

Title

Mapping the nanoscale organization of the human cell surface proteome reveals new functional associations and surface antigen clusters

Brendan M. Floyd^{1,5}, Elizabeth L. Schmidt¹, Nicholas A. Till¹, Jonathan L. Yang¹, Pinyu Liao¹, Benson M. George², Ryan A. Flynn^{2,3}, and Carolyn R. Bertozzi^{1,4,5}

¹ Sarafan ChEM-H and Department of Chemistry, Stanford University, Stanford, CA, USA

² Stem Cell Program and Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA

³ Harvard Stem Cell Institute, Harvard University, Cambridge, MA, USA

⁴ Howard Hughes Medical Institute, Stanford, CA, USA

⁵ Lead contact

Abstract

The cell surface is a dynamic interface that controls cell-cell communication and signal transduction relevant to organ development, homeostasis and repair, immune reactivity, and pathologies driven by aberrant cell surface phenotypes. The spatial organization of cell surface proteins is central to these processes. High-resolution fluorescence microscopy and proximity labeling have advanced studies of surface protein associations, but the spatial organization of the complete surface proteome remains uncharted. In this study, we systematically mapped the surface proteome of human T-lymphocytes and B-lymphoblasts using proximity labeling of 85 antigens, identified from over 100 antibodies tested for binding to surface-exposed proteins. These experiments were coupled with an optimized data-independent acquisition mass spectrometry workflow to generate a robust dataset. Unsupervised clustering of the resulting interactome revealed functional modules, including well-characterized

complexes such as the T-cell receptor and HLA class I/II, alongside novel clusters. Notably, we identified mitochondrial proteins localized to the surface, including the transcription factor TFAM, suggesting previously unappreciated roles for mitochondrial proteins at the plasma membrane. A high-accuracy machine learning classifier predicted over 6,000 surface protein associations, highlighting functional associations such as IL10RB's role as a negative regulator of type I interferon signaling. Spatial modeling of the surface proteome provided insights into protein dispersion patterns, distinguishing widely distributed proteins, such as CD45, from localized antigens, such as CD226 pointing to active mechanisms of regulating surface organization. This work provides a comprehensive map of the human surfaceome and a resource for exploring the spatial and functional dynamics of the cell membrane proteome.

Introduction

Proteins located on the plasma membrane and exposed to the extracellular space, collectively known as the surfaceome, are critical for a variety of cellular processes such as immune cell recognition of damaged cells and formation of complex neural circuits¹. Proteins of the surfaceome provide the primary means of cellular response to extracellular cues and stimulation with different ligands^{2,3}. Aberrations in the expression and arrangement of cell-surface proteins has been linked to various disease states, including breast⁴, colorectal⁵, and pancreatic cancers⁶. These disease-related changes are often associated with alterations in signaling pathways that depend on protein interactions formed at the surface. The plasma membrane is thought to consist of spatially grouped proteins and lipids that form micro- and nanodomains⁷. These domains are considered hubs for cooperative signaling among protein constituents. To date, studies of functional domains on the cell surface have primarily been conducted in either reconstituted systems that lack native cellular context or in individual units, limiting

our understanding of the cell surface as an integrated system. A comprehensive map of these functional domains on the cell surface remains uncharted.

Recently developed proximity labeling (PL) proteomics methods have been used to probe cell surface protein-protein interactions^{9,10}. These methods deploy chemical probes that are converted to reactive intermediates in the presence of a localized catalyst, typically an enzyme or a photocatalyst. The short half-life of the reactive intermediate limits its diffusion distance, leading to spatially-constrained protein labeling in proximity to the catalyst. Several enzymes have been employed as PL catalysts, including the engineered ascorbate peroxidase APEX2¹¹, the relatively promiscuous horseradish peroxidase (HRP)¹², and biotin ligase variants such as BioID¹³ and TurboID¹⁴. These methods differ in their labeling radii and reaction times, driving researchers to explore methods with finer control. Photocatalytic methods leverage either Dexter energy transfer (e.g. μ Map) or singlet oxygen generators (e.g., the Eosin Y-catalyzed Multimap and the Rhodamine dye-catalyzed LUX-MS)^{15–18}. Fluorescence microscopy has also been used to profile the surface proteome, with methods like Förster resonance energy transfer (FRET)¹⁹ and super-resolution techniques, such as STED²⁰, dSTORM²¹, and DNA-PAINT²², providing insights into protein clustering and nanodomain organization²³.

