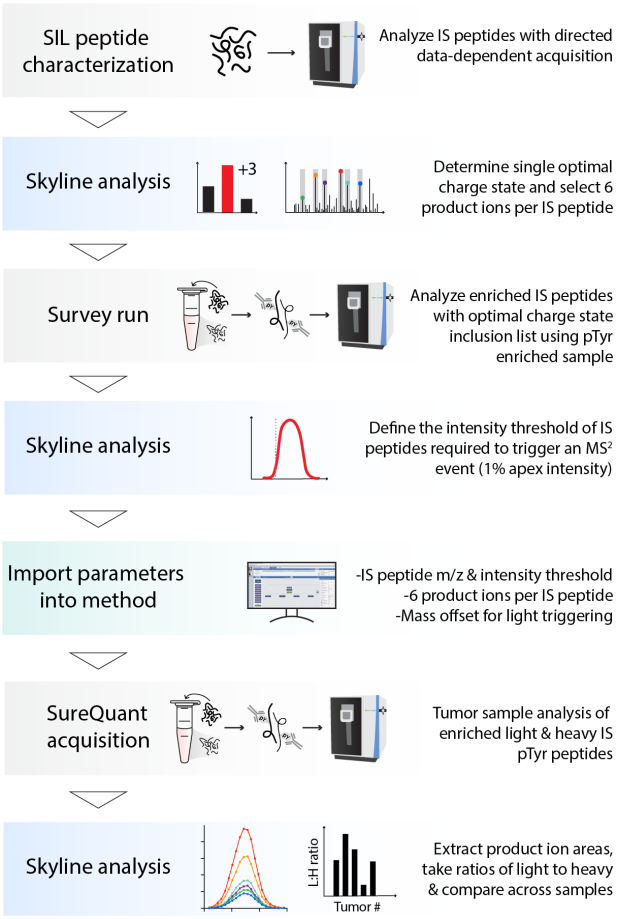
# High-density, targeted monitoring of tyrosine phosphorylation reveals activated signaling networks in human tumors

Stopfer et al.

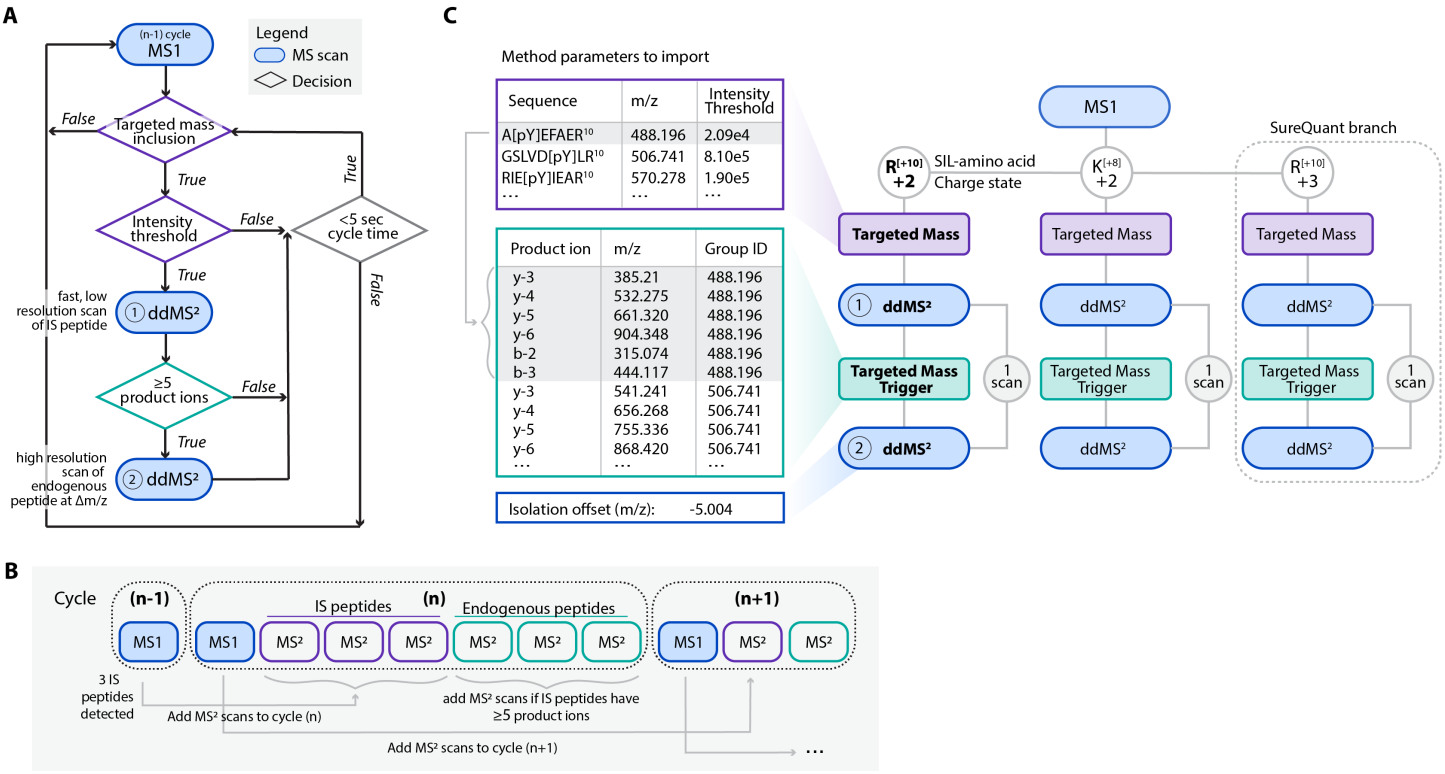
This file includes:

- Supplementary Figures and legends
- Supplementary Methods
- Supplementary Table legends

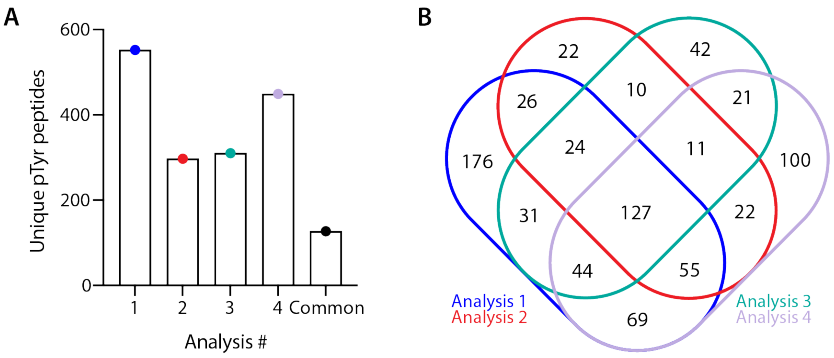
# SUPPLEMENTARY FIGURES



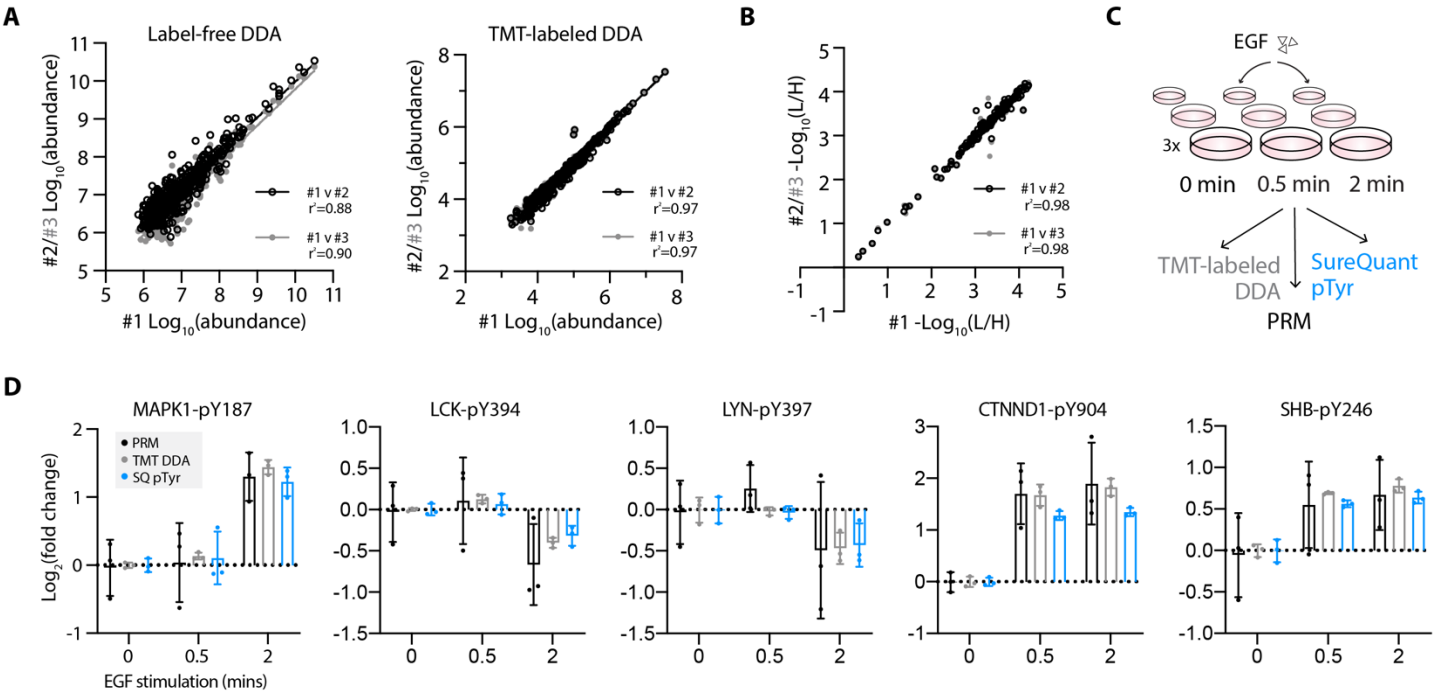

**Supplementary Figure S1:** Flowchart of SureQuant pTyr workflow. SureQuant pTyr pipeline for method building, data acquisition, and data analysis.



