nature methods

Article

https://doi.org/10.1038/s41592-024-02256-z

Post-translational modification-centric base editor screens to assess phosphorylation site functionality in high throughput

Received: 16 October 2023

Accepted: 20 March 2024

Published online: 29 April 2024

Check for updates

Patrick H. Kennedy[1,2,3], Amin Alborzian Deh Sheikh[1,2,3], Matthew Balakar[4], Alexander C. Jones[5,6], Meagan E. Olive[4], Mudra Hegde[4], Maria I. Matias[1,2,3], Natan Pirete[4], Rajan Burt[4], Jonathan Levy[7,8,9], Tamia Little[1,2,3], Patrick G. Hogan[3,10,11], David R. Liu[7,8,9], John G. Doench[4], Alexandra C. Newton[5], Rachel A. Gottschalk[12], Carl G. de Boer[13], Suzie Alarcón[14,15], Gregory A. Newby[7,8,9,16,17] & Samuel A. Myers[1,2,3,5,10,11]

Signaling pathways that drive gene expression are typically depicted as having a dozen or so landmark phosphorylation and transcriptional events. In reality, thousands of dynamic post-translational modifications (PTMs) orchestrate nearly every cellular function, and we lack technologies to find causal links between these vast biochemical pathways and genetic circuits at scale. Here we describe the high-throughput, functional assessment of phosphorylation sites through the development of PTM-centric base editing coupled to phenotypic screens, directed by temporally resolved phosphoproteomics. Using T cell activation as a model, we observe hundreds of unstudied phosphorylation sites that modulate NFAT transcriptional activity. We identify the phosphorylation-mediated nuclear localization of PHLPP1, which promotes NFAT but inhibits NFκB activity. We also find that specific phosphosite mutants can alter gene expression in subtle yet distinct patterns, demonstrating the potential for fine-tuning transcriptional responses. Overall, base editor screening of PTM sites provides a powerful platform to dissect PTM function within signaling pathways.

Nearly every eukaryotic cellular process is controlled by PTMs, which can modulate protein subcellular localization, protein–biomolecular interactions, enzymatic activity, stability and so on. Protein phosphorylation is arguably the best characterized PTM[1]. The human genome encodes roughly 500 protein kinases and 200 phosphatases that control the coupling and hydrolysis of phosphates, respectively, on substrates in a rapid and dynamic fashion[2,3]. These signaling cascades organize into elaborate biochemical networks that allow the cell to

[1]Laboratory for Immunochemical Circuits, La Jolla Institute for Immunology, La Jolla, CA, USA. [2]Center of Autoimmunity and Inflammation, La Jolla Institute for Immunology, La Jolla, CA, USA. [3]Division of Signaling and Gene Expression, La Jolla Institute for Immunology, La Jolla, CA, USA. [4]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [5]Department of Pharmacology, University of California San Diego, San Diego, CA, USA. [6]Biomedical Sciences Graduate Program, University of California San Diego, San Diego, CA, USA. [7]Merkin Institute of Transformative Technologies in Healthcare, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [8]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. [9]Howard Hughes Medical Institute, Harvard University, Cambridge, MA, USA. [10]Program in Immunology, University of California San Diego, San Diego, CA, USA. [11]Moores Cancer Center, University of California San Diego Health, La Jolla, CA, USA. [12]Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [13]School of Biomedical Engineering, University of British Columbia, Vancouver, British Columbia, Canada. [14]La Jolla Institute for Immunology, La Jolla, CA, USA. [15]AUGenomics, San Diego, CA, USA. [16]Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [17]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. e-mail: sam@lji.org

https://doi.org/10.1038/s41592-024-02256-z

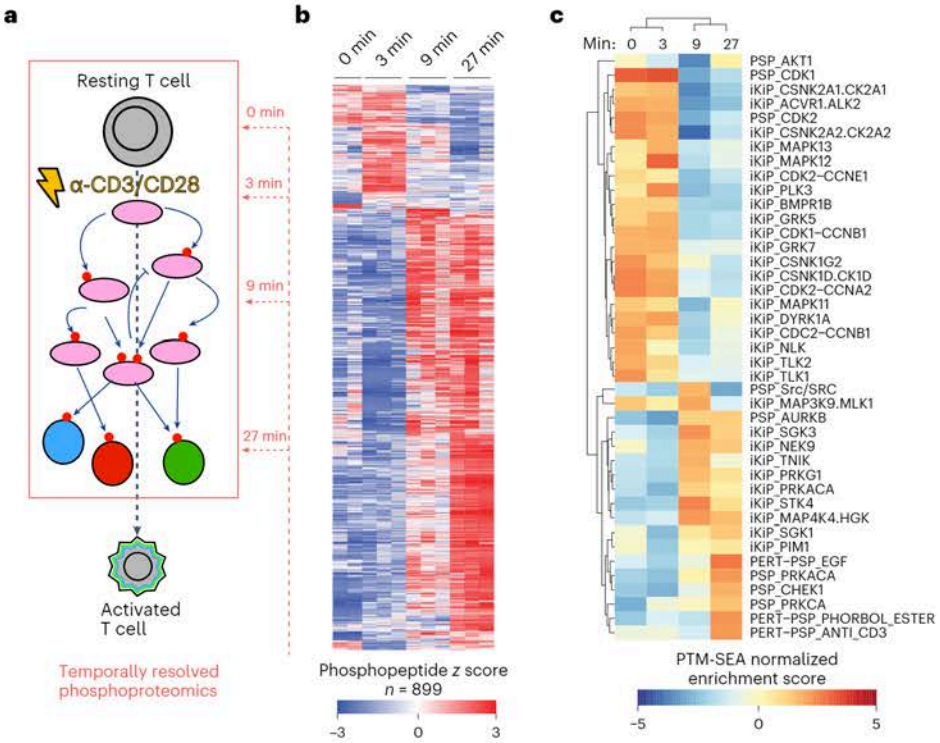


Fig. 1 | Signaling dynamics of early T cell activation. a, Diagram of early T cell activation and where the time points for the phosphoproteomic analysis were sampled. Pink ovals, kinases; colored circles, transcription factors. **b**, Heatmap of statistically regulated (moderated *F*-test) phosphopeptides during the first 27 min of T cell stimulation. Interactive data are associated with this figure (Supplementary Data File 1). **c**, Heatmap of PTM-SEA terms for the phosphoproteomics time series data. iKiP, in vitro kinase-to-phosphosite database[67]; PSP, kinase–substrate pairs from PhosphositePlus; PERT, PSP-curated perturbation.

process information about its intra- and extracellular environmental changes. Mass spectrometry-based proteomics has revolutionized our ability to map global signaling pathways at the phosphorylation site (phosphosite)-specific level, across time and cellular space. Current phosphoproteomics experiments can quantify tens of thousands of phosphosites, tracking their dynamics upon cell stimulation, drug treatment or mutational status[4–7]. Unfortunately, fewer than 3% of the nearly quarter of a million phosphorylation sites identified have an ascribed function[8,9]. Since functionally characterizing new modification sites is laborious and resource intensive, we often resort to describing complex biological systems by the limited number of well-characterized phosphosites for which we have good reagents.

Functional genomics has greatly increased our throughput for associating genes with specific cellular phenotypes. Genome-wide CRISPR–Cas9 technology coupled with phenotypic screens allow researchers to identify which genes or noncoding regions are important for a specific function such as gene expression[10], cytokine secretion[11], cell proliferation[12] or cell survival[13,14]. More recently, CRISPR–Cas9-mediated base editors, which introduce specific nucleotide substitutions in genomic DNA rather than double-stranded DNA breaks[15], have been used for mutational scanning across protein-coding genes and regulatory elements[16–19]. Base editor technology holds immense promise to study PTM site function[20] with high throughput by mutating specific amino acids, bypassing the need to create site-specific homology-directed repair templates.

Here we describe an experimental workflow to study phosphorylation site functionality with high throughput. By coupling quantitative phosphoproteomics with 'proteome-wide' base editing of individual phosphosites and phenotypic screens, we are able to functionally evaluate a large number of previously unstudied phosphosites that are involved in cell proliferation or the transcriptional responses following T cell activation. T cell activation via stimulation of the T cell receptor (TCR) and costimulation of CD28 activates several kinases that form both negative and positive feedback loops, as well as several transcription factors including NFAT, NFκB and AP1. Applying PTM-centric base editor screens to T cell activation, we show that we can recapitulate many known aspects of the pathway, while discovering new kinase activities and specific phosphorylation events that control different aspects of the transcriptional response. This allowed us to identify a specific phosphosite on the phosphatase PHLPP1 as a new regulator of T cell activation-induced NFAT and NFκB activities. Transcriptional profiling of PHLPP1 phosphosite-mutant T cells shows that individual phosphorylation events differentially impact downstream gene expression in subtle yet distinct patterns, creating the potential to map causal links between signaling and gene expression. PTM-centric base editor screens provide an experimental framework to functionally interrogate and systematically decode the vast network of biochemical signaling events to their downstream phenotypes.

Results

Optimization of base editing for experimentally derived phosphorylation sites

To develop a system by which we could profile and then functionally assess signaling pathways and their effects on gene expression, we focused on a classic model of T cell activation in the human T cell leukemia line Jurkat E6-1, during which several kinases and downstream transcription factors are activated (Fig. 1a)[21]. We performed a temporally resolved quantitative phosphoproteomics experiment, assaying global phosphorylation patterns for 0, 3, 9 and 27 min of T cell activation using α-CD3 and α-CD28 agonist antibodies[22,23]. Of the 26,037 quantified phosphopeptides, 899 were significantly differentially regulated (moderated *F*-test, false discovery rate (FDR) < 0.05) during this time

series (Fig. 1b, Supplementary Table 1 and Supplementary Data File 1). Replicates showed strong correlation and that the 3–9 min transition showed the largest changes according to principal component analysis (Extended Data Fig. 1). PTM-SEA[24]—a PTM site-centric analog to GSEA[25] (gene set enrichment analysis)—showed that various kinase activities were temporally regulated during the first 30 min of T cell activation (Fig. 1c). This analysis culminated with the perturbation signatures for 'anti-CD3' and 'phorbol esters,' indicating the temporally regulated phosphopeptides reflected the appropriate T cell activation pathways.

Using a custom bioinformatics pipeline, we queried which of the ~19,000 confidently localized phosphosites we could target using SpCas9-mediated C-to-T or A-to-G editors. We included all detected phosphosites, rather than only temporally regulated ones since there are likely to be phosphosites that were not statistically significant but still contribute to T cell activation. Considering editing windows and targetable locations, we found 7,618 unique phosphosites were targetable with 11,392 distinct single-guide RNAs (sgRNAs) using the SpCas9-A-to-G editor ABE8e, while 7,063 unique phosphosites could be targeted by the SpCas9-C-to-T editor BE4 (Fig. 2a and Supplementary Table 2)[26,27]. Roughly half of the editable phosphosites overlapped between the two base editors. The amino acid side chain representation of targetable phosphosites reflected those detected and statistically regulated (Fig. 2b). ABE8e seems to make more structurally conservative missense mutations and, unlike BE4, can target tyrosine-encoding codons (Fig. 2a).

To develop a flexible genomic engineering approach, we based our base editing strategy on previous genome-wide CRISPR–Cas9 screens in primary T cells where sgRNAs are delivered via lentivirus, followed by electroporation of Cas9 protein[12]. We nucleofected several different types of biomolecules to determine the most efficient base editors (Fig. 2c). Using Jurkat cells stably expressing a sgRNA targeting the model *HEK3* site in humans, we nucleofected either plasmid DNA, chemically synthesized and capped mRNA, or recombinant protein of different base editor versions. We found that purified recombinant ABE8e protein (NGG protospacer adjacent motif (PAM)) properly edited over 95% of adenosines in the base editing window (Fig. 2d). To test the reproducibility of the ABE8e protein, we re-expressed and purified the protein[27]. Again, we found that 92% of the adenosines in the base editing window were mutated to guanosine via Sanger sequencing[28]. These results demonstrate that we can reproducibly achieve sufficiently high base editing efficiency with ABE8e protein for high-throughput screens.

Base editing phosphosites that promote or inhibit markers of T cell activation

We tested whether mutating phosphosites in proteins known to be involved in the TCR signaling pathway with ABE8e protein would affect markers of T cell activation. We base edited the activating tyrosine of MAPK1 (also known as ERK2) Y187, and two targets in the TCR associated kinase ZAP70 Y292 and Y315 in Jurkat E6.1 cells. MAPK1 Y187 acts as a positive control as it is important for proper T cell activation signaling and transcriptional responses[29]. Mutant or control cells were activated with α-CD3/CD28 agonist antibodies for 12 h, stained for the early activation marker surface CD69 levels and analyzed by flow cytometry. Mutation of an inhibitory phosphotyrosine on ZAP70 Y292H[30] increased CD69 surface expression whereas MAPK1 Y187C showed diminished surface CD69 (Fig. 2e and Extended Data Fig. 2). ZAP70 Y315H showed no effect, consistent with previous reports[31]. Together these data establish that phosphosite mutations can have positive or negative effects on T cell activation levels.

Phosphosite-centric functional phenotypic screens using pooled base editing for cell proliferation

To assess base editing efficiency in pooled format, we created a lentiviral library consisting of roughly 11,000 phosphosite-targeting sgRNAs for missense mutations, 250 nontargeting controls and 250 intergenic controls as negative controls. We also included 250 guides that introduce terminating edits in essential genes via mRNA splice site disruption, effectively knocking out the gene. Triple parameter reporter (TPR) Jurkat cells, which have individual fluorescent reporters driven by separate NFAT, NFκB and AP1 transcriptional response elements[32], were transduced at a multiplicity of infection of 0.3. After puromycin selection, a 500× library coverage aliquot of cells was collected and the rest were electroporated with ABE8e protein. To confirm our base editing was efficient, we analyzed the representation of sgRNAs, comparing pre- and 6 days post-ABE8e protein electroporation (Fig. 3a). The representation of sgRNAs disrupting splice junctions in 250 essential genes was significantly lower 6 days after base editing compared with all sgRNAs in the library (Fig. 3b) indicating our base editing approach was working efficiently. The relative representation of intergenic and nontargeting controls was not affected by introduction of ABE8e protein.