Collectively, these approaches have enhanced our understanding of specific protein clusters. Furthermore, proximity labeling methods have been used to systematically map intracellular protein complexes and functional associations related to disease states^{24–26}. By contrast, efforts to define the complete interactome and functional associations of the surfaceome have been limited.

Here, we used PL, mass spectrometry and machine learning methods to construct a high-resolution interactome of the cell surface proteomes of human lymphocytes. We targeted 65 surface antigens on T lymphocyte cell lines with PL probes, identified their interactors, and applied unsupervised clustering of the resulting dataset to reveal functional associations on the T cell surface. This analysis identified a class of previously uncharacterized surface antigens, including a subset of proteins conventionally

identified as mitochondrial. We further developed a supervised machine learning classifier to predict protein associations across the cell surface proteome. Using this classifier, we identified and characterized a novel negative regulatory interaction between the IL-10 receptor beta subunit and the Type 1 Interferon complex. Expanding our analysis, we mapped the cell surface proteome on B lymphoblasts by proximity labeling 20 surface antigens and characterized both conserved protein associations and functional nanodomains specific to these cells. Finally, we used three-dimensional network graphs to model the spatial distribution of surface proteins on T cells and B cells, demonstrating the utility of this dataset for exploring fundamental questions about cell surface biology. This study provides a comprehensive spatial analysis of the cell surface proteome in T lymphocytes and B lymphoblasts, in addition to a combination of protocols, open-source code, and accessible data.

Results

Systematic proximity labeling of a T lymphocyte cell surface proteome

To explore the functional organization and spatial distribution of the T-cell surface proteome, we conducted proximity labeling experiments using wild-type Jurkat cells. We hypothesized that the broad labeling radius (200–300 nm²⁷) of the horseradish peroxidase (HRP)-catalyzed PL reaction would enable the creation of overlapping nanodomains around target proteins. Thus, targeting a critical mass of surfaceome of components would lead to broad coverage across the entire cell surface proteome and produce a spatial map of nearest neighbors. We labeled select target proteins with primary antibodies, then with a secondary antibody conjugated to HRP (**Figure 1A**). In the presence of hydrogen peroxide, HRP converts biotinyl tyramide into a phenoxyl radical with a short half-life (<1 ms), labeling proteins within a 200-300 nm radius of the target protein. After labeling, cells were lysed, the membrane fraction was isolated, and biotinylated proteins were enriched via streptavidin precipitation. We quantified protein abundances using an optimized data-independent acquisition (DIA) mass spectrometry

workflow, incorporating gas-phase fractionation injections to build a comprehensive spectral library^{28,29}.

Fold changes for all detected proteins were then calculated relative to an isotype control sample.

To validate that our workflow accurately enriched neighboring proteins of target proteins, we used an antibody targeting the receptor tyrosine phosphatase CD45, or PTPRC. This experiment enriched 145 surface proteins neighboring CD45. A comparison with a previously published dataset showed an 84% overlap in enriched proteins and fold change values that largely corresponded with each other (**Supplementary Figure 1A, 1B**)¹⁴, demonstrating the reproducibility of this proximity labeling workflow and proteomics approach. Additionally, we employed the high-resolution μ Map approach, using an iridium-catalyst conjugated secondary antibody for CD45-targeted PL. We confirmed that proteins enriched with this method also overlapped with those from the HRP-catalyzed biotinylation approach (**Supplementary Figure 1C**).

With confidence in our experimental workflow, we systematically mapped the cell surface proteome. To identify potential bait proteins, we used a wheat germ agglutinin horseradish peroxidase (WGA-HRP) conjugate to label the Jurkat cell surface proteome with biotin³⁰. Following streptavidin capture, enriched cell membrane proteins were identified using quantitative mass spectrometry. From the resulting list of cell surface proteins, we selected potential bait proteins across a range of protein abundances using the following criteria: (1) previously reported protein function on the cell surface, (2) availability of commercial antibodies against the target protein, and (3) expression of target proteins on other cell types to allow for comparative studies (**Supplementary Figure 1D**). This selection process yielded over 90 surface antigens, for which we screened 94 antibodies for cell surface binding affinity on Jurkat cells. Of these, antibodies against 65 surface antigens showed significant cell surface binding, making them suitable for proximity labeling (**Supplementary Figure 1E, Supplementary Figure 2, Supplementary Figure 13, Table S1**). We then applied these antibodies in our optimized workflow, confirming target antigen enrichment to verify antibody specificity. Including replicates, gas-phase

fractionation spectral library runs, isotype control samples, and antibody-targeted samples, a total of 339 mass spectrometry experiments were conducted with labeled Jurkat cells. The resulting data identified 390 Jurkat cell surface proteins, in good agreement with a previously reported machine learning predicted surfaceome dataset (i.e., 77% overlap) (**Supplementary Figure 1G**). Overall, our final 65-dimensional dataset contains 25,806 fold change enrichment values that define the spatial organization of the Jurkat cell surface proteome (**Table S2**).

