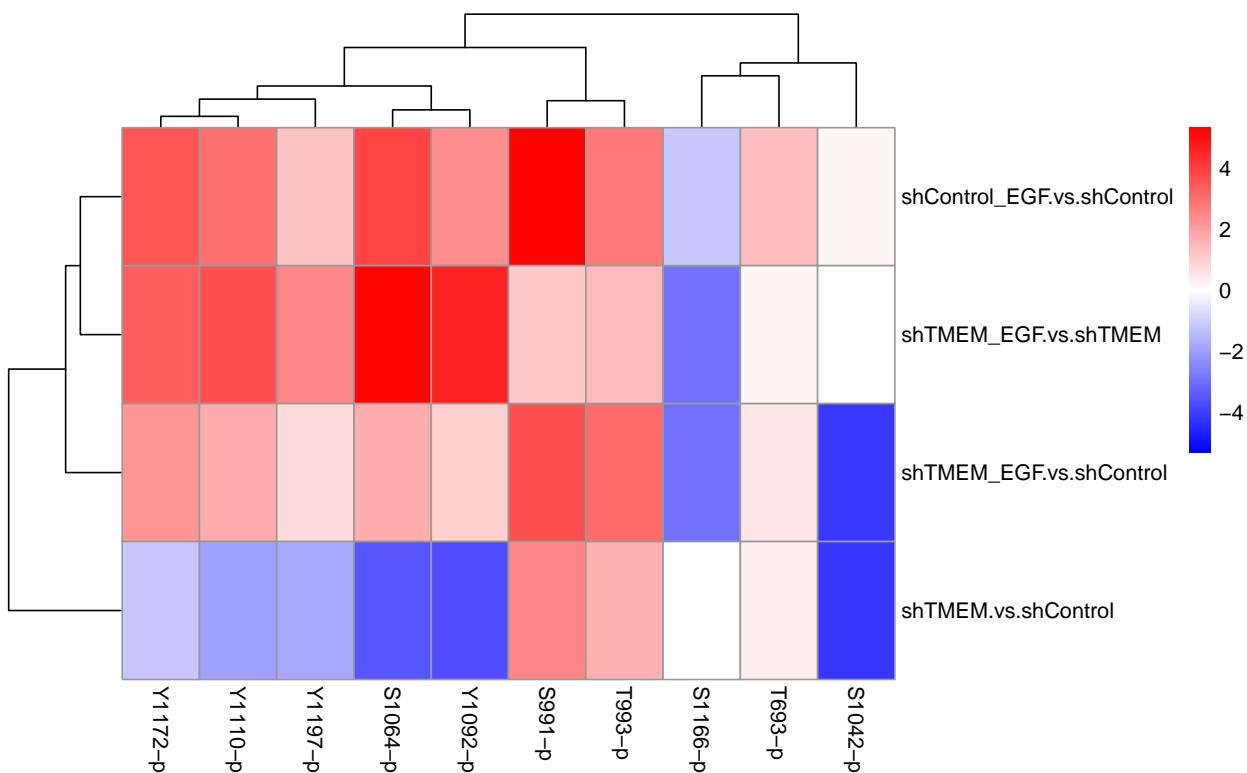
Here is **Figure 1** after data imputation has been performed. The imputation method resolves the missing data problem by inferring values that closely reflect replicate measurements. Therefore, it is perhaps not surprising that the correlation coefficients have improved upon imputation.

Figure 4 - Volcano plots on phosphorylation response to perturbation



Volcano plots show the effect of perturbation on phosphorylation levels and the statistical significance of the change along the x- and y-axis, respectively. Each point represents a phosphosite, and the thresholds are drawn at a fold change of 3 (or 1.58 in  $\log_2$ ) and P-value of 0.05. To tease out any functional patterns among proteins whose phosphosites were differentially regulated, a GO enrichment analysis was performed on a list of proteins associated with the colored points above. A list of unregulated proteins identified in the phosphoproteomic analysis was supplied as a reference set. Unfortunately, no GO terms were significantly enriched.