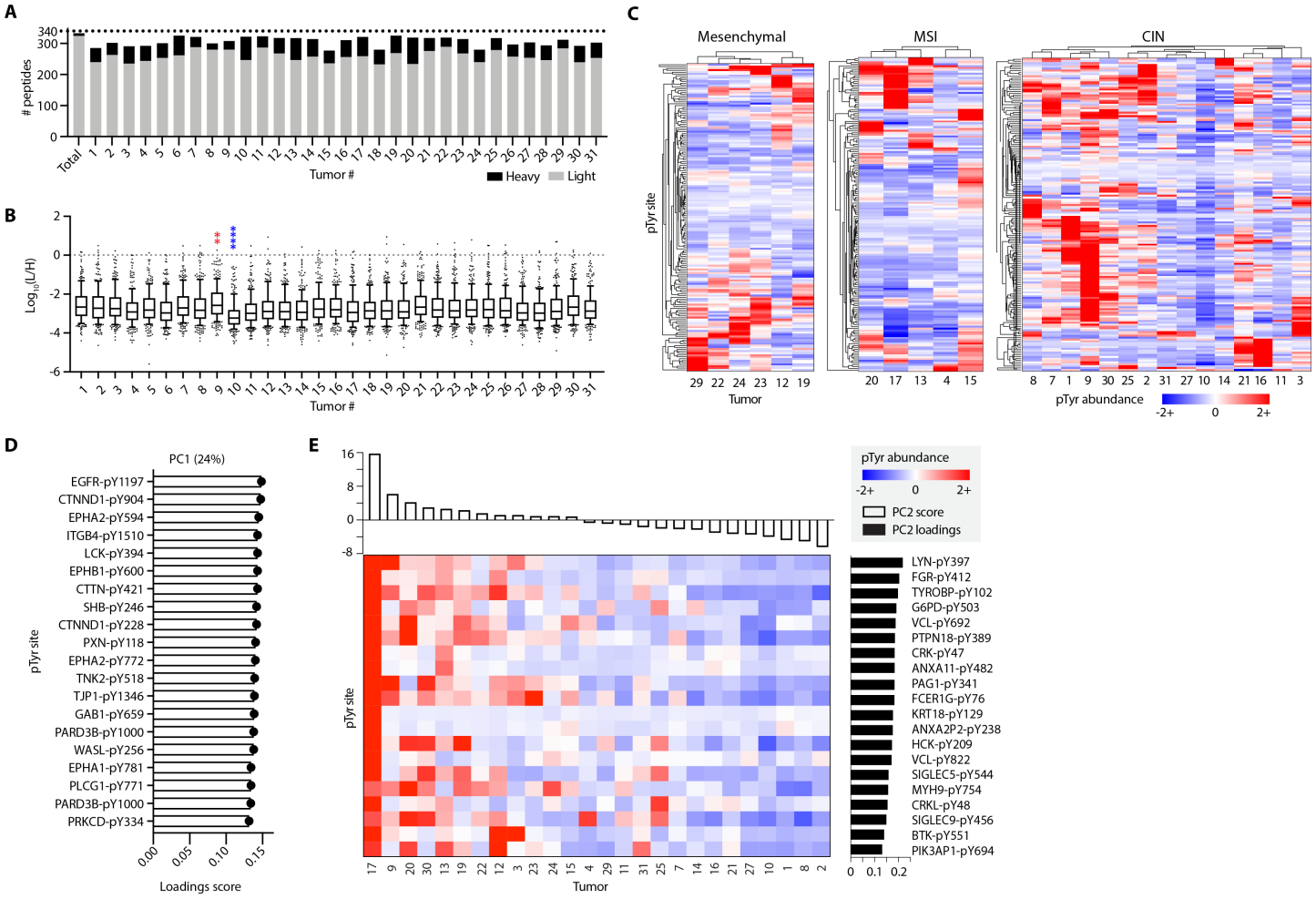

**Supplementary Figure S2.** SureQuant pTyr acquisition method framework and parameters. **A,** Decision tree for SureQuant acquisition method framework. **B,** Scan sequence for SureQuant analyses on the Exploris 480 MS. For example, if three internal standard (IS) peptides are detected in the MS1 scan of cycle (n-1), an $MS^2$ scan of each IS peptide is added to the scan sequence in cycle (n). If each IS $MS^2$ scan contains ≥5 pre-defined product ions, $MS^2$ scans of each endogenous peptide at the defined m/z offset are added to the scan sequence in cycle (n). This scan structure repeats for each cycle. **C,** Method scan structure and parameters to import for SureQuant acquisition illustrated for three branches, however SureQuant pTyr uses ten branches. [pY] denotes the residue position with pTyr modification, and $R^{10}$ and $K^8$ denote stable isotope labeled (SIL) arginine (+10) and lysine (+8), respectively.