Prey-centric analysis reveals functional protein modules on the cell surface

Our dataset, encompassing enrichment profiles from each of the proximity labeling experiments, provided a framework for examining the spatial relationships between proteins on the Jurkat cell surface (**Supplementary Figure 3A**). We first sought to verify that we were capturing known protein-protein interactions. We reasoned those proteins located proximally to each other would share similar enrichment values across the 65 proximity labeling experiments. As expected, known protein interactors including the integrin $\alpha 4 \beta 1$ heterodimer (**Figure 1B**) and CD3 complex subunits (**Figure 1C**) showed high correlation (Pearson coefficients of 0.95 and 0.92 respectively) across the proximity labeling experiments. Additionally, several other dimeric or multi-subunit complexes exhibited strong correlations across experiments, such as the CD98 heterodimer, the MHC class I complex members HLA-A and B2M, and the cell junction proteins Desmoglein 1 and Desmocollin 1 (**Figure 1D, Supplementary Figure 3A**). These results reinforced the reliability of our surfaceome map in accurately capturing known protein interactions.

Next, we aimed to assess the spatial resolution of our dataset. We benchmarked pairwise Pearson correlation coefficients of the subunits of the CD3-T-cell receptor complex by comparing them to actual distances determined from the center of mass of a 3.7 Å cryo-EM structure (PDB 6JXR)³¹. A significant correlation was observed between the pairwise Pearson coefficients and the true distances between

subunits, demonstrating the systematic proximity labeling approach's ability to discern subunit arrangements down to approximately 4 nm, notably a much higher resolution than the labeling radius of a single experiment (**Figure 1E, Supplementary Figure 3B**).

We then explored whether functional modules could be identified on the cell surface. To achieve this, we used a hierarchical clustering approach previously employed to analyze multimeric protein complexes in plants³² and red blood cells³³. In short, this approach relies on Euclidean distance-based clustering of proteins using their enrichment profiles in the 65-dimensional dataset. Rather than applying a single cutoff to define hierarchies, multiple cutoffs are used to define discrete modules of spatially co-localized proteins³². For each cluster, we identified enriched functional annotations using the DAVID web server^{34,35} and determined whether the surface antigen had been previously characterized on the cell surface (**Figure 1F, Supplementary Figure 3C, Table S3**). This analysis uncovered several functional modules, with a notable presence of cell adhesion-related clusters. Multi-subunit complexes, including the T-cell receptor alpha and beta subunits along with CD3 complex members, clustered with other adaptive immunity regulators, such as the leukocyte associated inhibitory receptor LAIR1 and the MHC class 1 subunits HLA-A and HLA-B. Interestingly, cell junction associated proteins normally involved in intercellular adhesion such as Desmoglein 1 and Desmocollin 1 clustered closely with proteins associated with secreted vesicles and extracellular exosomes, possibly pointing to a location of material exchange between cells. In all, this analysis grants a high-level view of the functional compartmentalization on the cell surface.

Conservation of cell surface protein organization across cell types

To explore differences and similarities in surface protein arrangement across cell types, we adapted our proximity labeling and surface mapping approach for Daudi cells, a B-lymphoblast line. Daudi cells offer a distinct perspective as they express B-cell-specific proteins like CD19, CD20, and CD22, while also sharing some surface markers with Jurkat cells, allowing for comparative analysis. We

proximity labeled 20 surface antigens on Daudi cells, including nine B-cell-specific antigens and 11 antigens shared with Jurkat T-lymphocytes (**Figure 2A, Supplementary Figure 4A, Table S4**). Using criteria similar to those applied to our Jurkat dataset but using a lower experimental enrichment threshold for inclusion due to the smaller amount of Daudi proximity labeling experiments, we identified and mapped 189 proteins on the Daudi cell surface (**Supplementary Figure 4B, 4C**). We then applied a prey-centric analysis to the Daudi cell surface proteome proximity labeling dataset and for validation we examined the correlation of known protein interactors across runs. Among the most correlated pairs were the two subunits of the Human Leukocyte Antigen DR isotype (HLA-DR) complex (**Figure 2B**). Additionally, the Ig α (CD79A) and Ig β (CD79B) subunits of the B-cell receptor complex demonstrated high pairwise correlations (**Figure 2C, Supplementary Figure 4D**).