**Figure 5 - Heatmap of EGFR phosphorylation status under a combination of shTMEM and EGF treatment**



To examine the changes in phosphorylation under different conditions at a finer granularity, phosphosites on EGFR were isolated and visualized on a heatmap. The units are log<sub>2</sub> fold change in intensity. According to PhosphoSitePlus (PTM database focused on curating protein modifications and their functions), phosphorylation on **Y1172** and **Y1197** are associated with induced EGFR activity. Taken together, these results show that EGFR is active upon EGF treatment even in the absence of TMEM16A. In contrast, TMEM16A depletion alone does not appear to induce EGFR activity. This finding, however, does not preclude the possibility that downstream growth signaling is somehow maintained or even increased in the absence of TMEM16A.

**Figure 6a - Kinase Activity Prediction: shControl+EGF vs shControl**

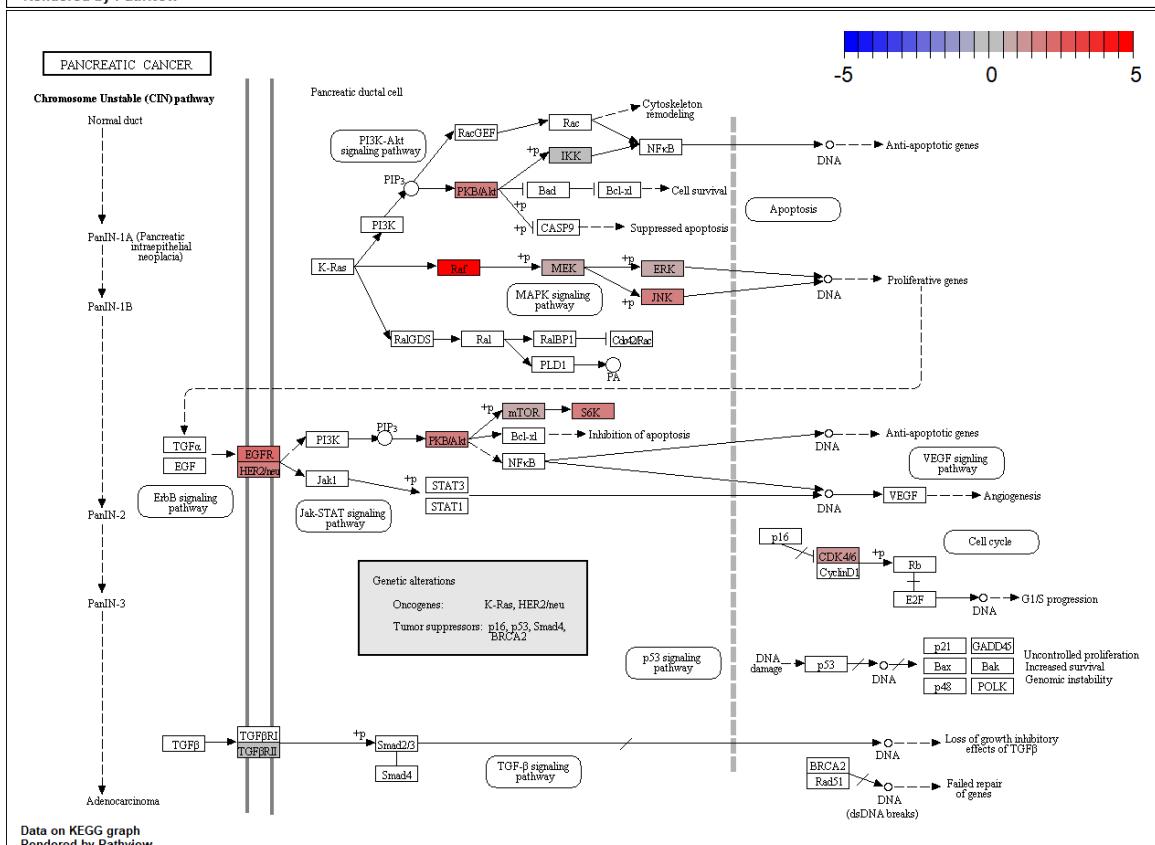
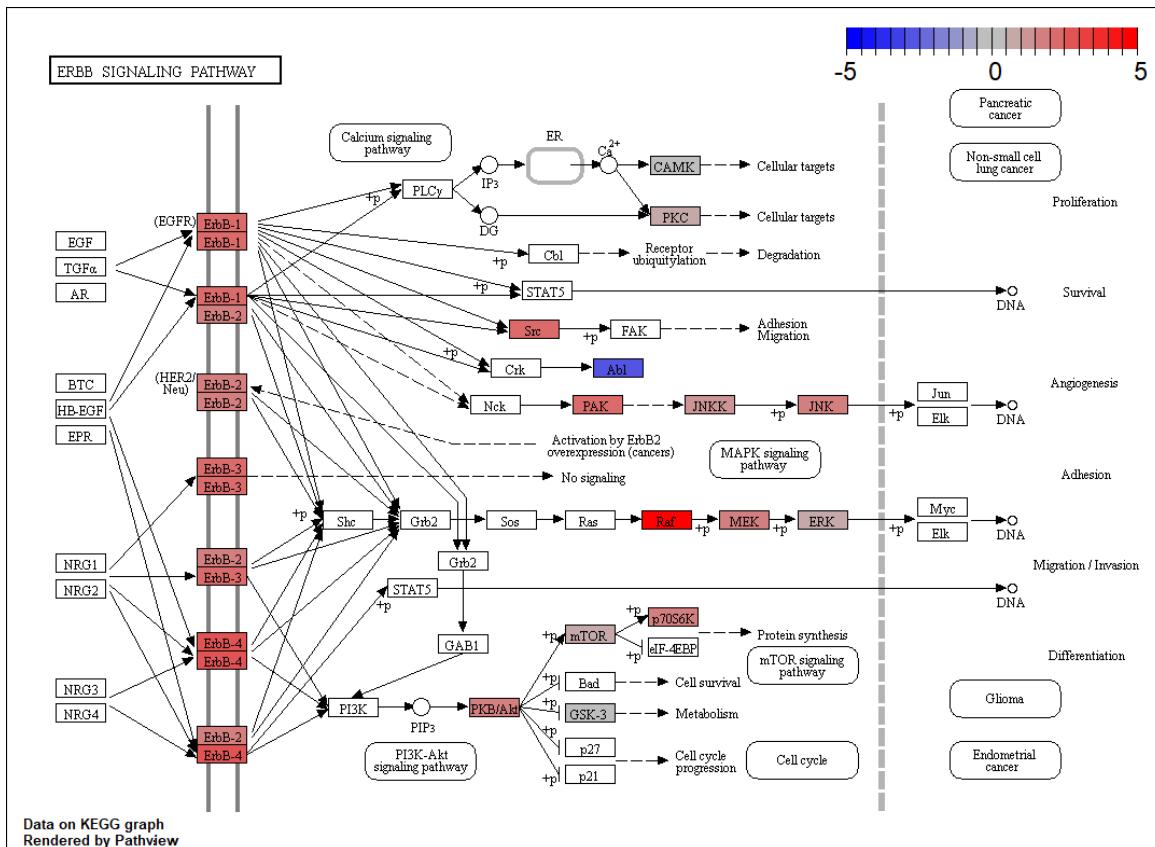