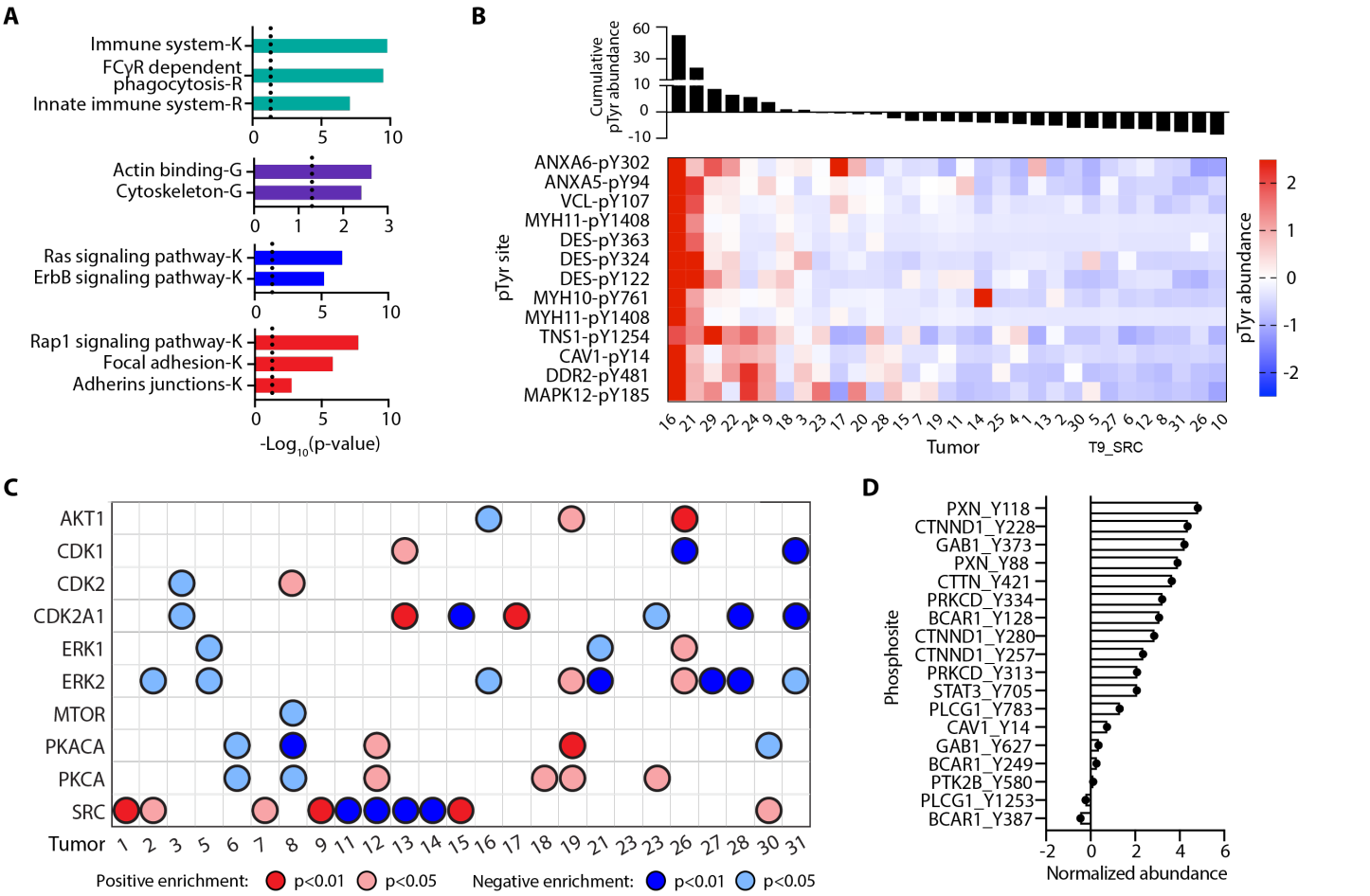

**Supplementary Figure S3.** Unique peptides in discovery analyses. **A,** Number of unique pTyr peptides in each discovery analysis. **B,** Overlap in identified peptides between the four analyses.