Next, we analyzed the overall similarities in pairwise protein interactions between Jurkat and Daudi cells. From the cell surface proteomes of both cell types, we identified 71 shared proteins (**Figure 2D**). These proteins included functional units essential for various immune cell types, such as cell adhesion and immune-related proteins. Using this subset of shared proteins, we calculated pairwise Pearson correlation values for all protein pairs and compared them between the two cell lines (**Figure 2E**). This comparison revealed strong correspondence between the 2,278 pairwise associations, with a significant Spearman rho value of 0.516. Focusing on interactions with nearly identical Pearson values (within 0.05) between the two cell types, we observed a striking prevalence of receptor interactions (**Figure 2F**). Finally, we clustered the Daudi cell surface proteome, highlighting functionally enriched annotations for each cluster (**Figure 2G, Supplementary Figure 4E**). This clustering captured a cell junction protein cluster similar to that observed on Jurkat cells in addition to a cluster of immunity-related proteins, including the B-cell receptor accessory proteins CD72 and CD19. Notably, the MHC class II HLA-DR complex subunits formed their own distinct cluster.

Discovery of clusters of novel cell surface mitochondrial proteins

Among the identified and clustered surface proteins on Jurkat cells, 72% had been previously characterized on the cell surface. However, we observed a distinct cluster of proteins typically assumed to be restricted to mitochondria. These mitochondrial proteins constitute approximately 10% of our observed surface proteins on Jurkat cells (**Figure 3A**). We also observed chaperone proteins such as the endoplasmic reticulum-localized HSP90B1³⁶ and RNA binding proteins such as the nuclear RNA binding-protein FUS²². We refer to these non-canonical cell surface proteins as “maverick proteins”.

We confirmed that maverick proteins are indeed present on live Jurkat cell surfaces using flow cytometry. We confirmed cell surface expression of the mitochondrial protein transcription factor A (TFAM), which exhibited substantial peptide coverage in our proximity labeling experiments (**Supplementary Figure 5A**), as well as the nuclear RNA-binding protein FUS which was observed both on Jurkat and Daudi cells²² (**Supplementary Figure 5B, 5C**). To confirm antibody specificity for TFAM, we generated a Jurkat TFAM knockout line using CRISPR-Cas9 (**Supplementary Figure 5D**) and again probed for cell surface TFAM using flow cytometry. As expected, the knockout line showed diminished TFAM antibody binding (**Figure 3B, Supplementary Figure 5E, 5F**). Further, we profiled the interactome of cell surface TFAM (csTFAM) via the HRP-mediated PL method (**Figure 3C**). We observed enrichment of 58 proteins (relative to an isotype control), which included TFAM itself as well as other maverick proteins such as FUS and another RNA-binding protein, CIRBP (**Table S2**). We then extended this analysis to 13 additional cell lines as well as human primary blood mononuclear cells (PBMCs). Cell-surface TFAM expression was observed in several cell lines and was particularly high in the U-937 monocytic cancer line (**Figure 3D**). We also detected cell surface TFAM on neutrophils from human PBMCs (**Figure 3E**).

Having identified TFAM as a *bona fide* surface antigen, we sought to understand its spatial relationship to other mitochondrial maverick proteins. We observed high pairwise correlation values among TFAM and these other proteins, particularly with ATP synthase subunits, aligning with prior

reports of ATP synthase translocation to the plasma membrane in cancer cell lines (**Figure 3F**)^{37,38}.
Notably, all detected mitochondrial proteins on the Jurkat cell surface originated from the mitochondrial inner membrane (MIM) or matrix, with none from the outer membrane (**Figure 3G**). GO enrichment analysis of this cluster highlighted significant roles in cellular respiration, suggesting a potential adaptation for supplemental energy acquisition as previously observed with translocated cell-surface ATP synthase³⁷ (**Figure 3H**).