We next examined whether mutation of phosphosites important for cell division would affect cell viability or proliferation in our pooled format without specific stimuli or selective pressure[33]. Examination of phosphorylation sites on CDK1—a kinase whose activity is necessary for proper cell division, showed that Y160C;T161A and Y15C reduced cell viability post-ABE8e electroporation similar in magnitude and direction to levels seen previously using homology-directed recombination[34] (Fig. 3c). The silent mutation of Y19Y showed no effect. Analysis of the whole phosphosite-mutant dataset using the Model-based Analysis of Genome-wide CRISPR–Cas9 Knockout (MAGeCK)[13] resulted in 89 sgRNAs that were significantly enriched in pre-ABE compared with postediting (Fig. 3d and Supplementary Table 3). sgRNAs targeting phosphosites that were depleted after ABE8e protein introduction were enriched for genes involved in the term 'Cell Cycle' amongst the top three Reactome pathways (Extended Data Fig. 3). A total of 58 sgRNAs introducing phosphosite mutations were enriched post-base editing, suggesting a proliferative advantage (Fig. 3d). These genes belonged to the Reactome signatures for 'membrane trafficking', 'metabolism of RNA' and 'transcriptional regulation by RUNX1', which are all pathways involved in the proliferation of glioblastoma and acute myeloid leukemia cells[35–37] (Extended Data Fig. 3). Gene Ontology analysis of the genes knocked out through splice site disruption showed an enrichment for mRNA processing and cell cycle (Extended Data Fig. 3). PTM site-centric pathway analyses can identify groups of phosphosites (in this case, mutation of phosphosites) enriched in the pre- or postedited pools. Kinase Library[38], which uses primary sequence motifs derived from biochemical kinase reactions to predict kinase activity through motif enrichment, identified CDK1/4/5/6/13/18 motifs from aggregated phosphosite mutants depleted relative to pre-edited cells (Fig. 3e). MELK1, NIM1, DYRK2/4 and YANK2/3 motifs were enriched in the pre-edited cells. Together, these data show that our approach of base editing phosphorylation sites in a pooled format can identify phosphorylated residues and putative kinases important for proper cell cycle proliferation or survival.

Coupling functional phosphosite screens with transcriptional reporters identifies new regulators of NFAT transcriptional activity

To screen for phosphosites that are functionally linked to transcriptional outputs, we performed a 'proteome-wide' base editor screen in activated TPR Jurkat cells, utilizing a green fluorescent protein (GFP) transcriptional reporter driven by the 4x *ARRE2* sequence derived from the murine *IL2* promoter[39]. TPR Jurkat cells stably integrated with the sgRNA library described above were electroporated with ABE8e protein, stimulated with α-CD3/CD28 antibodies for 12 h, and sorted for high and low GFP (NFAT activity) levels (Fig. 4a). Genomic DNA was collected from the cells sorted into high and low bins; sgRNA identifiers were amplified by PCR, and were sequenced by next-generation sequencing (NGS). Total read-normalized, log transformed sgRNA counts were

https://doi.org/10.1038/s41592-024-02256-z

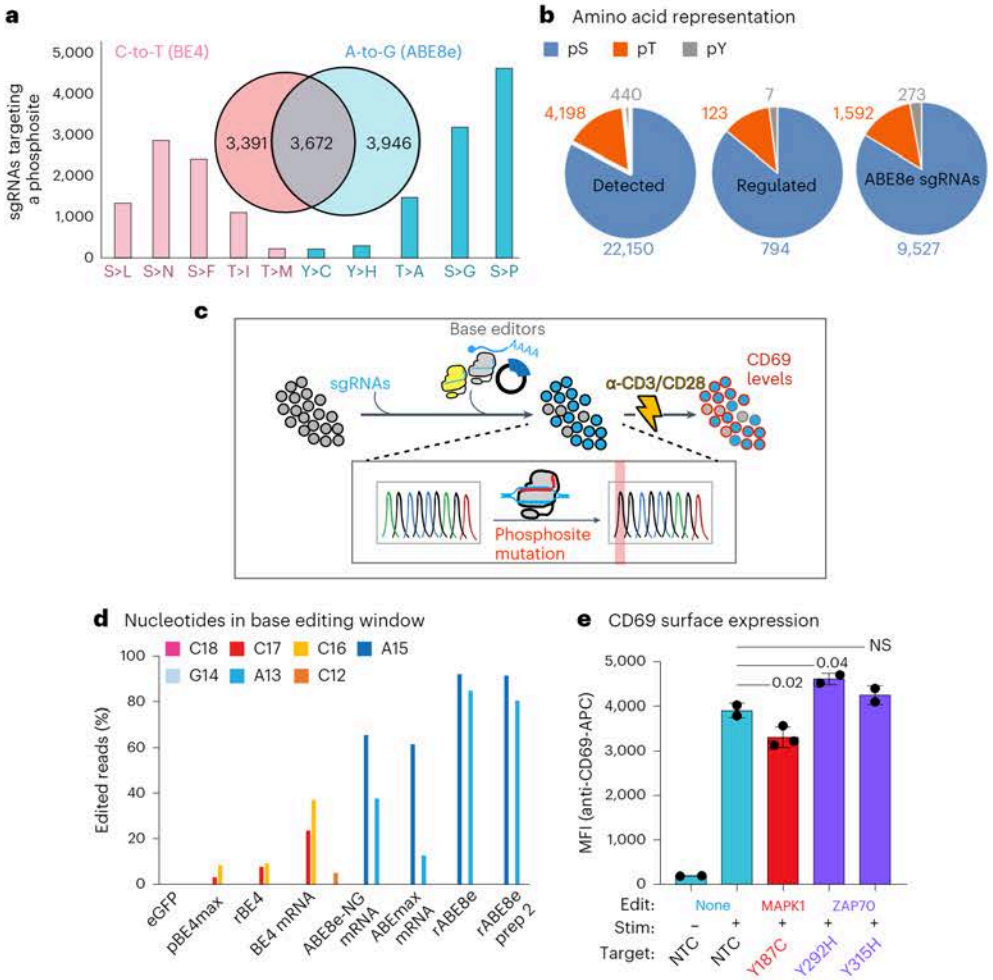


Fig. 2 | Base editing capabilities of empirically derived phosphorylation sites. a, Bar graph showing the number of sgRNAs that target phosphosites for the respective base editors BE4 (pink) or ABE8e (blue) and what the resultant amino acid codon will be for BE4 or ABE8e. BE4 cannot mutate Y to any other amino acid and was omitted. The Venn diagram inset shows the number of distinct phosphosites targeted by at least one sgRNA. **b**, Distribution of all putative phosphorylated amino acid side chains from all phosphopeptides detected by mass spectrometry (left), the number of statistically regulated phosphopeptides in Fig. 1b (middle) and the number of sgRNAs that can be used by ABE8e to target a phosphosite (right). **c**, Diagram for testing various base editor delivery methods in an arrayed format, one sgRNA at a time followed by T cell activation. **d**, Base editing efficiency, as determined by NGS amplicon sequencing in percentage of edited reads, testing the nucleofection of different biomolecules to deliver base editors. The base editing window, counting right to left from the PAM sequence, is shown in the inset and the nucleic acid targets are color coordinated. p, plasmid; r, recombinantly expressed and purified; mRNA, synthetic capped mRNA. **e**, Effect of phosphosite base edits on T cell activation-induced CD69 surface levels. Edits, targeted gene; Stim, 12 h treatment with α-CD3/CD28 agonist antibodies; Target, amino acid targeted; NTC, nontargeting control at the *HEK3* locus; MFI, mean fluorescence intensity. Two sample *t*-test *P* values are shown; NS, not significant. *n* = two or three editing replicates, indicated by the datapoints where s.d. is shown.

moderately to strongly correlated indicating sufficient data quality between quadruplicates, and the mean correlation between sgRNA counts in GFP high and low bins was strong (Extended Data Fig. 4a, b and Supplementary Table 4). We used MAGeCK[13] to identify and rank which phosphosite mutations, the targets of the sgRNAs, regulate the NFAT transcriptional reporter. Phosphosites with several sgRNAs (Extended Data Fig. 4c) were combined in MAGeCK and the high and low GFP bins were compared. We identified 411 sgRNAs enriched in the GFP high bin, and 293 in the GFP low bin (Fig. 4b and Supplementary Table 4). Rolling up our phosphosite level perturbations to the gene level (gene-centric), we performed pathway analyses, using several tools, to assess the fidelity of our approach. Enrichment analysis using MAGeCKFlute[40] identified TCR pathways as enriched in the GFP low (Fig. 4c). g:Profiler[41], a Gene Ontology-based analysis, also identified TCR pathway in the GFP low bin (Extended Data Fig. 4d). GSEA[25] identified the 'TCR Calcium Pathway' signature in the GFP high bin (Extended Data Fig. 4e), probably due, in part, to dephosphorylation of NFAT as a regulatory mechanism[42]. Edits containing bystander mutations next to the target phosphosite were not enriched in the GFP high and low bins compared with the library as a whole, suggesting that they do not confer extra effects (Extended Data Fig. 4f).

Utilizing PTM site-centric analyses, Kinase Library identified several kinases implicated in NFAT activity regulation, such as JNK[43], NLK[44] and CAMK2G[45], enriched in the GFP low bin (Fig. 4d). The GFP high bin was enriched for the CDK9 motif, a kinase involved in general transcriptional regulation[46] as well as motifs of kinases known to be involved in T cell activation such as PAK2 (ref. 47) and CDK5 (ref. 48). CLK3, PAK4, SRPK1 and MYLK4 were also enriched in the GFP high bin but have poorly characterized roles in controlling NFAT transcriptional activity (Fig. 4d). PTM-SEA[24] agreed with the Kinase Library results for PAK2 and CDK9 (Fig. 4e). However, PTM-SEA also identified mutations of MAPK1, MTOR and DYRK1A/2/3 substrates to be enriched in the GFP low bins, corroborating these kinases' involvement in T cell activation (Fig. 4e). As expected, these results demonstrate that phosphosite

https://doi.org/10.1038/s41592-024-02256-z

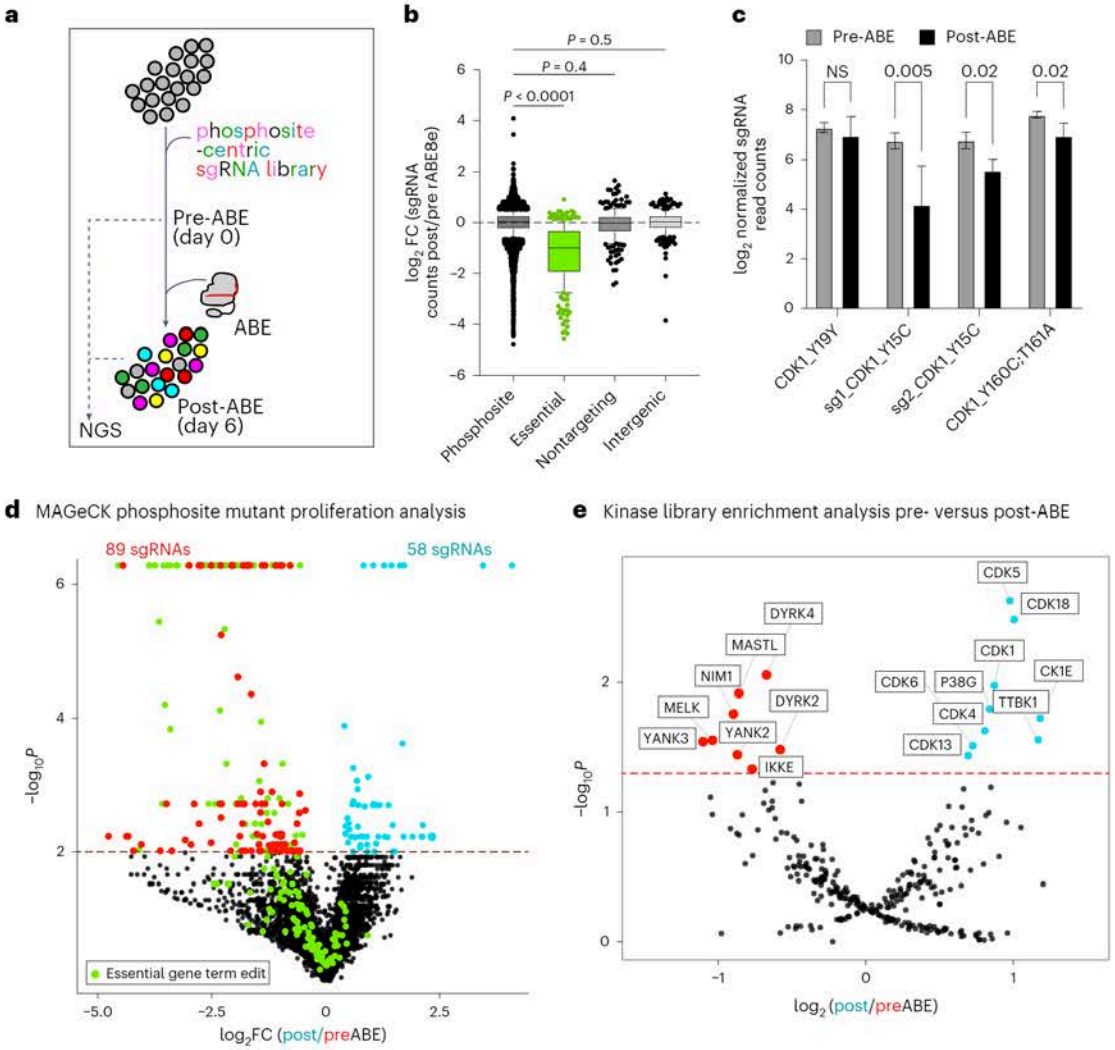


Fig. 3 | Base editing screening reveals phosphosites involved in proliferation or survival. a, Diagram of pooled base editor screens for phosphosites or terminating edits of essential genes important for cell proliferation or survival followed by NGS. **b**, $\log_2$FC between pre- and post-ABE8e protein introduction of cells expressing the sgRNA library. Essential, terminating edit into essential genes. Intergenic base edits and nontargeting controls are also shown. Two sample *t*-test *P* values are shown; $n = 4$ transduction replicates. Box and whiskers plot shows median, quartiles, maximum and minimum and outliers (individual datapoints). **c**, Mutations made to CDK1 phosphosites and their influence of sgRNA representation before and after base editing in a pooled format. 'sg1 and sg2' indicate that two different sgRNAs were used for the Y15C mutation. Two sample *t*-test *P* values are shown; $n = 4$ transduction replicates where s.d. is shown. **d**, Volcano plot showing the distribution of all sgRNAs pre- and post-ABE8e protein introduction as determined by MAGeCK analysis. Green points, terminating edits in essential genes; red points, statistically significant sgRNA targeting phosphosites depleted after base editing; blue points, sgRNAs targeting phosphosites that were enriched after base editing. Enrichment values and statistical thresholds were determined by MAGeCK[13]. **e**, Kinase library, site-centric enrichment analysis of phosphosite mutants, as an aggregate motif, enriched in post-ABE8e-edited cells (blue) or pre-ABE8e-edited cells (red) bins. Enrichment values by MAGeCK[13], and the one-sided Fisher's exact test *P* value after Benjamini–Hochberg correction is shown by the dashed red line, adj. $P < 0.05$.

mutations can directly implicate their involvement in regulating NFAT transcriptional activity, recapitulating known signaling pathways that were constructed from various studies of general T cell activation. These results also strongly suggest that our approach of base editing phosphosites is not only capable of rediscovering crucial signaling molecules in T cell activation, but also provides new insights by identifying new regulatory kinases and phosphorylation events.