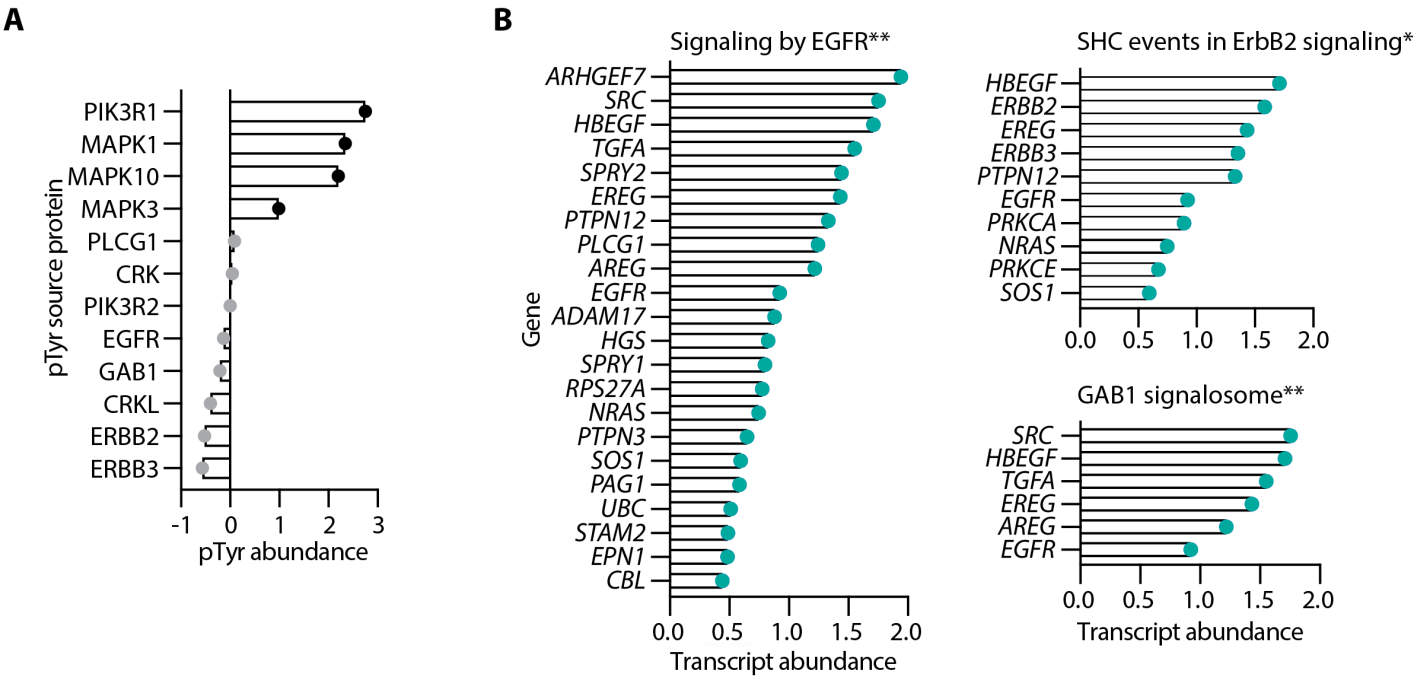

**Supplementary Figure S4.** Benchmarking SureQuant pTyr quantitative accuracy**. A,** Correlation of peptide abundances for three label-free DDA analyses (left) and one multiplexed, TMT-labeled DDA (right) analysis. R-squared values were 0.88 and 0.90 for the label-free analysis, and $r^2$=0.97 for the TMT-labeled analysis. **B,** Correlation of peptide L/H ratios between three technical replicate tumor samples, $r^2$=0.98.
**C,** Experimental setup for quantitative comparison between TMT-labeled DDA, PRM, and SureQuant pTyr methods. **D,** $Log_2$ fold change values, relative to the mean peptide abundance at the 0-minute timepoint, of pTyr peptides for three data acquisition methods: PRM (black), TMT-labeled DDA (grey) and SureQuant pTyr (blue). No significant differences in quantitation between PRM and TMT-DDA vs. SQ pTyr were reported (Dunnett's multiple comparisons test). Each sample includes n=3 biological replicates, error bars represent the standard deviation.



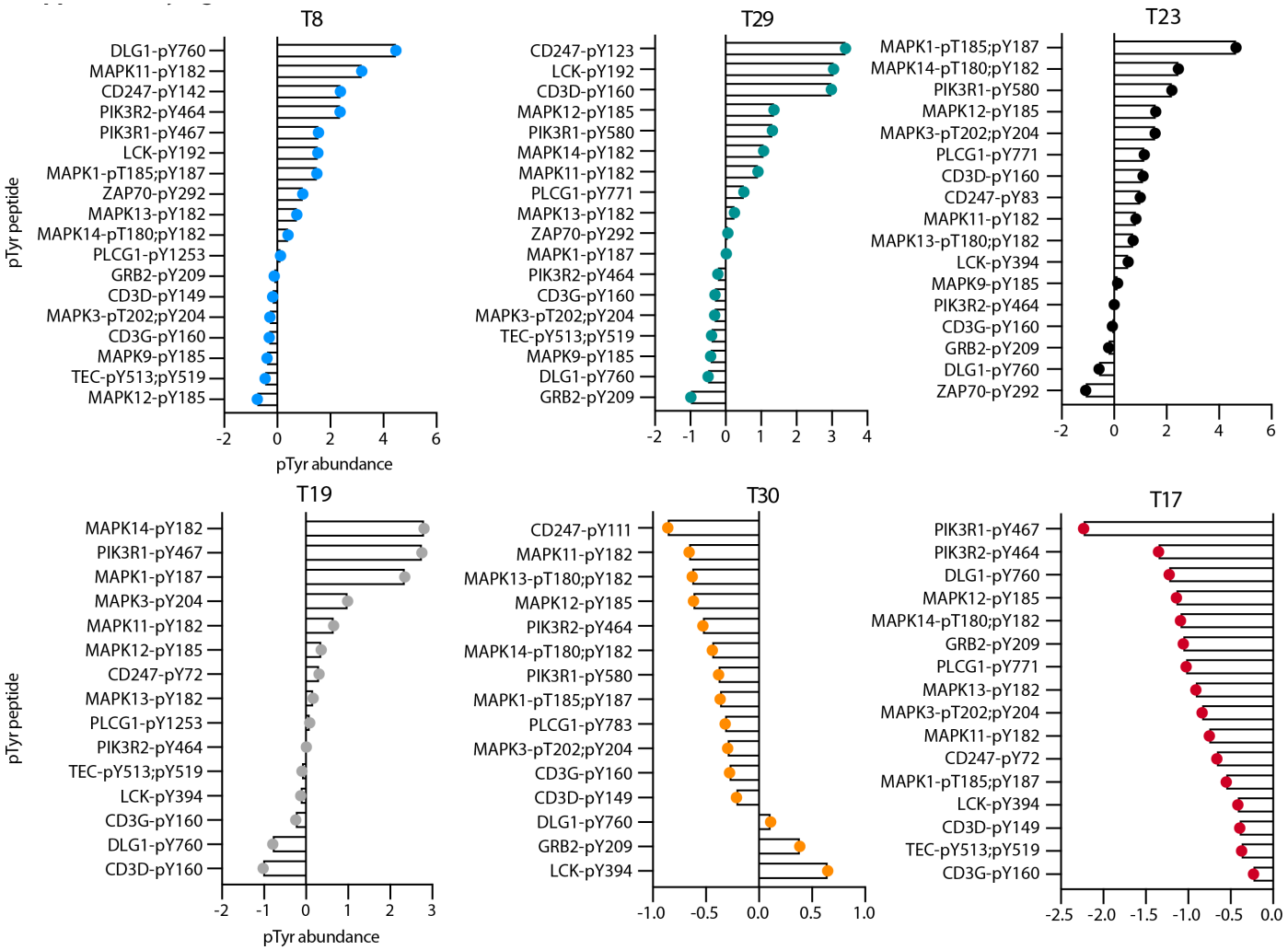

**Supplementary Figure S5.** pTyr sites measured in human colorectal tumors. **A,** Number of unique heavy (black) and light (grey) peptides identified in each tumor. **B,** Distributions of light to heavy signal ratios (L/H) for each tumor. Data is displayed as a box and whiskers plot, where the box describes the interquartile range and the whiskers define the 10-90 percentile of data. * Indicates significantly increased (red) or decrease (blue) from the mean distribution using Dunnett's multiple comparison test for significance. **p<0.01, ****=p<0.0001. **C,** Hierarchical clustering of pTyr peptides within each unified multi-omics subtype defined by Vasaikar et al. Distance metrics for clustering of peptides and tumors, respectively, are Euclidean and correlation. **D,** Peptides with the top 20 loadings scores for PC1, pY denotes the residue position with pTyr modification. **E,** Top 20 loadings scores of the peptides derived from unique proteins for PC2, ranked from highest to lowest PC2 score with corresponding pTyr abundance levels (z- score normalized L/H) across tumors.