Figure 6b - Kinase Activity Prediction: shTMEM+EGF vs shTMEM

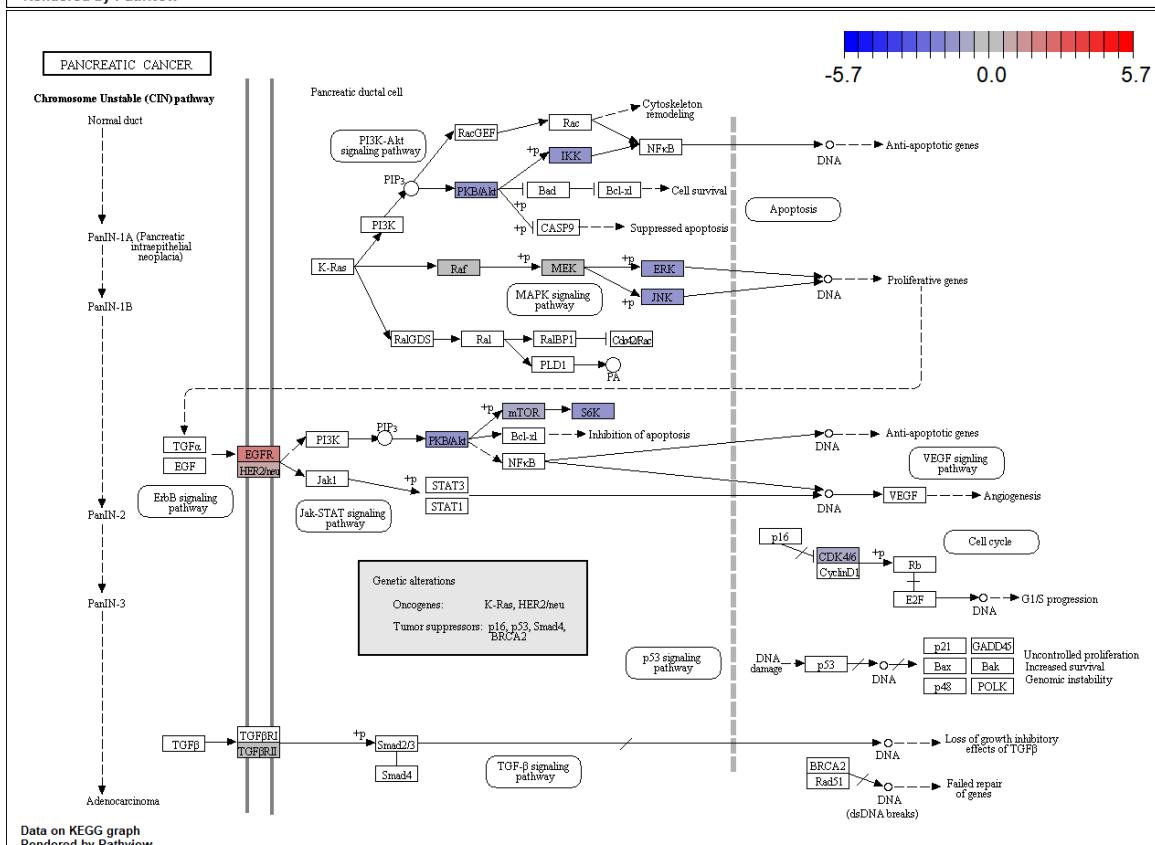
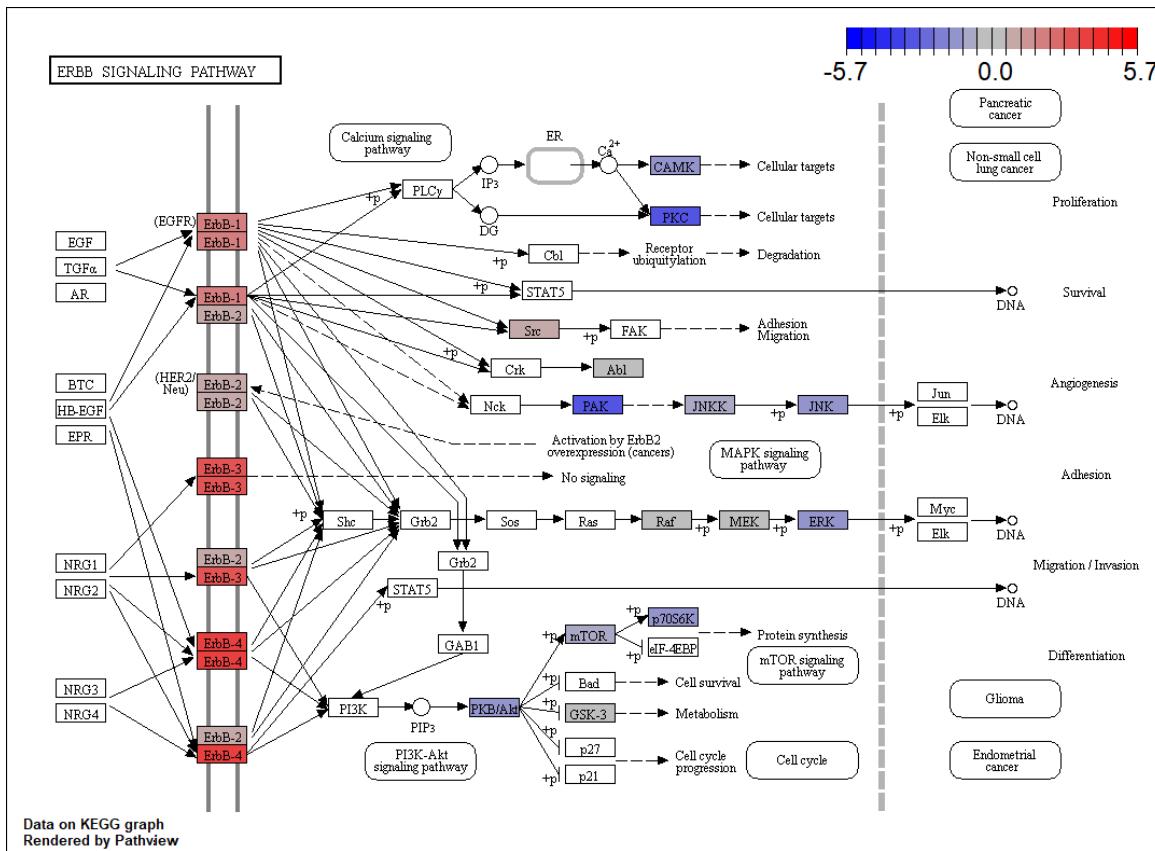