We next asked how phosphopeptide abundance dynamics related to functional readouts in our NFAT-GFP screen. As phosphopeptide abundances could have increased and/or decreased in our timecourse data, we plotted the *F*-statistic from the mass spectrometric analyses against the $\log_2$ fold change (FC) in the GFP high and low bins calculated by MAGeCK. We found that there was an overlap between the two datasets, where MAPK1 Y817C was amongst the largest changes in phosphopeptide levels and GFP expression (Extended Data Fig. 4g). We also tested whether phosphosites predicted to be functional by Ochoa et al.[49] correlated with our empirical data for NFAT-GFP activity. Although there was a positive trend between our phenotypic screen data and predicted functional scores, the correlation was not statistically significant (Extended Data Fig. 4h,i). These results underscore the need for PTM-centric phenotypic screens.

PTM-centric base editor screening identifies the uncharacterized phosphorylation-mediated nuclear localization of PHLPP1

PHLPP1 is the protein phosphatase implicated most widely in AKT signaling in cancer[50,51]. In macrophages, PHLPP1 attenuates the JAK/STAT axis by dephosphorylating STAT1 (ref. 52). The post-translational

https://doi.org/10.1038/s41592-024-02256-z

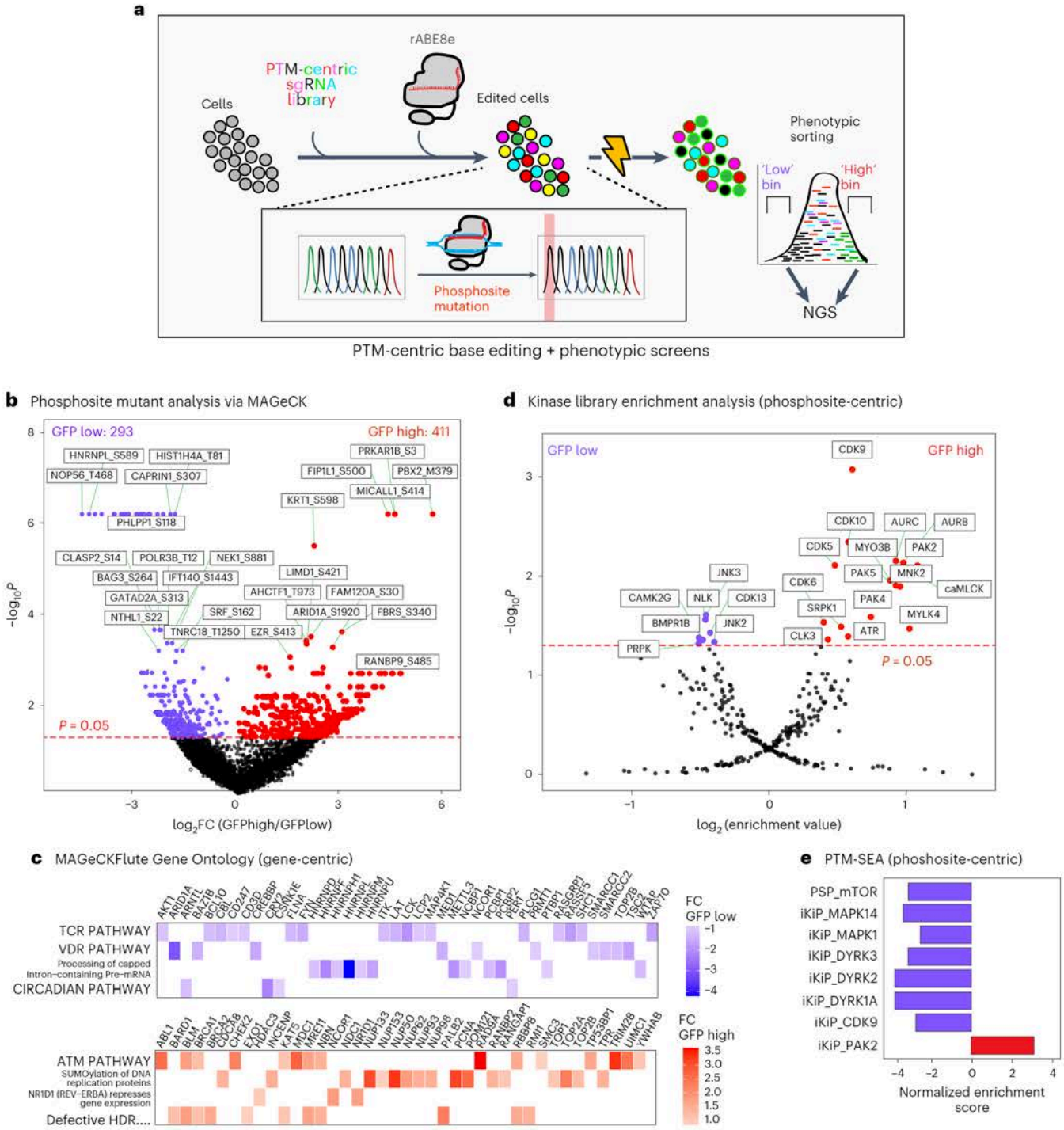


Fig. 4 | Proteome-wide base editing of phosphosites modulating NFAT transcriptional activity. a, Diagram of PTM-centric, proteome-wide base editing coupled to NFAT-GFP transcriptional reporter followed by NGS. **b**, Volcano plot comparing phosphosite edits in GFP high (red) with GFP low (purple) bins as determined by MAGeCK. Statistical thresholds were also determined by MAGeCK[13]. **c**, MAGeCKFlute gene-centric pathway analysis of genes with mutated phosphosites enriched in the GFP low (purple) or GFP high (red) bins. Genes in the respective pathways and their FC are shown. **d**, Kinase library, site-centric enrichment analysis of phosphosite mutants enriched in the GFP low (purple) or GFP high (red) bins. Enrichment values by MAGeCK[13], and the one-sided Fisher's exact test *P* value after Benjamini–Hochberg correction is shown. **e**, PTM-SEA of phosphosites mutated by ABE8e protein and enriched in the GFP low (purple) or GFP high (red) bins. iKiP, in vitro kinase-to-phosphosite database[67]; PSP, kinase–substrate pairs from PhosphositePlus. MAPK14 is also known as P38A.

mechanisms controlling PHLPP1 function and its involvement in T cell biology[53] are poorly understood. We identified the mutation PHLPP1 S118P in our base editing screen to have a strong negative impact on NFAT transcriptional activity, on par with that of MAPK1 Y187C (Fig. 5a). We nucleofected ribonucleoproteins comprised of in vitro transcribed sgRNA coupled to ABE8e protein to validate our screen results.

https://doi.org/10.1038/s41592-024-02256-z

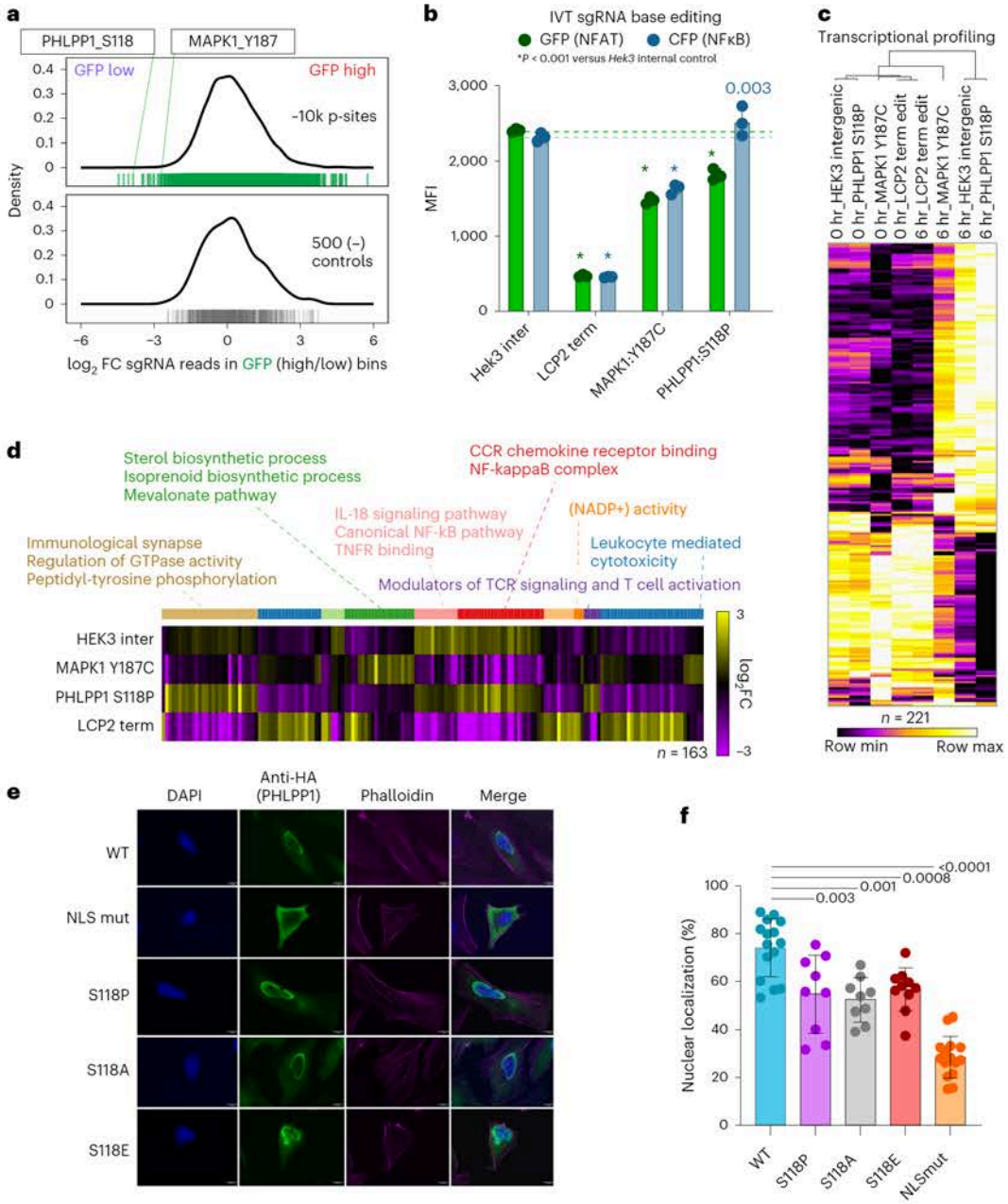


Fig. 5 | Phosphorylation-induced nuclear translocation of PHLPP1 promotes NFAT and represses NFκB transcriptional responses. a, Distribution of $\log_2$ FCs of the ~11,000 sgRNAs inducing phosphorylation mutations in the NFAT (GFP) transcriptional activity screen (top, green) and nontargeting and intergenic controls (bottom, gray) between GFP low and GFP high bins. PHLPP1 S118P and MAPK1 Y187C mutations are labeled for comparison. **b**, Validation of NFAT activity screening hits using electroporation of in vitro transcribed sgRNAs coupled with ABE8e protein, followed by α-CD3/CD28 stimulation for 16 h and analysis of GFP (NFAT) or cyan fluorescent protein (CFP; NFκB) transcriptional activity reporters. Two sample *t*-test *P* values compared with *HEK3* relative color control are shown if $P > 0.001$; $n = 3$ activation replicates, where s.d. is shown. **c**, Heatmap of differentially expressed genes between *HEK3* intergenic mutant control, MAPK1 Y187C or PHLPP1 S118P edited Jurkat cells activated for 0 or 6 h as determined by single-cell RNA sequencing. The *LCP2* terminating edit is shown for comparison but was not used for statistical testing (Supplementary Data File 2). **d**, *k*-means clustering and g:Profiler Gene Ontology analysis (all $P < 0.05$) of the differentially expressed gene clusters at 6 h between *HEK3* intergenic mutant control, MAPK1 Y187C or PHLPP1 S118. The *LCP2* terminating edit is shown for comparison but was not used for statistical testing. Cluster numbers count from left to right and are designated by color. **e**, Spinning disk confocal microscopy images showing the subcellular localization of PHLPP1 NTE constructs. NLS mut, full mutation of the two nuclear localization sequences in the NTE of PHLPP1. The S118P, S118A and S118E mutations are also shown. **f**, Quantification of PHLPP1 NTE constructs and the percentage of α-HA signal in the nucleus. Unpaired, two sample *t*-test *P* values are shown. Each datapoint represents an individual cell where s.d. is plotted.