**Supplementary Figure S6.** Correlation analysis identifies tumors with significant pathway enrichment. **A,** Selected significantly enriched pathways identified in Fig. 4. K=Kegg pathway, G=Gene Ontology term, R=reactome pathway. Significance values are FDR-adjusted, and a cutoff of $p < 0.05$ was used for filtering. **B,** pTyr peptides from cluster 2, rank ordered from highest to lowest cumulative pTyr abundance, where abundance values are z-score normalized light to heavy signal ratios. **C,** Significantly enriched kinase-substrate interactions within tumors using pTyr and pSer/pThr datasets. Significance (p-value) and directionality indicated by color, FDR q-value < 0.25 for all enrichment analyses. Tumors without any significant enrichment are not shown. **D,** SRC kinase-substrate enrichment signature for T9. Normalized abundance corresponds to z-score normalized light to heavy signal ratios.




**Supplementary Figure S7**. ErbB signaling pathway enrichment analysis. **A,** Tumor 19 pTyr abundance of peptides in the ErbB signaling pathway. All peptides are annotated by source protein, with peptides in the enrichment core colored in black. **B,** Tumor 9 positively enriched reactome pathways with corresponding transcript abundance levels (z-score normalized). *= $p<0.05$. **=$p<0.01$ FDR q-value < 0.05 for all.




**Supplementary Figure S8.** Tumor-specific pTyr signatures of T cell signaling peptides. pTyr abundance (z-score normalized light to heavy signal ratios) of T cell signaling peptides in tumors with significant pathway enrichment. Peptides are annotated by source protein.

# SUPPLEMENTARY METHODS

## Internal standard peptide characterization and survey run analysis

All analyses were processed using Skyline software (version 20.1, daily build) (27). The IS peptides properly detected in initial directed DDA analysis, *i.e.*, those for which at least one precursor ion yielded several $MS^2$ scans including at least 6 theoretical y- or b-type fragment ions, were retained for SureQuant method development. For peptides detected under multiple charge states, only the precursor ion yielding the highest signal response was retained. For each peptide, 6 associated optimal fragments ions were selected for psudo-spectral matching (typically the most intense ones showing sufficient specificity, *i.e.*, without neutral loss or low m/z value). Individual intensity thresholds for each IS peptide was set to 1% of the precursor MS1 intensity value at the apex of its chromatographic profile in the survey run analysis.

## Mass spectrometry data acquisition parameters

*Survey analysis*
The mass spectrometry parameters used for these preliminary analyses were as follows: spray voltage: 1.9kV, no sheath or auxiliary gas flow, heated capillary temperature: 280°C. DDA analyzes collected full-scan mass spectra with m/z range 300-1200, AGC target value: 1000% (1e7), maximum injection time (IT): 50 ms, resolution: 120,000. For every scan, the top 40 most intense ions on the inclusion list (if above a 1e5 intensity threshold) were isolated [isolation width of 1.0 m/z] and fragmented [normalized collision energy (nCE): 28%] by higher energy collisional dissociation (HCD), scan range: 100-1700 m/z, maximum IT: 10 ms, AGC target value: 1000% (1e6), resolution: 7,500.

*Multiplexed discovery analysis of colorectal tumors*
Peptides were resuspended in 10 µL 0.1% acetic acid and onto an analytical capillary column with an integrated electrospray tip (~1 µm orifice) prepared in house ((50 µm ID × 12 cm with 5 µm C18 beads (YMC gel, ODS-AQ, 12 nm, S-5 µm, AQ12S05)). Peptides were eluted using a 140 minute gradient with 13-42% buffer B (70% Acetonitrile, 0.2M acetic acid) from 10-105 minutes and 42-60% buffer B from 105-115 minutes, 60-100% B from 115-122 minutes, and 100-0% B from 128-130 minutes at a flow rate of 0.2 mL/min for a flow split of approximately 10,000:1. Peptides were analyzed using a Thermo Fisher Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer, and data was acquired using Thermo Fisher Scientific Xcalibur version 2.5.0.2042. Standard mass spectrometry parameters were as follows: spray voltage, 2 kV; no sheath or auxiliary gas flow; heated capillary temperature, 250 °C.

The Q Exactive Plus was operated in data-dependent acquisition (DDA) mode with the following scan settings: Full-scan mass spectrometry spectra (mass/charge ratio (m/z), 350 to 2,000; resolution, 70,000) were detected in the Orbitrap analyzer after accumulation of ions at $3e^6$ target value with a maximum injection time (IT) of 50 ms. For every full scan, the top 15 most intense ions were isolated (isolation width of 0.8 m/z) and fragmented (collision energy (nCE): 33%) by higher energy collisional dissociation (HCD) with a maximum IT 350 ms, AGC target $1e^5$, and 35,000 resolution. Unassigned and +1 charge states were excluded, and dynamic exclusion was set to 30 seconds.