Figure 6c - Kinase Activity Prediction: shTMEM vs shControl

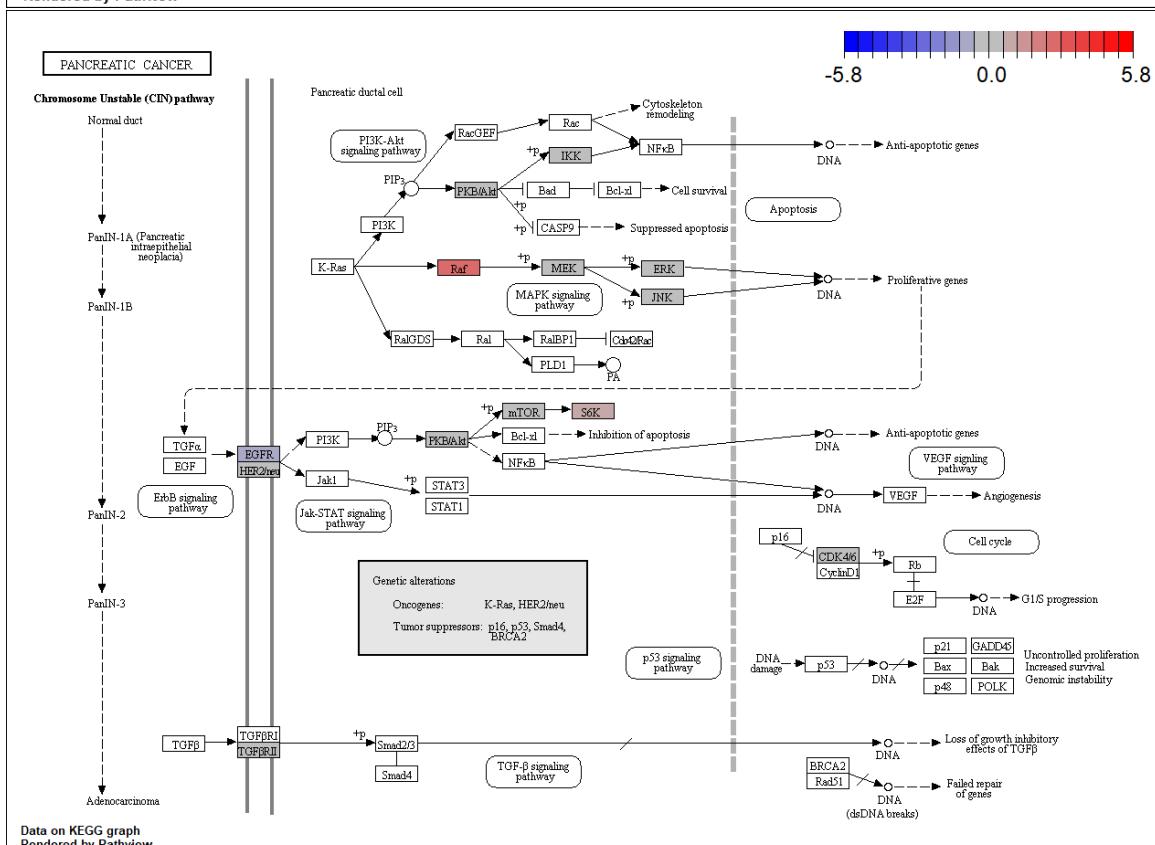
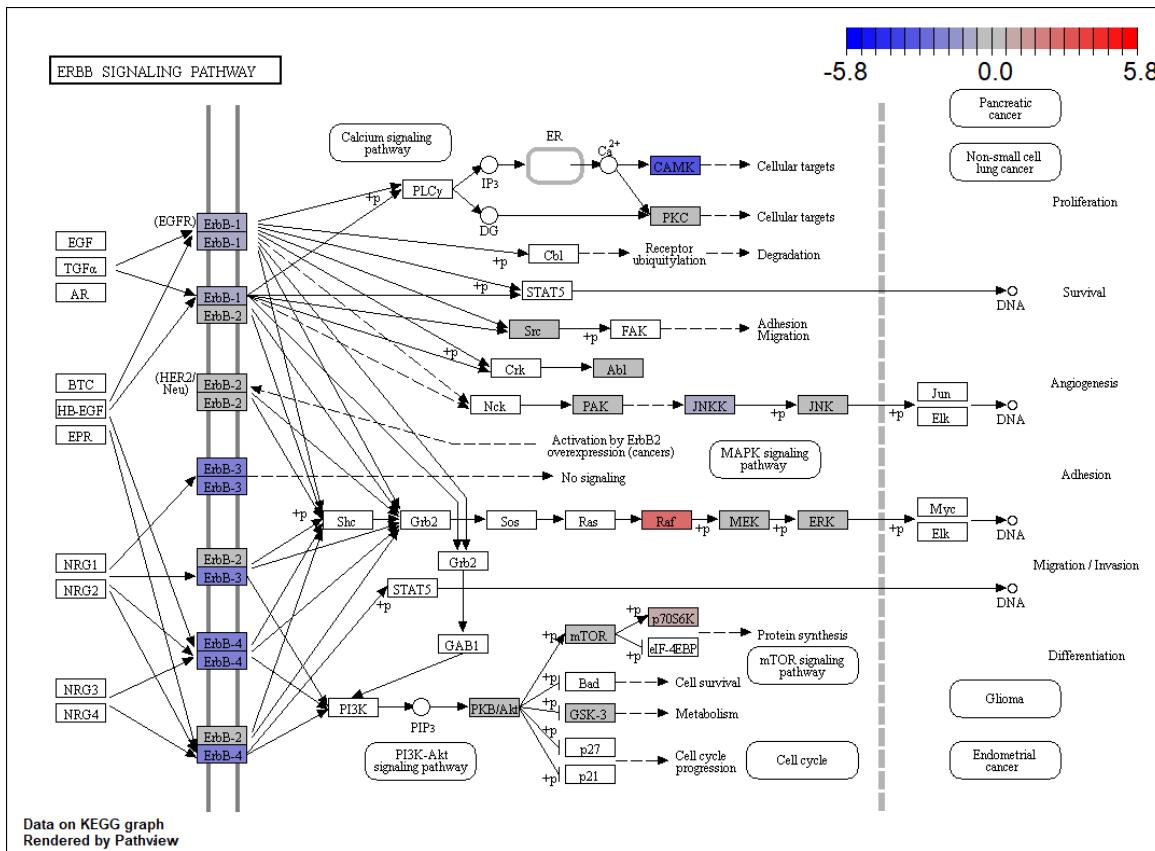