The intergenic mutation at the *HEK3* site was used as an editing control. For positive effect controls, we introduced a terminating edit in the scaffolding protein LCP2 (SLP76), a critical molecule for proper T cell activation[12] or the MAPK1 Y187C mutation, which prevented or altered T cell activation, respectively. After bulk editing was assessed by amplicon sequencing ($\bar{x} = 70\%$; Extended Data Fig. 5), four to eight validated single-cell clones were combined to avoid clone-specific effects. Editing at the top two predicted off-target sites for all clones

was within the noise of the assay, about <3% editing. MAPK1 Y187C and PHLPP1 S118P both showed diminished NFAT transcriptional activity (Fig. 5b). However, unlike MAPK1 Y187C, PHLPP1 S118P showed a small but statistically significant increase in cyan fluorescent protein (CFP) levels (NFκB activity reporter). To extend this analysis, we performed transcriptional profiling of clonal Jurkat T cells harboring the homozygous *HEK3* intergenic edit, *LCP2* terminating edit, the MAPK1 Y187C or the PHLPP1 S118P mutations for 0 and 6 h post CD3/CD28 stimulation. *LCP2*-terminated cells showed no effect of transcriptional activation 6 h after activation, whereas MAPK Y187C showed an intermediate pattern compared with the *HEK3* controls (Fig. 5c, Supplementary Table 5 and Supplementary Data File 2). The PHLPP1 mutant cells were similar in their transcription patterns compared with the *HEK3* editing controls, although differences were apparent (Fig. 5c). We highlighted the genes differentially expressed between *HEK3*, MAPK1 Y187C, and PHLPP1 S118P, and plotted them alongside the *LCP2* terminating edit cells (Fig. 5d). After *k*-means clustering and Gene Ontology analysis of differentially expressed genes, clusters 1, 5 and 6, which were higher in the *PHLPP1* mutant compared with the *HEK3* control, identified several terms associated with NFκB signaling (NF-kappaB complex, TNFR signaling), corroborating our transcriptional reporter results (Fig. 5b,d). The MAPK1 Y187C mutant cells were enriched for genes in sterol and isoprenoid biosynthesis (cluster 4 of Fig. 5d). These results suggest different phosphorylation sites in the T cell activation pathway can regulate downstream gene expression in disparate ways.

S118 lies within the bipartite nuclear localization sequences (NLSs) at the N-terminus of PHLPP1 (ref. 52). To test the hypothesis that phosphorylation controls PHLPP1 subcellular localization, we expressed N-terminal extension (NTE) constructs ectopically with the wild-type sequence, both halves of the NLS mutated, or S118P, the result of A-to-G editing. We also included the more archetypal amino acid substitution, S118A, which removes the phosphorylatable residue without potential structural changes introduced by the rotationally constrained amino acid proline, and the S118E mutation as a potential phosphomimetic. We found that both S118P and S118A reduced PHLPP1 NTE nuclear localization to comparable levels, but less severe than the full NLS mutant (Fig. 5e,f). S118E showed a similar loss of nuclear localization compared with S118A or S118P, indicating that, in this case, the glutamate may not accurately reflect bona fide phosphorylation[54]. These results suggest that phosphorylation of the NTE of PHLPP1 regulates nuclear localization. These data also suggest that serine to proline substitutions are reasonable proxies for loss of a phosphorylatable residue for screening purposes.

T cell transcriptional activation can be dissected through precise, base editor-mediated signaling modulation

Transcriptional profiling *HEK3* (intergenic editing control), *MAPK1* and *PHLPP1* mutant Jurkat T cells showed more differentially expressed genes 6 h poststimulus than in resting conditions (BH corrected *P* value < 0.05) (Fig. 5c), suggesting the mutated phosphosites indeed affect T cell activation-induced transcriptional responses. Inspection of the differentially expressed genes at 6 h post T cell activation showed specific T cell-related genes are expressed at subtle but different levels between phosphosite-mutant genotypes (Fig. 6a and Extended Data Fig. 6a). For example, *BCL11A* and *THEMIS* were activated to a higher extent in MAPK1 Y187C mutants compared with PHLPP1 S118P or control. In contrast, *NR4A3*, *ZFP36L1* and *IL21R* were highest in the PHLPP1 S118P cells, whereas *GZMA*, *TNFSF14*, *NFKBIA* and *JUND* were highest in the editing control cells (intergenic *HEK3*). Intracellular GZMB staining 24 h after T cell activation corroborated the gene expression results, showing a loss of GZMB protein in MAPK1 Y187C but not PHLPP1 S118P cells (Fig. 6b and Extended Data Fig. 6b). These results suggest that mutating different phosphosites in ostensibly the same signaling pathway can alter transcriptional responses, and may provide a means for fine-tuning gene expression.

Discussion

Linking specific signaling events to their downstream functions is a fundamental goal of biology. Choosing new phosphosites for mechanistic follow-up studies is often resource intensive and laborious. We aimed to create a screening platform to functionally assess and prioritize individual phosphosites and how they contribute to specific phenotypes with high throughput by integrating mass spectrometry-based phosphoproteomics with CRISPR-mediated base editor screens. We found that mutation of phosphosites via base editing in a pooled format can assess multiple functional readouts, in positive and negative directions, for proliferative or transcriptional phenotypes. This new experimental and computational framework will greatly enable future studies, allowing the community to address the complexity of signaling pathways.

Our study focused on phosphosites identified empirically in a parallel experiment rather than from PTM repositories or databases, mitigating the introduction of indiscriminate coding mutations that can have PTM-independent effects[16,17,19,33]. Moreover, as phosphoproteomic analyses of new, unstudied systems continue to grow (for example, in patient-specific cancers, primary mouse or human immune cells and so on), the bioinformatic tools needed to identify base editor-targetable phosphosites from empirical mass spectrometry data will become increasingly important (Fig. 1d).

All gene-centric pathway analyses from genes with mutated phosphosites that altered NFAT activity revealed that the TCR pathway was enriched, indicating the results from our phosphosite base editor screen can recapitulate known aspects of well-studied pathways. Interestingly, gene-level GSEA analysis found 'TCR Calcium Pathway' enriched in the GFP high bin. The phosphosite mutations driving this result were all in NFAT isoforms, or molecules known to regulate NFAT dephosphorylation: NFATC1/2/3, CABIN1 and RCAN1. This pathway regulates gene expression, cytokine production and T cell activation through the NFAT signaling pathway by dephosphorylation of NFAT molecules, replicating the effect of Calcineurin (*PPP3CC*) in the translocation of NFAT and T cell activation[42,55]. Site-centric pathway analyses of enriched phosphosite mutations revealed two main trends in the signatures identified in our Jurkat T cell activation model: TCR signaling and cell cycle. As Jurkat cells are rapidly dividing transformed cells[21], it is not surprising that many of the phosphosites identified by mass spectrometry, and then subsequently in base editor screens, identified cell cycle genes and phosphorylation events. The DYRK family of kinases is a good example (Figs. 3e and 4e). DYRKs are well known to control NFAT transcriptional activity[56,57], but to also have roles in T cell proliferation[58]. Our results suggest that DYRKs have a stronger role in T cell activation compared with proliferation. We envision that applying base editor screens of PTM sites in primary immune cells will provide clearer connections within signaling-to-transcription networks.

Our PTM-centric base editor phenotypic screen identified a new mechanism for PHLPP1 as a regulator of T cell activation-mediated gene expression. The function of PHLPP1 has been studied primarily in cancer contexts[50,51], although its role in regulatory T cell development was identified through a genetic study[53]. Our analyses originally identified it as a positive regulator of NFAT activity and, through transcriptional profiling, we found that PHLPP1 can negatively regulate genes downstream of NFκB. This regulation is probably due to the phosphorylation-induced nuclear translocation of PHLPP1, although the precise substrates of dephosphorylation via PHLPP1 remain to be determined. It is worth noting that the phosphopeptide identifying PHLPP1 pS118 was induced at 9 min by 1.6-fold on average but was not statistically regulated in our phosphoproteomics data (Supplementary Table 1). Conversely, MAPK1 Y817C had both one of the strongest FCs at the phosphopeptide abundance level and in the NFAT-GFP screen. The complementary but orthogonal information between the phosphoproteomics and the base editor screen results is probably due to the specific nature of our chosen screening platform—an NFAT activity reporter. T cell activation is highly complex and not

https://doi.org/10.1038/s41592-024-02256-z

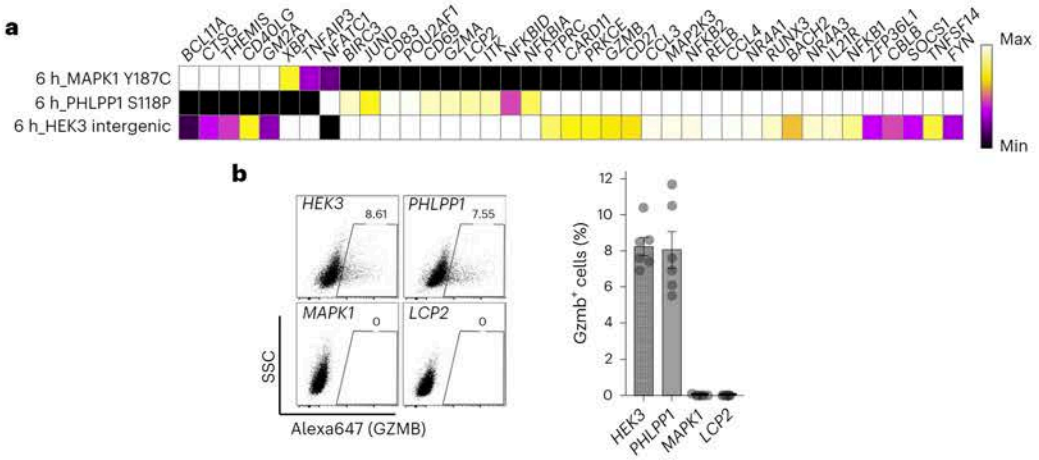


Fig. 6 | Dissecting T cell activation transcriptional responses. a, Select T cell-related differentially expressed genes at 6 h post T cell activation between *HEK3* (control), PHLPP1 S118P and MAPK1 Y187C mutant cells. This data in $\log_2$ FC space is available as Supplementary Data File 2. **b**, Intracellular GZMB staining in *HEK3* (control), *PHLPP1* S118P, *MAPK1* Y187C or *LCP2* term edited mutant cells 24 h post T cell activation. Representative plot (left) and quantification of percentage of GZMB$^+$ cells (right); $n = 6$ activation replicates where s.d. is shown.

limited to transcriptional outputs and, thus, a substantial proportion of phosphorylation events may be completely unrelated to induction of NFAT activity. This underscores the need for functional screens of PTM sites as small changes to catalytic enzymes may have prominent downstream effects.

Gene expression profiling of various phosphosite-mutant Jurkats after T cell activation revealed subtle but different transcriptional responses. For example, the *MAPK1* phosphomutant showed decreased expression of *GZMB* compared with the *PHLPP1* mutant, at the protein and mRNA level (Fig. 6), despite having similar effects on NFAT reporter activity (Fig. 5). *MAPK1* phosphomutant cells expressed less *NR4A1* and *NR4A3* than the PHLPP1 phosphomutant cells. GZMB is the main effector molecule of the cytotoxic program and NR4As have been shown to be drivers of T cell exhaustion[59]. These results together lead to an intriguing possibility of fine-tuning expression levels of specific genes, through genetic or small molecule inhibitor manipulation of signaling pathways.

Proteome-wide, PTM-centric base editing coupled to phenotypic screens provides a powerful experimental framework with which to untangle the vast network of biochemical signaling reactions and how they lead to the control of specific cellular functions. This phosphosite-targeting base editing approach directly assesses the impact of a phosphosite mutation on a given phenotype rather than relying purely on evolutionary or structural conservation, and provides interpretable results with high confidence for further mechanistic studies[33,49,60,61]. With currently available base editors it should be possible to functionally screen other PTMs including acetylation/ubiquitination (lysine), O-GlcNAcylation (Ser/Thr), cysteine[18] or even specific proteolysis events (caspases), utilizing established infrastructure common to many research institutions. We envision this approach to be widely enabling to the cell biology community.

Limitations and future steps for PTM-centric base editor screens

We focused on ABE8e for our technology development because it can be expressed and purified readily from *Escherichia coli*, and its activity after electroporation into cells harboring sgRNA-expressing plasmids is highly efficacious (Fig. 2c,d)[26,27]. Also, as an A-to-G editor it can mutate tyrosine residues—a small but important fraction of total phosphosites. This property is missing from C-to-T editors. The amino acid substitutions made by current base editors to study phosphorylation are not archetypal (S/T to A; Y to F), threonines targeted by ABE8e being the exception (Thr to Ala). Over half of the serines mutated in our study were to proline—a limitation of current base editors. However, since several kinase families are proline directed, and most phosphosites are in flexible loops, this is likely to be problematic in only a subset of cases[62–64]. Empirically, the vast majority of Pro substitutions in our data had no effect, arguing it is not inherently disruptive. In fact, our mutational analysis of the PHLPP1 phosphosite showed similar effects between the S118P and the archetypal S118A mutation used for proper validation, suggesting that proline mutations are not invariably detrimental to screen for phosphosite function. Of course, it is prudent to validate the findings of a screen with orthogonal approaches. Prime editors, which can install specific, desired sequences, will probably become increasingly useful as they become adapted for genome-wide screens[65]. They may also allow for installation of phosphomimetic substitutions, S/T/Y to D or E. However, carboxylic side chains can often not properly mimic a bona fide phosphorylated residue[54]. Moreover, employing Cas molecules with less restrictive PAM sequences (NG rather than NGG), should double to triple the percentage of targetable phosphosites[19].