Crude peptide analysis was performed on a Q Exactive Plus mass spectrometer to correct for variation in peptide loadings across TMT channels. Approximately 30 ng of the supernatant from pTyr IP was loaded onto an in-house packed precolumn (100 um ID x 10 cm) packed with 10 µm C18 beads (YMC gel, ODS-A, AA12S11) connected in series to an analytical column (as previously described) and analyzed with a 75 min LC gradient [0-30% B from 0-40 minutes, 30-60% B from 40-50 minutes, 60-100% B from 50-55 minutes, and 100-0% B from 60-65 minutes]. MS1 scans were performed with m/z range: 350-2000; resolution: 70,000; AGC target: $3x10^6$; maximum IT: 50 ms. The top 10 abundant ions were isolated (isolation width 0.8 m/z) and fragmented (nCE = 33%) with 35,000 resolution, maximum IT 350 ms, AGC target $1e^5$. Unassigned and +1 charge states were excluded, and dynamic exclusion was set to 30 seconds.

*Label-free DDA analysis of A549 cells*
Three label-free replicate samples of A549 cells stimulated with 5 nM EGF were analyzed using an Orbitrap Exploris 480 mass spectrometer (Thermo Scientific) coupled with an UltiMate 3000 RSLC Nano LC system (Dionex), Nanospray Flex ion source (Thermo Scientific), and column oven heater (Sonation). A 10 µL injection volume of sample was directly loaded onto a 25 cm Aurora Series emitter column (IonOpticks) with a column oven temperature of 40°C. Peptides were eluted at a flow rate of 400 nL/min across a linear gradient consisting of 0.1% formic acid (buffer A) and 80% acetonitrile in 0.1% formic acid (buffer B). The gradient is as follows: 2-8% B from 2-15 mins, 8-35% B from 15-95 mins, 35-60% B from 95-105 mins, 60-100% B from 105-110 mins, and 100-2% B from 112-114 mins. Standard MS parameters were as follows: 1.5 kV; no sheath or auxiliary gas flow; heated capillary temperature, 275 °C.

The Exploris 480 was operated in DDA mode with the following scan settings: full-scan mass spectra were collected with m/z range 300-1600, AGC target value: 300% ($3e^6$), maximum IT: auto. For every MS1 scan, peptides were isolated during a 3 s cycle time [isolation width 0.4 m/z] and fragmented [nCE 30%] by HCD with 60,000 resolution, 150 ms maximum IT, AGC target 100% ($1e^5$). Charge states < 2 and > 6 were excluded., and peptides were excluded from $MS^2$ analysis for 30 s duration after fragmentation n=2 times within 20 s duration.

*TMT-labeled DDA analysis of A549 cells stimulated with EGF*
Multiplexed, A549 cells stimulated with 5 nM EGF for 0, 0.5, or 2 mins were analyzed using the Exploris 480 and UltiMate 3000 RSLC Nano LC system as previously described, using the SureQuant gradient. Standard MS parameters were as follows: 2.5 kV; no sheath or auxiliary gas flow; heated capillary temperature, 275 °C.

The Exploris 480 was operated in DDA mode with the following scan settings: full-scan mass spectra were collected with m/z range 350-1600, AGC target value: 300% ($3e^6$), maximum IT: auto. For every MS1 scan, peptides were isolated during a 3 s cycle time [isolation width 0.4 m/z] and fragmented [nCE 33%] by HCD with 60,000 resolution, 150 ms maximum IT, AGC target 100% ($1e^5$). Exclusion parameters are as described with label-free DDA analysis of A549 cells.

Crude peptide analysis of the immunoprecipitation supernatant was performed on a Q Exactive Plus mass spectrometer to correct for variation in peptide loadings across TMT channels. Approximately 30 ng of the supernatant from pTyr IP was loaded onto an in-house packed precolumn (100 um ID x 10 cm) packed with 10 µm C18 beads (YMC gel, ODS-A, AA12S11) connected in series to an analytical column(as previously described) and analyzed with a 70 min LC gradient, as previously described. MS1 scans were performed with m/z range: 350-2000; resolution: 70,000; AGC target: $3x10^6$; maximum IT: 50 ms. The top 10 abundant ions were isolated (isolation width 0.4 m/z) and fragmented (nCE = 33%) with 70,000 resolution, maximum IT 150 ms, AGC target $1e^5$. Unassigned, >8, and +1 charge states were excluded, and dynamic exclusion was set to 30 seconds.

*Label-free PRM analysis of A549 cells stimulated with EGF*
A549 cells stimulated with 5 nM EGF for 0, 0.5, or 2 mins with analyzed by PRM using the Exploris 480 MS and UltiMate 3000 RSLC Nano LC system as previously described, using the SureQuant gradient. Standard MS parameters were as follows: 1.5 kV; no sheath or auxiliary gas flow; heated capillary temperature, 275 °C.

The Exploris 480 was operated using a targeted inclusion mass list (Supplementary Table S4), with an MS1 resolution of 60,000, scan range 3350-1600, AGC target value: 300% ($3e^6$). During a 3 s cycle time, targeted peptides were isolated [isolation width 0.4 m/z] and fragmented [nCE 30%] by HCD with 60,000 resolution, 150 ms maximum IT, AGC target 100% ($1e^5$).

Additional search space, filtering, and analysis details for MS analyses.