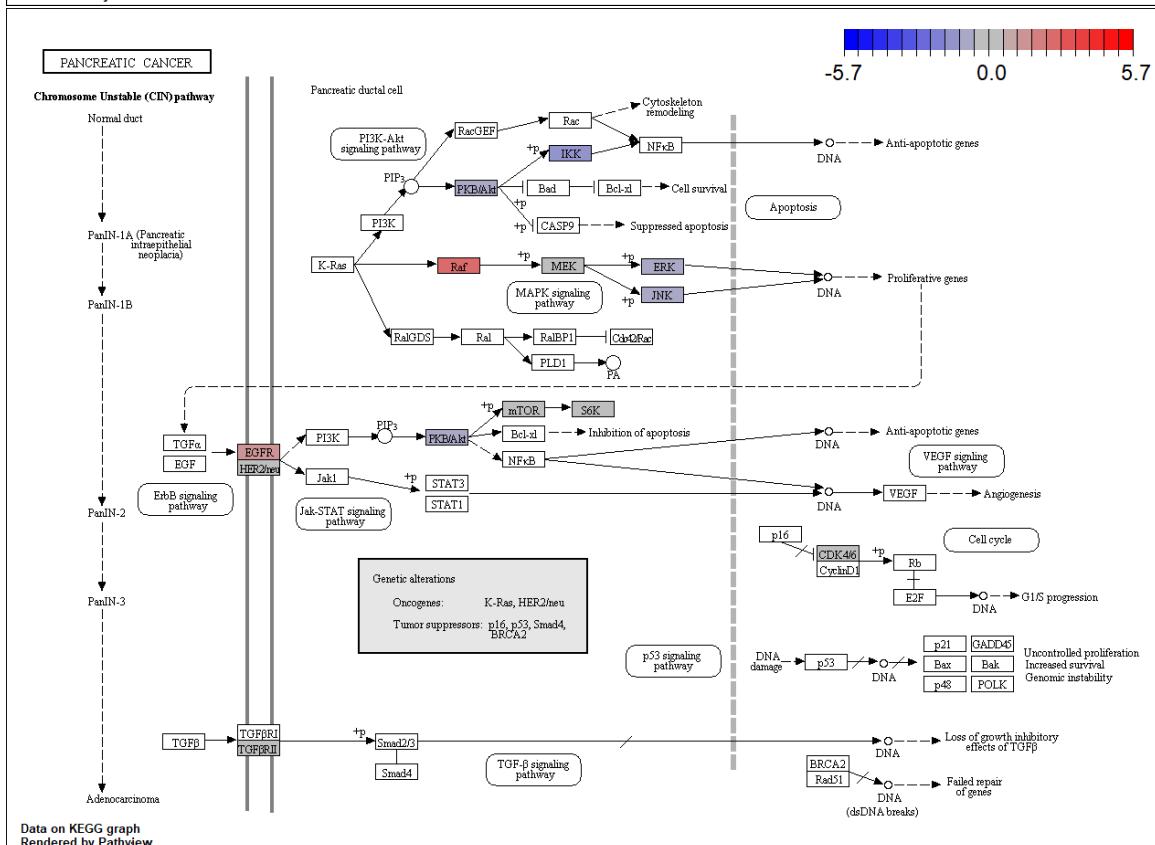
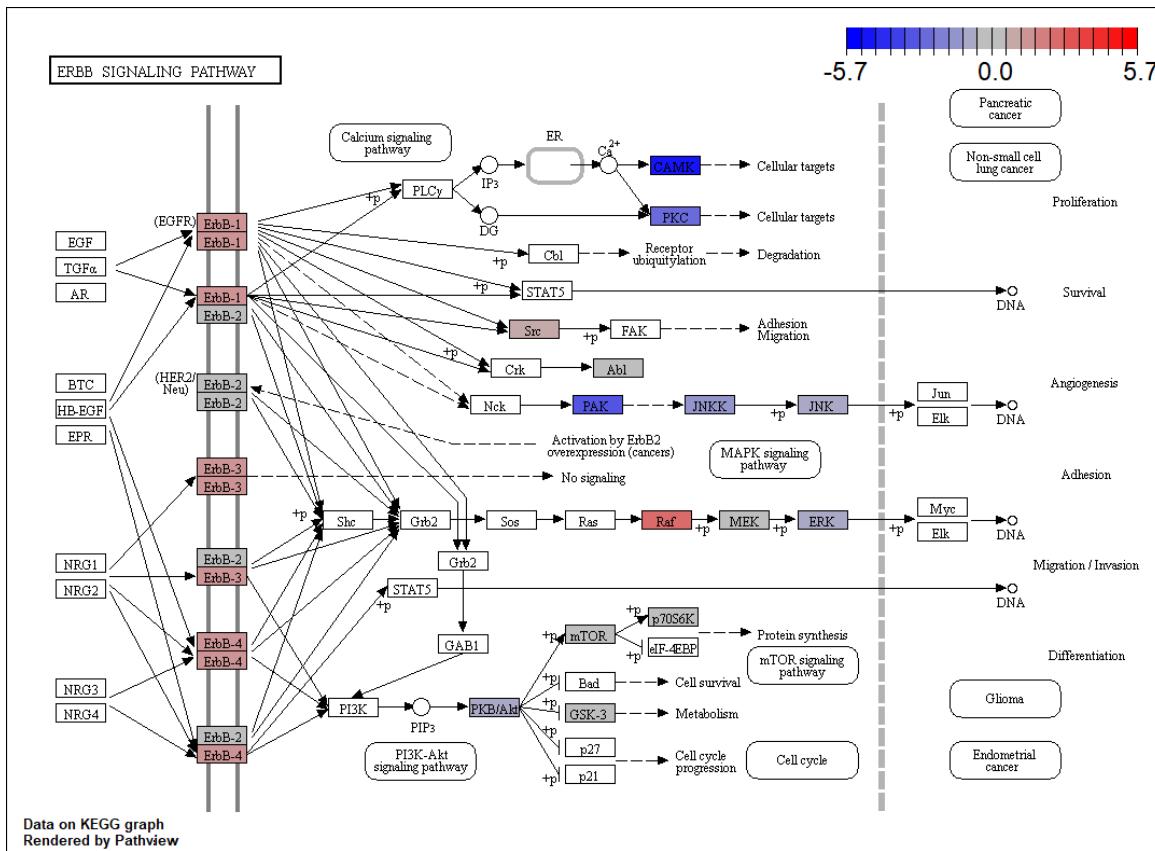