Our nucleofection of purified ABE protein approach borrows from previous literature performing CRISPR-mediated gene knockouts in primary human myeloid and T cells, the latter of which can be performed at a genome-wide scale[12,66]. As our phosphosite-targeting sgRNA library is much smaller than standard genome-wide libraries, our PTM-centric workflow should be adapted readily to other systems. Moreover, as mass spectrometry-based phosphoproteomics becomes more sensitive and robust, and unstudied systems such as patient-specific cancers or a variety of human or mouse immune cells start to be analyzed, the tools developed here should enable the community to study new signaling pathways in a more comprehensive and nonbiased manner.

Considerations for implementation

We chose to focus on mutating phosphosites identified in a parallel phosphoproteomics experiment. This was driven by the extensive evidence that indiscriminate amino acid mutations can have large, PTM-independent effects on protein function[16,17,19,33]. If a quantitative, temporally resolved phosphoproteomics experiment is not feasible, we envision a 'happy medium' could be a nonquantitative phosphorylation analysis of the system of interest. Now, if one detects a phosphosite, one can test whether or not it is functionally important. Dynamics can be figured out after importance is established.

It is also possible that ABE8e expression, purification and electroporation is too cumbersome for some laboratories. While the protocol seems robust, having produced ABE8e protein in three different laboratories for this study[27], methods that have come out since we started our study show that viral vectors can be used[19]. This should remove the cell number bottleneck of nucleofection and may be more well-suited for laboratories without protein purification capacity.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-024-02256-z.

References

1. Hunter, T. Why nature chose phosphate to modify proteins. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 2513–2516 (2012).
2. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).
3. Chen, M. J., Dixon, J. E. & Manning, G. Genomics and evolution of protein phosphatases. *Sci. Signal* **10**, eaag1796 (2017).
4. Katrancha, S. M. et al. Trio haploinsufficiency causes neurodevelopmental disease-associated deficits. *Cell Rep.* **26**, 2805–2817.e9 (2019).
5. Martinez-Val, A. et al. Spatial-proteomics reveals phospho-signaling dynamics at subcellular resolution. *Nat. Commun.* **12**, 7113 (2021).
6. Koch, H. et al. Phosphoproteome profiling reveals molecular mechanisms of growth-factor-mediated kinase inhibitor resistance in EGFR-overexpressing cancer cells. *J. Proteome Res* **15**, 4490–4504 (2016).
7. Paulo, J. A. & Gygi, S. P. A comprehensive proteomic and phosphoproteomic analysis of yeast deletion mutants of 14-3-3 orthologs and associated effects of rapamycin. *Proteomics* **15**, 474–486 (2015).
8. Needham, E. J., Parker, B. L., Burykin, T., James, D. E. & Humphrey, S. J. Illuminating the dark phosphoproteome. *Sci. Signal* **12**, eaau8645 (2019).
9. Hornbeck, P. V. et al. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520 (2015).
10. Dixit, A. et al. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866.e17 (2016).
11. Parnas, O. et al. A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675–686 (2015).
12. Shifrut, E. et al. Genome-wide CRISPR screens in primary human T cells reveal key regulators of immune function. *Cell* **175**, 1958–1971.e15 (2018).
13. Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
14. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
15. Rees, H. A. & Liu, D. R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* **19**, 770–788 (2018).
16. Hanna, R. E. et al. Massively parallel assessment of human variants with base editor screens. *Cell* **184**, 1064–1080.e20 (2021).
17. Lue, N. Z. et al. Base editor scanning charts the DNMT3A activity landscape. *Nat. Chem. Biol.* **19**, 176–186 (2023).
18. Li, H. et al. Assigning functionality to cysteines by base editing of cancer dependency genes. *Nat. Chem. Biol.* **19**, 1320–1330 (2023).
19. Schmidt, R. et al. Base-editing mutagenesis maps alleles to tune human T cell functions. *Nature* **625**, 805–812 (2024).
20. Yeh, W.-H., Chiang, H., Rees, H. A., Edge, A. S. B. & Liu, D. R. In vivo base editing of post-mitotic sensory cells. *Nat. Commun.* **9**, 2184 (2018).
21. Abraham, R. T. & Weiss, A. Jurkat T cells and development of the T-cell receptor signalling paradigm. *Nat. Rev. Immunol.* **4**, 301–308 (2004).
22. Larange, A. et al. A regulatory circuit controlled by extranuclear and nuclear retinoic acid receptor α determines T cell activation and function. *Immunity* **56**, 2054–2069 (2023).
23. Abelin, J. G. et al. Workflow enabling deepscale immunopeptidome, proteome, ubiquitylome, phosphoproteome, and acetylome analyses of sample-limited tissues. *Nat. Commun.* **14**, 1851 (2023).
24. Krug, K. et al. A curated resource for phosphosite-specific signature analysis. *Mol. Cell Proteom.* **18**, 576–593 (2019).
25. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
26. Richter, M. F. et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891 (2020).
27. Huang, T. P., Newby, G. A. & Liu, D. R. Precision genome editing using cytosine and adenine base editors in mammalian cells. *Nat. Protoc.* **16**, 1089–1128 (2021).
28. Kluesner, M. G. et al. EditR: a method to quantify base editing from Sanger sequencing. *CRISPR J.* **1**, 239–250 (2018).
29. Helou, Y. A., Nguyen, V., Beik, S. P. & Salomon, A. R. ERK positive feedback regulates a widespread network of tyrosine phosphorylation sites across canonical T cell signaling and actin cytoskeletal proteins in Jurkat T cells. *PLoS ONE* **8**, e69641 (2013).
30. Bottini, N. et al. Activation of ZAP-70 through specific dephosphorylation at the inhibitory Tyr-292 by the low molecular weight phosphotyrosine phosphatase (LMPTP). *J. Biol. Chem.* **277**, 24220–24224 (2002).
31. Di Bartolo, V. et al. Tyrosine 319, a newly identified phosphorylation site of ZAP-70, plays a critical role in T cell antigen receptor signaling. *J. Biol. Chem.* **274**, 6285–6294 (1999).
32. Jutz, S. et al. Assessment of costimulation and coinhibition in a triple parameter T cell reporter line: Simultaneous measurement of NF-κB, NFAT and AP-1. *J. Immunol. Methods* **430**, 10–20 (2016).
33. Li, J. et al. Functional phosphoproteomics in cancer chemoresistance using CRISPR-mediated base editors. *Adv. Sci.* **9**, e2200717 (2022).
34. Pihlajamaa, P., Kauko, O., Sahu, B., Kivioja, T. & Taipale, J. A competitive precision CRISPR method to identify the fitness effects of transcription factor binding sites. *Nat. Biotechnol.* **41**, 197–203 (2023).
35. Quesada, A. E. et al. Clinico-pathologic characteristics and outcomes of the World Health Organization (WHO) provisional entity de novo acute myeloid leukemia with mutated RUNX1. *Mod. Pathol.* **33**, 1678–1689 (2020).
36. Huang, K. et al. Genome-wide CRISPR-Cas9 screening identifies NF-κB/E2F6 responsible for EGFRvIII-associated temozolomide resistance in Glioblastoma. *Adv. Sci.* **6**, 1900782 (2019).
37. Cheng, F. H. C. et al. E2F6 functions as a competing endogenous RNA, and transcriptional repressor, to promote ovarian cancer stemness. *Cancer Sci.* **110**, 1085–1095 (2019).
38. Johnson, J. L. et al. An atlas of substrate specificities for the human serine/threonine kinome. *Nature* **613**, 759–766 (2023).
39. Mognol, G. P. et al. Targeting the NFAT:AP-1 transcriptional complex on DNA with a small-molecule inhibitor. *Proc. Natl Acad. Sci. USA* **116**, 9959–9968 (2019).

40. Wang, B. et al. Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. *Nat. Protoc.* **14**, 756–780 (2019).
41. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
42. Hogan, P. G., Chen, L., Nardone, J. & Rao, A. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev.* **17**, 2205–2232 (2003).
43. Ortega-Pérez, I. et al. c-Jun N-terminal kinase (JNK) positively regulates NFATc2 transactivation through phosphorylation within the N-terminal regulatory domain. *J. Biol. Chem.* **280**, 20867–20878 (2005).
44. Ishitani, T. et al. The TAK1-NLK mitogen-activated protein kinase cascade functions in the Wnt-5a/Ca(2+) pathway to antagonize Wnt/beta-catenin signaling. *Mol. Cell. Biol.* **23**, 131–139 (2003).
45. MacDonnell, S. M. et al. CaMKII negatively regulates calcineurin-NFAT signaling in cardiac myocytes. *Circ. Res.* **105**, 316–325 (2009).
46. Anshabo, A. T., Milne, R., Wang, S. & Albrecht, H. CDK9: a comprehensive review of its biology, and its role as a potential target for anti-cancer agents. *Front Oncol.* **11**, 678559 (2021).
47. Phee, H. et al. Pak2 is required for actin cytoskeleton remodeling, TCR signaling, and normal thymocyte development and maturation. *eLife* **3**, e02270 (2014).
48. Pareek, T. K. et al. Cyclin-dependent kinase 5 activity is required for T cell activation and induction of experimental autoimmune encephalomyelitis. *J. Exp. Med.* **207**, 2507–2519 (2010).
49. Ochoa, D. et al. The functional landscape of the human phosphoproteome. *Nat. Biotechnol.* **38**, 365–373 (2020).
50. Chen, M. et al. Identification of PHLPP1 as a tumor suppressor reveals the role of feedback activation in PTEN-mutant prostate cancer progression. *Cancer Cell* **20**, 173–186 (2011).
51. Nitsche, C. et al. The phosphatase PHLPP1 regulates Akt2, promotes pancreatic cancer cell death, and inhibits tumor formation. *Gastroenterology* **142**, 377–87.e1–5 (2012).
52. Cohen Katsenelson, K. et al. PHLPP1 counter-regulates STAT1-mediated inflammatory signaling. *eLife* **8**, e48609 (2019).
53. Patterson, S. J. et al. Cutting edge: PHLPP regulates the development, function, and molecular signaling pathways of regulatory T cells. *J. Immunol.* **186**, 5533–5537 (2011).
54. Balasuriya, N. et al. Genetic code expansion and live cell imaging reveal that Thr-308 phosphorylation is irreplaceable and sufficient for Akt1 activity. *J. Biol. Chem.* **293**, 10744–10756 (2018).
55. Feske, S. Calcium signalling in lymphocyte activation and disease. *Nat. Rev. Immunol.* **7**, 690–702 (2007).
56. Gwack, Y. et al. A genome-wide Drosophila RNAi screen identifies DYRK-family kinases as regulators of NFAT. *Nature* **441**, 646–650 (2006).
57. Liu, H. et al. NFATc1 phosphorylation by DYRK1A increases its protein stability. *PLoS ONE* **12**, e0172985 (2017).
58. Thompson, B. J. et al. DYRK1A controls the transition from proliferation to quiescence during lymphoid development by destabilizing Cyclin D3. *J. Exp. Med.* **212**, 953–970 (2015).
59. Chen, J. et al. NR4A transcription factors limit CAR T cell function in solid tumours. *Nature* **567**, 530–534 (2019).
60. Beltrao, P. et al. Systematic functional prioritization of protein posttranslational modifications. *Cell* **150**, 413–425 (2012).
61. Beltrao, P., Bork, P., Krogan, N. J. & van Noort, V. Evolution and functional cross-talk of protein post-translational modifications. *Mol. Syst. Biol.* **9**, 714 (2013).
62. Liu, N., Guo, Y., Ning, S. & Duan, M. Phosphorylation regulates the binding of intrinsically disordered proteins via a flexible conformation selection mechanism. *Commun. Chem.* **3**, 1–9 (2020).
63. Nicolaou, S. T., Hebditch, M., Jonathan, O. J., Verma, C. S. & Warwicker, J. PhosIDP: a web tool to visualize the location of phosphorylation sites in disordered regions. *Sci. Rep.* **11**, 9930 (2021).
64. Trinidad, J. C. et al. Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Mol. Cell Proteom.* **11**, 215–229 (2012).
65. Ren, X. et al. High-throughput PRIME-editing screens identify functional DNA variants in the human genome. *Mol. Cell* **83**, 4633–4645.e9 (2023).
66. Hiatt, J. et al. Efficient generation of isogenic primary human myeloid cells using CRISPR-Cas9 ribonucleoproteins. *Cell Rep.* **35**, 109105 (2021).
67. Mari, T. et al. In vitro Kinase-to-Phosphosite database (iKiP-DB) predicts kinase activity in phosphoproteomic fatasets. *J. Proteome Res* **21**, 1575–1587 (2022).

Methods

Cell culture

Jurkat E6.1 and HEK293T cells were purchased through ATCC. TPR Jurkat cells[32] were purchased from P. Steinberger at the Medical University of Vienna. TPR Jurkat cells and E6.1 Jurkat cells were cultured and passaged in Roswell Park Memorial Institute medium 1640 plus GlutaMax (ThermoFisher Scientific) supplemented with 10% heat-inactivated fetal bovine serum. Cells were passaged and maintained at cell densities between 1×10^5 and 5×10^5 cells ml^{-1}. HEK293T cells were cultured and passaged in Dulbecco's modified Eagle medium supplemented with 10% heat inactivated fetal bovine serum. All cells were incubated and maintained at 37 °C with 5% CO_2.

Phosphoproteomic analysis

Jurkat E6.1 cells (ATCC) were activated in 96-well tissue culture plates for the stated times at 2×10^5 cells ml^{-1}. Plates were coated with 3.33 μg ml^{-1} α-CD3 (HITa3) and α-CD28 antibodies (Biolegend), in 100 μl PBS at 4 °C overnight, followed by one cold PBS wash. To stop the reaction, the cells were transferred to ice-cold PBS, washed twice, flash frozen and stored at −80 °C until processing. Phosphoproteomic analysis was performed as previously described[23]. Data were analyzed using Spectrum Mill (Agilent and Broad Institute) for phosphopeptide identification and quantification. The tandem mass tag denominator for each sample was the median of all tandem mass tag channels. After global median normalization and median absolute deviation scaling, a moderated *F*-test (limma, *R*) was performed to identify 'regulated' phosphopeptide levels across the time series. Statistically significant phosphopeptides (Benjamini–Hochberg adjusted *P* value < 0.05) were visualized using Morpheus (https://software.broadinstitute.org/morpheus/). The .json file associated with this manuscript, Supplementary Data File 1, can be used to explore these data in Morpheus.