*Multiplexed discovery analysis of colorectal tumors*
Mass spectra were processed with Proteome Discoverer version 2.2 (PD 2.2) (Thermo Fisher) and searched against the human SwissProt database (2018_4) using Mascot version 2.4 (Matrix Science). Spectra were searched using the following parameters: enzyme: trypsin, maximum missed cleavages: 2, precursor mass tolerance: 10 ppm, fragment mass tolerance: 20 mmu. Static modifications included TMT-10-labeled lysine and N-terminal residues, as well as cysteine carbamidomethylation. Dynamic modifications included methionine oxidation, and tyrosine, serine, and threonine phosphorylation.

Peptide spectrum matches (PSMs) with phosphorylated tyrosine residues and no reporter ion missing values were filtered according to search engine rank = 1, ion score ≥ 20. Reporter ion intensities were summed across matching PSMs. Phosphotyrosine peptide reporter ion areas were corrected for variations in sample loading within each analysis using the median of peptide ratios in the crude peptide analysis for each channel relative to channel. Quantitation is represented as the fold change between the tumor and normal phosphorylation levels.

*TMT-labeled DDA analysis of A549 cells stimulated with EGF*
Mass spectra were processed with PD 2.5 and searched against the human SwissProt database (2020_4) using Mascot version 2.4. Spectra were searched and analyzed using the same parameters described in the multiplexed discovery analysis of colorectal tumors.
To evaluate the quantitative reproducibility of the SureQuant method, the three 2-minute EGF stimulated samples were used (Supplementary Figure S1A).

*Label-free DDA analysis of A549 cells*
Mass spectra were processed with PD 2.5 and searched against the human SwissProt database (2020_4) using Mascot version 2.4. Spectra were searched using the following parameters: enzyme: trypsin, maximum missed cleavages: 2, precursor mass tolerance: 10 ppm, fragment mass tolerance: 20 mmu. Cysteine carbamidomethylation was set as a static modification, and dynamic modifications included methionine oxidation, and tyrosine, serine, and threonine phosphorylation. MS1 integrated peak area quantitation was performed using the minora feature detector in PD 2.5 with match between runs enabled and filtered for ion score ≥ 20, search engine rank = 1. Only peptides containing phosphorylated tyrosine residue(s) were included in subsequently analyses.

*Label-free PRM analysis of A549 cells stimulated with EGF*
Mass spectra were processed with Proteome Discoverer version 2.5 (Thermo Fisher) and searched against a custom library containing the SureQuant pTyr panel, with precursor mass tolerance: 10, fragment mass tolerance: 20 mmu, along with the modifications as described in the label-free DDA analysis of A549 cells. MS1 integrated peak area quantitation was similarly performed using the minora feature detector. Spectra for targeted peptides were verified by looking for the presence of the 6 preselection product ions using in the SureQuant pTyr analyses (Supplementary Table S1) using Skyline software.

*Protein expression profiling*
LC-MS/MS raw files from the TMT-labeled global proteome analysis performed by Vasaikar et al. (15) were re-processed using Proteome Discoverer (version 2.2). All mass spectra were searched using Mascot (version 2.4) against the human SwissProt database (version 2019_6) with a tryptic enzymatic digestion, allowing for 2 missed cleavages, +/- 10 ppm parent ion tolerance. Static modifications of Cys carbamidomethylation and TMT on N-terminus and Lys residues were included, along with variable Met oxidation. Peptide spectrum matches (PSMs) were filtered according to the following criteria: Search engine rank = 1, isolation interference ≤ 30, ion score ≥ 20, peptide length ≥ 6. Relative protein abundance was calculated as the ratio of tumor abundance to reference channel abundance (TMT-131) using the summed TMT reporter ion intensities from all peptides uniquely mapped to a gene. Relative abundances were next divided by the median relative abundance ratio from each TMT channel to correct for sample loading variation within each analysis. Adjusted relative abundances for proteins quantified across all 31 tumors were z-score normalized for subsequent analyses.

SureQuant pTyr peptide panel selection

Discovery analysis data of colon cancer tumor samples and adjacent normal tissue identified nearly 800 unique pTyr peptides. This data was prioritized for panel selection by considering the following criteria: peptides identified across multiple analyses, sites with differential phosphorylation levels between tumor and non-tumor tissue, and sites known to be implicated in oncogenic signaling. Positional isomers were avoided, and sequences with a single pTyr phosphorylation event were prioritized over those with pTyr and a pSer/pThr modification in most cases (MAPK3/1 peptides being a notable exception. We did not employ a set of strict criteria for panel selection, rather we considered the features described above and then manually curated the final list based on experience generating and analyzing tyrosine phosphorylation data. Additional peptides were then added to this list using EGFR and T cell signaling data previously generated by mass spectrometry in our group. For example, we did not identify ZAP70 in our discovery analyses, but we know from the literature ZAP70 is critical for T-cell signaling. Therefore, we selected a tyrosine phosphorylated tryptic sequence previously identified by mass spectrometry to add to our panel.

SUPPLEMENTARY TABLE LEGENDS

**Supplementary Table S1**. Targeted tyrosine phosphorylated peptides, method parameters, and file map of raw data files.

**Supplementary Table S2**. Custom pathway and substrate-kinase libraries for tumor-specific pathway enrichment analyses.

**Supplementary Table S3**. Tyrosine phosphorylated peptides identified in discovery analyses of colorectal tumors.

**Supplementary Table S4**. Quantitative reproducibility comparison between SureQuant pTyr and DDA.

**Supplementary Table S5.** Quantitative dynamics comparison between SureQuant pTyr, PRM, and DDA.

**Supplementary Table S6.** Light to heavy ratios and Z-score normalized pTyr abundances for peptides identified in colon tumors.

**Supplementary Table S7**. Clinical data for tumor specimens analyzed.