Humphrey et al

Supplementary Figure Legends

Supplementary Figure 1: Molecular subclassification and bioinformatic workflow.

A, Distribution of phosphoprotein types according to A, GO Cellular Compartment (GOCC) and B, GO Molecular Function (GOMF). Each observed phosphoprotein was assigned protein categories from the GO Database via the mapping tools available in Perseus. The total relative phosphorylation of unique proteins in each category (calculated by the addition of the TIC of sites from these phosphoproteins), as a fraction of the total phosphorylation (total TIC of all sites of all phosphoproteins), are represented by the columns. C, Bioinformatic workflow applied to ATCC and TKCC phosphotyrosine datasets. D, Coefficient of variation (CV) calculations were used to assess for any correction of methodological variance between biological replicates provided by normalization using exogenous pTyr peptide internal standards. For each ATCC cohort cell line, serial CV values were calculated between biological replicates for each quantified pTyr site using either the un-normalized or normalized TICs. The CV values across all sites of each cell line were then compared through boxplots.

Supplementary Figure 2: Mutation analysis and protein expression across the ATCC panel.

A, Identification and characterisation of KRAS somatic mutations within ATCC cohort cell lines using the OncoCarta Panels v1.0, v2.0 and v3.0 (Sequenom). Colours of the bars within the chart refer to the key denoting the somatic mutation identified. B, K-ras, E-cadherin and vimentin protein expression in selected ATCC cohort cell lines (subtype 1 is green, 2 is pink and 3 is blue), determined by Western blotting. Grey bar: unclassified cell line. C, hnRNP

protein expression blotting in selected ATCC cohort cell lines (subtype 1 is green, 2 is pink and 3 is blue).

Supplementary Figure 3: Characterization of the TKCC phosphotyrosine dataset.

A, TKCC cell line biological replicate technical reproducibility according to Pearsons correlation (R). B, A non-redundant list of the most abundant tyrosine phosphopeptides in each PDAC cell line cohort was generated by identification of the 25 most abundant phosphopeptides in each cell line. These lists were overlayed and compared. C, Quantified, class I sites (> 0.75 localization probability) from all RAW files are categorized as either known or novel, when compared to the PhosphoSitePlus database using Perseus. These sites were appropriately FDR controlled (1%) within a single MaxQuant search, with datasets set as separate fractions.