ABE8e protein expression and purification

Recombinant ABE8e was expressed and purified as previously described[27]. Briefly, ABE8e (Addgene Plasmid catalog no. 161788) was expressed as an 8× His tagged protein from a rhamnose-inducible promoter in BL21-Star DE3 cells with low RNase activity (ThermoFisher Scientific). At an optical density OD_{600} of ~0.8, the 2× terrific broth *E. coli* cultures were cold shocked on ice for 1 h, and induced with 0.8% final concentration of rhamnose. Roughly 24 h later, cells were lysed via lysozyme and sonication, and ABE8e protein was purified on Ni-NTA resin. After imidazole elution, ABE8e protein was further purified using cation exchange. Fractions were monitored by ultraviolet- and SDS–PAGE. ABE8e protein containing fractions were pooled, concentrated to ~90 pmol μl^{-1} using 100 MW cutoff filters, aliquoted, flash frozen and stored at −80 °C.

Base editing with arrayed lentivirus

Oligonucleotides containing the protospacer were chosen by hand[68] or from our original bioinformatic analysis, and ordered from IDT DNA technologies. CACCG was added to the forward oligonucleotide, whereas AAAC was added to the reverse complement of the protospacer sequence. A C on the 3′ end of the reverse complement oligonucleotide was also appended. pRDA118 (Addgene Plasmid calatog no. 133459) was digested with BsbmI_V2 (NEB) for 1 h and FastAP (Thermo) for the last 5 min, followed by gel purification. Oligonucleotides were mixed 1:1 at 100 μM, phosphorylated via polynucleotide kinase (NEB) and annealed after a 5 min 95 °C step at 5 °C per 5 min until 25 °C. Annealed oligonucleotides were diluted 1:200 and 1 μl was mixed with 25–50 ng of digested backbone. T4 ligation (NEB) was performed for 20 min at 37 °C and the 5–10 μl ligation reaction was transformed into Stbl3 *E. coli*, made in house. Colonies were verified via Sanger sequencing. sgRNA plasmids (1 μg) were transfected with Lipofectamine 3000 (Invitrogen) along with 1 μg of pPAX2 and 0.1 μg VSVG into HEK293Ts seeded the night before at 2×10^5 cells per six-well plate. After 3 h, the 2.5 ml medium was replaced with 5 ml medium supplemented with 1% BSA. Supernatants were collected after 3 days, filtered to remove HEK293T cells, concentrated tenfold with Lenti-X precipitation solution (Alstem), and aliquots were flash frozen and stored at −80 °C.

Wild-type or TPR Jurkat cells were spinfected for 2 h at 666*g*; 2 days later, 2 μg ml^{-1} puromycin was added until the nontransduced cells were all dead (2–4 days). To test base editor delivery molecules, only the *HEK3* site-targeting sgRNA was used, and electroporation was performed using Lonza Nucleofection with the SE or P3 nucleofection solution. pCMV-BE4max was used for plasmid DNA. Base editor mRNAs were designed[68] and generated by in vitro transcription using the HiScribe T7 High-Yield RNA synthesis kit (NEB, catalog no. E2040S)[69]. NEBnext polymerase was used to PCR-amplify template plasmids and install a functional T7 promoter and a 120 nucleotide polyadenine tail. Transcription reactions were set up with complete substitution of uracil by N1-methylpseudouridine (Trilink BioTechnologies, catalog no. N-1080) and cotranscriptional 5′ capping with the CleanCap AG analog (Trilink BioTechnologies, catalog no. N-7113) to generate a 5′ Cap1 structure. mRNAs were purified using ethanol precipitation according to kit instructions, dissolved in nuclease-free water and normalized to a concentration of 2 μg μl^{-1} using Nanodrop RNA quantification of diluted test samples.

Base editing with in vitro transcribed guides and ABE8e protein

sgRNAs were transcribed in vitro using the EnGen sgRNA Synthesis kit (NEB) with oligonucleotides containing the protospacers (Supplementary Table 6). The sgRNAs were then purified (NEB Monarch RNA Cleanup kit) and quantified by nanodrop. Purified sgRNA (3 μg) was then mixed with 1 μl of 90 μM ABE8e protein in Lonza P3 electroporation buffer in a total reaction volume of 10 μl and incubated at room temperature to allow ribonucleoprotein complex to form.

TPR Jurkat cells were collected, spun, washed twice with 37 °C PBS and 2×10^5 cells were resuspended per cuvette well in Lonza electroporation buffer P3 with the ABE8e and sgRNA ribonucleoprotein complex to a total reaction volume of 20 μl. Electroporation and cell recovery was performed according to the manufacturer's instructions. At 2 days postnucleofection, genomic DNA (gDNA) of the cells was extracted, and base editing efficiency of each population was determined through PCR amplification of the edited genomic region of interest and quantified through Sanger sequencing and EditR software analysis[28]. To produce purely edited cell populations, single-cell clones were isolated from the corresponding bulk-edited cell populations by diluting 0.8 cells per well in 96-well plates. Once isolated, these cells were grown until confluence, then genotyped individually through extraction of its gDNA, PCR amplification of the edited genomic region of interest and Sanger sequencing via EditR[28]. Once single-cell clones were validated by Sanger sequencing, four to eight single-cell clones were mixed together to avoid clone-specific effects.

CD69 staining

Cells activated as described above were washed once with cold Cell wash buffer (Biolegend) and incubated with 0.5 μl α-CD69-APC antibodies (Biolegend) at 4 °C, wrapped in foil, for 20 min. Cells were washed twice with cold PBS and stored on ice before flow cytometry.

Phosphosite library design and construction

All phosphorylation sites determined through the accompanying phosphoproteomic analysis were filtered for phosphosites that were localized in the peptide with a confidence of 90% or greater. For all genes with phosphosites, we used an in-house base editor design tool to first design all possible guides targeting these genes. All adenines in the window of 4–8 of the sgRNA (where 1 is the most PAM-distal position, and positions 21–23 are the PAM) were considered to be edited. sgRNAs with cloning sites, poly Ts and greater than five perfect matches

in the genome were excluded. sgRNAs targeting the phosphosite of interest were then picked. sgRNAs predicted to make silent edits at phosphosites but also predicted to have bystander edits in the 4–8 window were excluded. We then included 250 nontargeting controls, 250 intergenic controls, 250 controls targeting splice sites of essential genes and 250 controls targeting splice sites of TCR genes. The code is available at https://github.com/mhegde/base-editor-design-tool.

Proteome-wide base editor screen

TPR Jurkat cells were spinfected in 4 μg ml^{-1} polybrene at 30 °C for 45 min at 666*g* at a multiplicity of infection of 0.3, assuming 30% transduction efficiency, and maintained a 500× library coverage. Transduction quadruplicates were used for downstream replicates. Puromycin was added to 2 μg ml^{-1} and cells were selected for stable integrants for 7 days. At 3 days after removal of puromycin, an aliquot of cells corresponding to 500× library coverage was saved as the pre-ABE8e inputs. Cells were washed with 37 °C PBS twice, and resuspended in SE nucleofection reagent (Lonza) with 94 pmol of ABE8e protein per well. Each well was 2×10^5 library-containing TPR cells in 19 μl of SE + 1 μl ABE8e protein, and all 16 wells were used simultaneously for a library coverage of 250×. Electroporation and cell recovery were performed according to the manufacturer's instructions. At 6 days after ABE8e protein introduction, another 500× aliquot of cells was frozen for the post-ABE8e sample.

At 14 days after ABE8e protein introduction, the mutant library TPR Jurkat pool was washed in room temperature PBS, and activated as described above. A 2,000× library coverage of TPR cells was prepared for fluorescence-activated cell sorting (FACS). FACS was performed on a Bigfoot Spectral Cell Sorter (ThermoFisher) and roughly 1×10^6 cells were collected per bin. The top one and bottom two 12.5% bins were sorted. The two bottom bins ('bottom' and 'low') were sorted to avoid unactivated cells, which was always about 15% of all cells. Only the second lowest bin ('low') was used for downstream analyses due to the superior performance of controls, and to ensure cells were activated.

Collected cells were pelleted and stored at −80 °C until further processing. The gDNA was isolated and PCR-amplified for barcode abundance determination as previously described[16]. Standard Illumina adapters were added and stagger regions were introduced for base diversity, according to the protocol 'sgRNA/shRNA/ORF PCR for Illumina Sequencing' (The Broad Institute GPP). Libraries were sequenced at La Jolla Institute for Immunology at a depth of 10×10^6 reads or more.

Differential sgRNA abundance analyses via MAGeCK and MAGeCKFlute programs

To identify the functional phosphorylation sites through mutational analyses, we analyzed raw reads of GFP high and GFP low samples using the MAGeCK program (v.0.5.9.5), used for analyzing CRISPR screens[13]. To summarize the results of MAGeCK's phosphosite and sgRNA data, we used MAGeCKFlute program (v.2.4.0)[40]. Intergenic and nontargeting controls were used for normalization and size factor estimation. Before conducting the pathway enrichment and gene-centric GSEA analysis with MAGeCKFlute, we converted the phosphorylated and mutated gene products to their corresponding gene symbols. For site-centric analyses (PTM-SEA, Kinase Library), we used the MAGeCK output for ssGSEA2.0R package with default settings[24]. PTM-SEA included iKiP data[67] but excluded the LINCS P100 terms for clarity. To investigate the putative kinases responsible for phosphorylating phosphosites abundant in post-base editing and NFAT signaling, we employed The Kinase Library Enrichment analysis, which offers an atlas of primary sequence substrate preferences for the human serine/threonine kinome, was used on the PhosphositePlus.org website[38].

Data visualization was done using R v.4.3.0, Graphpad Prism v.10, Protigy v.1.1.7 and iDEP v.0.96.

Dataset comparisons

To compare the mass spectrometry-based phosphoproteomic data with the phenotypic screen results, we used the *F*-statistic calculated during the moderated *F*-test as our variable. We reasoned that since we had multiple time points, where multiple pairwise comparisons could be made, the *F*-statistic would capture the magnitude and reproducibility of a phosphopeptide's abundance over time. Directionality is lost, although directionality may differ between any two different timepoint comparisons.

To compare predicted functional scores, we downloaded the Supplemental Tables from Ochoa et al.[49] (Supplementary Table 3) and compared the column 'functional_score' with the $\log_2$ FC of the base editor screens calculated by MAGeCK.

PHLPP1 localization

Plasmids encoding the NTE of PHLPP1 were HA tagged[52], and the appropriate codons were mutated using site-directed mutagenesis (Agilent). HeLa cells were transiently transfected with wild type, NLS mutant, S118P, S118A or S118E PHLPP1 NTEs and, after 2 days, were fixed and stained with anti-HA antibody(1:500 dilution) (Cell Signaling, 3742), Alexa Fluor-647-phalloidin (Invitrogen, catalog no. A22285), and 4,6-diamidino-2-phenylindole stains. Images were acquired using spinning disk confocal microscopy. Quantification was performed using individual cells and was tested statistically using a one-way analysis of variance with Tukey's multiple test corrections.

Transcriptional profiling of phosphosite-mutant T cell lines

The various phosphosite mutants were introduced into TPR Jurkat cells via in vitro transcribed sgRNAs described above. After single-cell cloning, four to eight single-cell clones were mixed together to avoid clone-specific effects. Cells were incubated and activated with α-CD3/CD28 agonist antibodies as described above, except that the antibody concentrations were 1 μg ml^{-1}. Cells were activated for 0 and 6 h.

Cells were stained with BioLegend TotalSeq-C Human Universal Cocktail and anti-human hashtag antibodies according to NYCG CITEseq protocols (http://cite-seq.com/) and processed using 10x Genomics 5′ HT with Feature Barcode assay according to protocol, with the addition of selective transcript removal using Jumpcode Genomics CRISPRclean Single-Cell Boost kit.

Briefly, 100,000 cells from each phosphosite-mutant line per timepoint were blocked with BioLegend Human TruStain FcX Fc receptor blocking solution. Cells were then stained with BioLegend TotalSeq-C Human Universal Cocktail resuspended in BioLegend cell staining buffer at a concentration of one vial per 500,000 cells and anti-human hashtag antibodies at a concentration of 0.75 μg 1×10^{-6} cells for 30 min at 4 °C. Cells were then washed three times in cell staining buffer. Cell viability and concentration was assessed using the Moxi Go II and Moxi Cyte Viability Reagent containing propidium iodide). Cells were pelleted and resuspended in 0.04% BSA in PBS for a final concentration of 1,300–1,600 cells μl^{-1} and processed using 10× Genomics Next GEM Single-Cell 5′ HT v.2 assay. Gene Expression (GEX) and Cell Surface Protein libraries were constructed according to protocol (CG000424 Rev D) with the following deviation: postligation product of the GEX library was subjected to Jumpcode Genomics CRISPRclean Single-Cell Boost kit following protocol. Libraries were sequenced at a targeted depth of 50,000 reads per cell for GEX and 5,000 reads per cell for Cell Surface Protein on an Illumina NovaSeq 6000 and an Element Aviti. Cell Ranger count v.7.1.0 was used to generate cloupe files for analysis.

Data were analyzed using the Loupe Browser 6.5.0 (10× Genomics). For Fig. 5c, *HEK3*, *MAPK1*, *LCP2* and *PHLPP1* phosphosite-mutant cells were selected for 0 and 6 h postactivation, and differential gene expression ($\log_2$ FC) was determined using local (sample specific) expression. Genes with a $P \leq 0.1$ were determined to be regulated, as recommended by the software. For Fig. 5d, differentially expressed genes were determined using only *HEK3*, *MAPK1* and *PHLPP1* cells,

although the $\log_2$ FC values of *LCP2* terminating edit cells were included in the plot for context. For Fig. 6a, genes were selected from the plot in 5D through their known involvement in T cell signaling. The .json file associated with this analysis, Supplementary Data File 2, can be used to explore these data in Morpheus.