Figure 6d - Kinase Activity Prediction: shTMEM+EGF vs shControl



## Figure 6 Discussion

To investigate the effects of TMEM16A depletion on EGFR signaling, kinase activities were inferred from our phosphoproteomic data. The kinase set enrichment analysis (KSEA) was applied as described by Casado et al. (Science Signaling 2013). The method operates under the assumption that the activity of a kinase can be expressed as a function of the average level of phosphorylation across its substrates. These activities can be compared between two conditions to obtain a measure of relative difference due to perturbation. Once computed, these relative activities were mapped onto the ErbB signaling pathway and the pancreatic cancer model from KEGG. Kinases whose activities could not be inferred due to the absence of associated substrates have a white background. Expression of kinases in AsPc-1 was not considered in this analysis, which means that annotated kinases may not necessarily be present inside the cell.

Four different comparisons were analyzed in **Figures 6a-d**. In **6a**, EGF stimulation turns on EGFR and activates downstream signaling as expected. In **6b**, EGF stimulation in shTMEM16A cells again increases EGFR activity (albeit not as high as control) but interestingly down-regulates growth signaling. In **6c**, TMEM16A knockdown appears to lower basal EGFR activity but in general has minimal effect on downstream signaling activity. In **6d**, the combination of TMEM16A knockdown and EGF treatment weakly activates EGFR but surprisingly translates to a growth-inhibitory response.

In short, the KSEA analysis shows that TMEM16A is not necessary for EGFR activation but is required for growth signal propagation. TMEM16A knockdown appears to have no negative effect on growth signaling, which is consistent with previous findings that TMEM16A depletion has no effect on proliferation in AsPc-1. However, the addition of EGF to shTMEM16A cells unexpectedly reduces downstream growth signaling to a point below the basal level in control cells.