Granzyme B staining

A total of 3.5×10^5 cells were incubated with fixable viability dye before fixation and permeabilization using Ghost Dye UV 450 (Cytek) at a 1:100 dilution, in a total volume of 50 µl, for 15 min in PBS containing 2% FBS (PBS 2%) at 4 °C and then washed once in the same medium at 500*g* for 4 min. Cells were fixed in 50 µl of BD Cytofix during 20 min at 4 °C and then washed in PBS 2% at 500*g* for 4 min. After, cells were fixed and permeabilized using the eBioscience Foxp3/Transcription Factor Staining Buffer Set (ThermoFisher) for 40 min at 4 °C and then washed in permeabilization/wash buffer at 500*g* for 4 min. Expression of granzyme B was evaluated using the Granzyme B- A647 monoclonal antibody (clone GB11, BD biosciences) at a 1:100 dilution in a total volume of 50 µl. Cells were incubated 1 h at room temperature followed by a 4 °C incubation overnight. Cells were washed in permeabilization/wash buffer at 500*g* for 4 min and then resuspended for analysis. Analyses were performed on LSRII cytometer (BD Biosciences). A total of 10,000 events were recorded and data analyses were performed in FlowJo software (Tree Star).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw mass spectrometry data and metadata can be accessed at ftp://MSV000092965@massive.ucsd.edu. Raw RNA sequencing data can be accessed at GEO accession ID GSE244164.

Code availability

The code for the base editor design tool is available at https://github.com/mhegde/base-editor-design-tool.

References

68. Hwang, G.-H. et al. Web-based design and analysis tools for CRISPR base editing. *BMC Bioinform.* **19**, 542 (2018).
69. Chen, P. J. et al. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* **184**, 5635–5652. e29 (2021).

Acknowledgements

We thank A. Haber, P. Vijayanand, E. Kvedaraite, B. Hamilton, A. Rubin, T.M. Yaron and M. Gentili for useful discussion. We also thank G. and S. Clouse and S. Carr for support, as well as P. Guo and the Nikon Imaging Center at the University of California San Diego for the support on microscopy experiments. This work was supported by National Institutes of Health (NIH) grant nos. R35GM147554 and R01CA279795 (S.A.M.); NIH grant no. R35GM122523 (A.C.N.); NIH grant nos. U01AI142756, R35GM118062, RM1HG009490 and HHMI (D.R.L.); NIH grant nos. R01AI040127 and R01AI109842 (P.G.H.); Stem Cell Network Jump Start Award (no. ECR-C4R1-7) for C.G.d.B. who is a Michael Smith Health Research BC Scholar; and the University of California San Diego Graduate Training Program in Cellular and Molecular Pharmacology (grant no. T32 GM007752) and the National Science Foundation Graduate Research Fellowship Program (no. DGE-1650112) (A.C.J.). The NovaSeq 6000 was acquired through the Shared Instrumentation Grant Program (S10) S10OD025052; La Jolla Institute for Immunology Next-Generation Sequencing Core Facility RRID:SCR_023107. FACSAria-3 was acquired through the Shared Instrumentation Grant Program (S10): RR027366; La Jolla Institute for Immunology Flow Cytometry Core RRID:SCR_014832.

Author contributions

S.A.M. conceptualized the study. P.H.K., A.C.J., M.H., C.G.d.B., G.A.N. and S.A.M. developed the methodology. M.H. and J.G.D. wrote the software. P.H.K., A.A.D.S., A.C.J. and S.A.M. validated the results. A.A.D.S., M.E.O., R.B., S.A., R.A.G., G.A.N. and S.A.M. performed the formal analyses. P.H.K., A.A.D.S., M.B., A.C.J., M.E.O., M.I.M., N.P., P.G.H., R.B., A.C.N., S.A., R.A.G., C.G.d.B., G.A.N. and S.A.M. performed the investigations. M.B., N.P., J.L., T.L., P.G.H., D.R.L., J.G.D., G.A.N., C.G.d.B. and S.A.M. provided resources. A.A.D.S., M.E.O., M.H., M.I.M., R.B., J.G.D., S.A., R.A.G. and S.A.M. curated the data. All authors wrote the paper. P.H.K., A.A.D.S., A.C.J., M.E.O., S.A. and S.A.M. visualized the findings. S.A.M. supervised the project and was the project administrator. S.A.M. acquired funding.

Competing interests

D.R.L. is a consultant and/or equity owner for Prime Medicine, Beam Therapeutics, Pairwise Plants, Chroma Medicine and Nvelop Therapeutics—companies that use or deliver genome editing or epigenome engineering agents. The other authors declare no competing interests.

Additional information

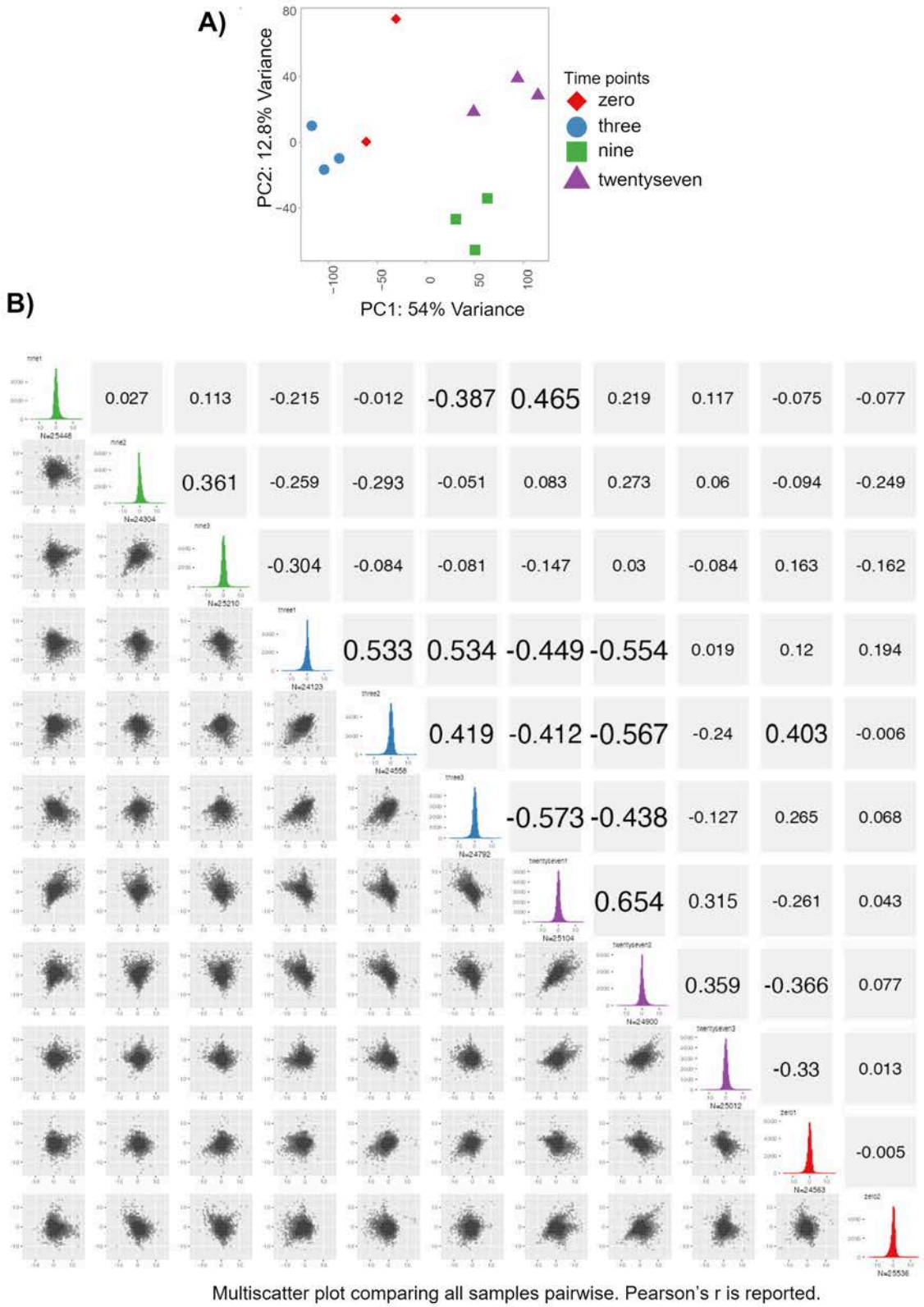
Extended data is available for this paper at https://doi.org/10.1038/s41592-024-02256-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-024-02256-z.

Correspondence and requests for materials should be addressed to Samuel A. Myers.

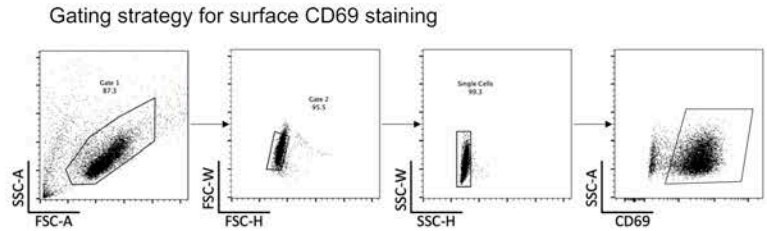
Peer review information *Nature Methods* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Rita Strack, in collaboration with the *Nature Methods* team.

https://doi.org/10.1038/s41592-024-02256-z



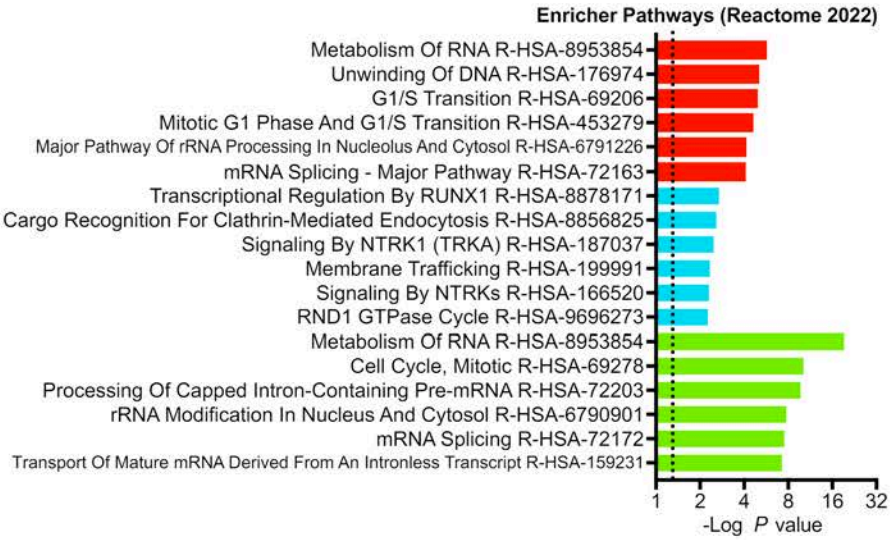
Extended Data Fig. 1 | associated with Fig. 1 Kennedy et al. Phosphoproteomics quality controls. **a)** Principal component analysis of all phosphoproteomics samples prior to differential expression analysis. **b)** Multi-scatter plot comparing all samples to each other pairwise. Pearson's r is shown. Colors of samples are the same as in Extended Data Fig. 1a.

https://doi.org/10.1038/s41592-024-02256-z



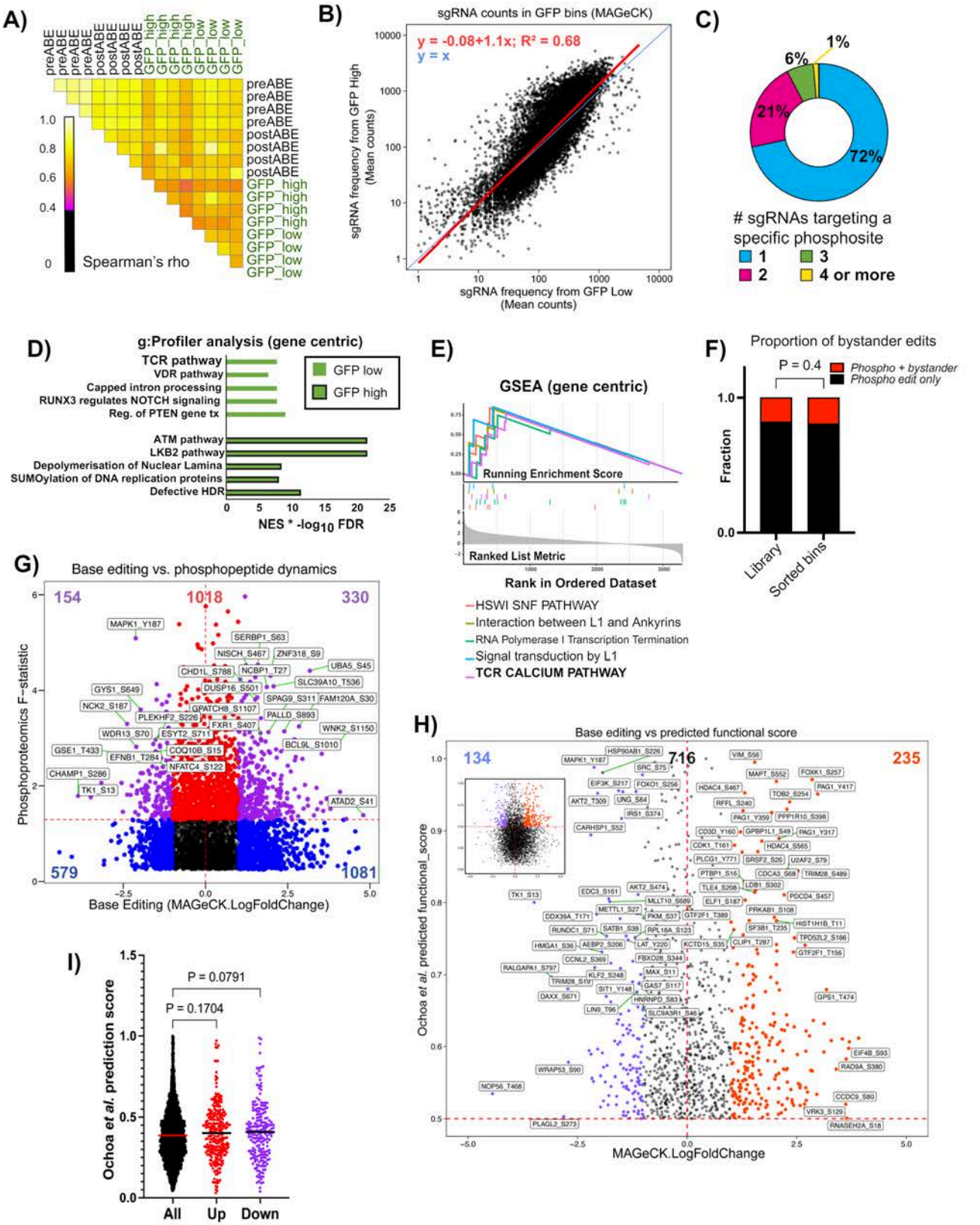
Extended Data Fig. 2 | associated with Fig. 2 Kennedy et al. Gating strategy for CD69 staining analyzed by flow cytometry.

https://doi.org/10.1038/s41592-024-02256-z



Extended Data Fig. 3 | associated with Fig. 3 Kennedy et al. Gene ontology analysis of the genes targeted by the sgRNAs depicted in Fig. 3d. Colors coordinate with Fig. 3d. Dotted line is the hypergeometric distribution test FDR threshold.

https://doi.org/10.1038/s41592-024-02256-z

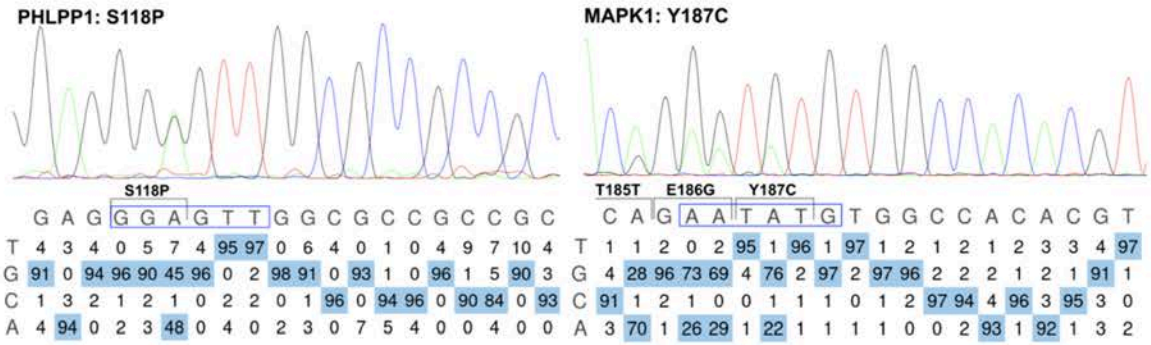


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41592-024-02256-z

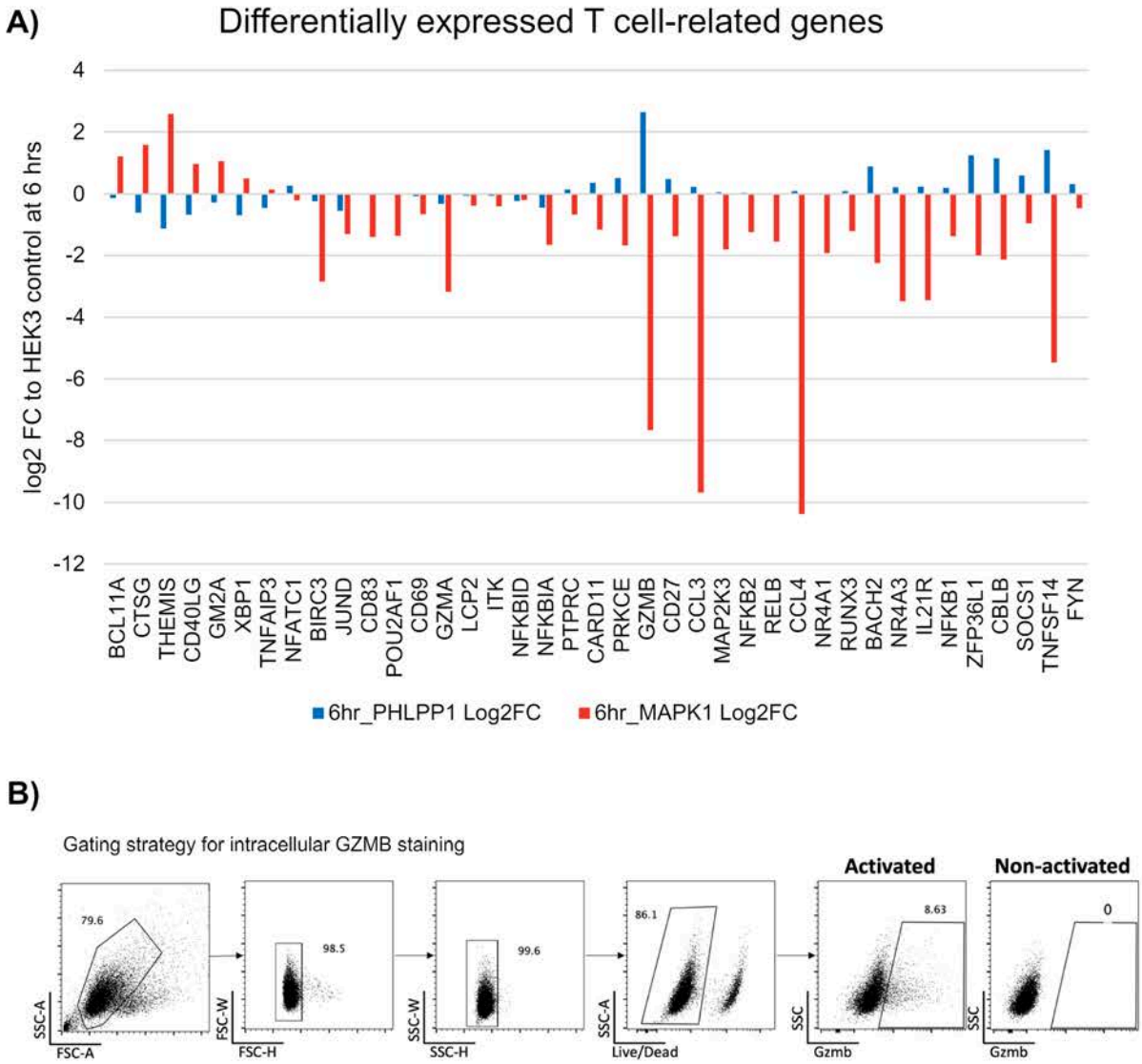
Extended Data Fig. 4 | associated with Fig. 4 Kennedy et al. Quality control and characterization of phosphosite base editing coupled to NFAT activity reporters. **a)** Pairwise Spearman correlations between all normalized log transformed read counts across replicates and experimental conditions. 0.4 is the lower limit cut off in black. **b)** Mean (across replicates) sgRNA counts for individual sgRNAs prior to collapsing redundant phosphosite targets in the GFP high and low bins. Regression line is shown. **c)** Percentage of phosphosite targets with one or more protospacer sequences. **d)** g:Profiler (gene-centric) analysis of genes with phosphosite mutations enriched in the GFP low or GFP high bins. For the x-axis the normalized enrichment score (NES) was multiplied by the -log10 FDR. **e**) GSEA (gene-centric) analysis of gene sets enriched in the GFP high bin. TCR Calcium Pathway is bolded. **f)** Proportion of phosphosite targets that contain a putative bystander edit in the library as a whole and in the sorted GFP bins. Student's two sample T test P value is shown. **g)** Scatterplot comparing the F statistic from the phosphoproteomic analysis, a proxy for magnitude and reproducibility of abundance changes across the four time points, and the log2 fold change GFPhigh/low bins calculated by MAGeCK. Horizontal red dashed line delineates nominal p value of < 0.05 from the moderated F test of the phosphoproteomics data. **h)** Scatterplot comparing the log2 fold change GFPhigh/low bins calculated by MAGeCK to the predicted functional score from the machine learning analysis in Ochoa et al.[49]. Inset shows the full data structure while the scatter plot is a zoom of points above a predicted functional score of 0.5. Horizontal red dashed line delineates a score threshold determined in Ochoa et al.[49]. **i)** Distribution of predicted functional scores from Ochoa et al.[49] for all data points in the GFP screen, the phosphosite mutants that increased ('up' in red) or decreased GFP levels ('down' in purple). P values for comparison to the whole data set are shown. Data points represent the mean log2 FC (GFPhigh/GFP low) of four transduction replicates. P values for an ANOVA test followed by uncorrected Fisher's least significant difference for multiple comparisons.

https://doi.org/10.1038/s41592-024-02256-z



Extended Data Fig. 5 | associated with Fig. 5, Kennedy et al. EditR software analysis[28] plots outlining bystander base editing levels for PHLPP1 S118P and MAPK1 Y187C prior to single cell clone isolation.

https://doi.org/10.1038/s41592-024-02256-z



Extended Data Fig. 6 | associated with Fig. 6, Kennedy et al. a) Log2 fold change of select T cell genes differentially expressed between PHLPP1 S118P and MAPK1 Y187C mutant cells, compared to *HEK3* control cells. **b)** Gating strategy for intracellular GZMB staining and analysis by flow cytometry.

nature portfolio

Corresponding author(s): Samuel A. Myers

Last updated by author(s): Feb 27, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Cellranger count v7.1.0 was used to generate cloupe files for analysis
Data analysis	Phosphoproteomic data was analyzed using Spectrum Mill (Agilent and Broad Institute) for phosphopeptide identification and quantification. Statistically significant phosphopeptides (Benjamini-Hochberg adj. p value <0.05) were visualized using Morpheus (https://software.broadinstitute.org/morpheus/) (Version: 0.1.1.1). EditR software analysis was used to analyze sanger sequencing trace data from the amplification of the Phospho-mutant T cells genomic region of interest (Version 1.0.10). For all genes with phosphosites, we used an in-house base editor design tool to first design all possible guides targeting these genes (https://github.com/mhegde/base-editor-design-tool). To identify the functional phosphorylation sites through mutational analyses, we analyzed raw reads of GFPhigh and GFPlow samples using the MAGeCK program (v0.5.9.5). To summarize the results of MAGeCK's phosphosite and sgRNA data, we used MAGeCKFlute program (version 2.4.0). For site-centric analyses (PTM-SEA, Kinase Library), we used the MAGeCK output for ssGSEA2.0 R package with default settings. RNA sequencing data was analyzed using the Loupe Browser (10X Genomics). Flow cytometry data was analyzed using FlowJo(Tree Star, Ashland, OR) (Version: 10.8.1). Data visualization is done using R 4.3.0, Graphpad Prism 10, Protigy (v1.1.7), iDEP 0.96.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw mass spectrometry data and metadata can be accessed at ftp://MSV000092965@massive.ucsd.edu . Raw RNA sequencing data can be accessed at GEO accession ID GSE244164, which are publicly available.
We employed The Kinase Library Enrichment analysis, which offers an atlas of primary sequence substrate preferences for the human serine/threonine kinome, was used on the PhosphositePlus.org website.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculations were necessary to be performed. The targeted phosphosites were derived through our mass spectrometry data and all targetable phosphosites were included in our sgRNA library. We made an effort to perform replicates of experiments to show reproducibility.
Data exclusions	No data was excluded from the analyses of experiments described in the manuscript.
Replication	All experiments were performed in at least duplicates, with experiments done in triplicates where shown.
Randomization	Since internal controls were present, and much of the study was screen or NGS based, randomization was not a necessary part of our study
Blinding	Due to the fact that analyses were quantitative and internal controls were present, blinding was not necessary part of our study

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: anti-human CD3 antibody (Biolegend, clone HITa3, Cat# 300301), anti-human CD28 antibody (Biolegend, clone CD28.2, Cat# 302901), APC Mouse Anti-Human CD69 antibody (BD Biosciences, clone FN50, Cat#555533, Lot#1285787) , anti-HA antibody (Cell Signaling, 3742) and Alexa Fluor-647-phalloidin (Invitrogen, A22285).

Validation: For the Biolegend anti-human CD3 and anti-human CD28 antibodies, each lot was quality tested by immunofluorescent staining with flow cytometric analysis as outlined on the manufacturer's website. Per manufacturer's website, the APC Mouse Anti-Human CD69 antibody is routinely tested via flow cytometry. Regarding Cell Signaling Technology products, per their product perfomance guarantee "To ensure product performance, we validate all CST® antibodies, assay kits, and reagents to ensure optimal performance in the approved applications shown on our product web pages". In addition, a certificate of analysis for the Alexa Fluor-647-phalloidin probe indicated lot data and specifications.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s): Jurkat E6.1 and HEK293T cells were purchased through ATCC. Triple parameter reporter (TPR) Jurkat cells36 were purchased from Professor Peter Steinberger at the Medical University of Vienna

Authentication: Phospho-mutant Jurkat (TPR) cell lines were genotyped and authenticated through extraction of its gDNA and PCR amplification of the edited genomic region of interest followed by EditR software analysis. Remaining cell lines were not independently authenticated in this study.

Mycoplasma contamination: All cell lines tested negative for mycoplasma contamination

Commonly misidentified lines (See ICLAC register): No commonly misidentified cell lines were used in the study.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation: Sample preparation is described in detail in the methods section for each experiment.

Instrument: Analyses were performed on LSRII cytometer (BD Biosciences) and cell sorts were performed on a Bigfoot Spectral Cell Sorter(ThermoFisher).

Software: FACSDiva was used for data collection and FlowJo(V10.8.1) was used to analyze the data.

Cell population abundance: 10,000 events were recorded for analysis.

Gating strategy: Jurkat T cells were first gated based on cell size (FSC) and complexity (SSC) and then gated based on live/dead staining. TFor Granzyme b staining, the % of Granzyme b positive cells were assessed on the live population.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.