The pathway by which mitochondrial proteins traffic to the cell surface is unclear. However, our analysis suggests that they first transit through mitochondria. We predicted mitochondrial targeting sequences (MTS)³⁹ for this set of surface-localized mitochondrial proteins and reprocessed a subset of our proximity labeling data, allowing semi-tryptic N-termini to capture proteolytically processed peptides. For proteins with high-confidence MTS predictions, 83% showed no detectable peptides derived from the MTS region, implying they had undergone mitochondrial processing before reaching the plasma membrane (**Figure 3I**). This class of maverick cell surface proteins represents a previously uncharacterized category of surface antigens, with potential applications in cancer targeting as previously demonstrated with ATP synthase^{40–42}.

A machine learning classifier for predicting cell surface protein associations

In order to discover novel protein functions based on their interactors, we used a machine learning approach to predict protein associations on the cell surface. We developed a random forest classifier trained to predict protein associations across the cell surface proteome. A random forest classifier was chosen due to its ability to handle high-dimensional data, robustness against overfitting, and interpretability of its feature importance rankings, which are critical for biological datasets with complex data structures⁴³. For classifier training, we used binary protein-protein interaction (PPI) data from multiple sources to compile a set of gold-standard PPIs. From this set, 1,532 PPIs comprised pairs of

proteins that were both identified on the surface of Jurkat cells. To prepare the data for classification, we calculated fold change differences and similarity metrics (e.g., Pearson correlation, Euclidean distance) for all possible protein pairs, creating a comprehensive feature matrix for the 60,031 possible protein pairs (**Figure 4A**). We then trained our random forest classifier on this feature matrix using the gold-standard cell surface PPIs as positive labels. With further optimization of the classifier's hyperparameters we were able to create a high accuracy classifier with an average ROC AUC of 0.92 (**Figure 4B**) and a precision-recall AUC of 0.62 when validated on a set of gold-standard PPIs withheld from training. (**Figure 4C, Supplementary Figure 6A**). We then predicted cell surface associations across the proteome by putting the full set of all-by-all protein pairs through the classifier, which resulted in 6,003 high confidence cell surface protein association predictions (**Supplementary Figure 6B, Table S6**).

To further validate the classifier predictions, we compared the predicted association scores to the estimated Pearson correlation coefficients and observed that higher predicted association scores corresponded to stronger Pearson correlations (**Figure 4D**). As well, comparing our random forest predicted association scores for bait-prey pairwise interactions to the experimentally measured fold change enrichment for our 65 different baits, higher association scores had significantly higher log fold change values compared to interactions with a low association score (**Figure 4E**). Confident that our machine learning classifier was providing high accuracy association predictions, we then looked at known and novel associations predicted for the Jurkat proteome. Known cell surface interactions included those of the TCR subunits (**Figure 4F**). Additionally, we observed high connectivity of the integrin alpha and beta subunits (**Figure 4G**).

Prediction and characterization of IL10RB as a negative regulator of Type I Interferon signaling

Among the 6,003 high confidence associations predicted by our machine learning classifier, the interaction between the beta subunit of the IL-10 receptor complex, IL10RB, and a subunit of the Type I

interferon system, IFNAR2, stood out. IL10RB was predicted to associate with IFNAR2 with an association score of 0.82, placing it in the top 2% of all scored associations (**Figure 5A**). This prediction was particularly intriguing, as IL10RB is a known signaling component for various class II cytokines, including IL-22, IL-26, IL-28, and Type III interferons^{44,45}. However, it has not previously been associated with the Type I interferon system. While transcripts for IL10RB co-receptors, such as IL22RA1 and IFNLR1, have been previously detected in Jurkat cells⁴⁶, no protein-level evidence of their translation was detected by mass spectrometry in both whole cell and surface protein enriched samples (**Supplementary Figure 7A**). The lack of canonical interaction partners for IL10RB on the Jurkat cell surface, coupled with its broad involvement in signaling pathways, prompted us to investigate further the classifier prediction of its association with IFNAR2. Interestingly, IL10RB also shares structural similarities with IFNAR1, another core component of the Type I interferon receptor complex including, importantly, an intracellular binding domain for the Jak kinase family member Tyk2 (**Supplementary Figure 7B**)^{47,48}.

To explore the functional relevance of the IL10RB-IFNAR2 association, we generated an IL10RB knockout (KO) Jurkat cell line using CRISPR-Cas9. As IL10RB and IFNAR2 are located in cis on chromosome 21, we verified that IFNAR2 expression remained intact in the knockout line (**Supplementary Figure 7C**). Type I interferon signaling is mediated through the JAK1/STAT pathway and involves STAT1 and STAT2 phosphorylation, leading to heterodimerization and complex formation with IRF9, which then promotes transcription and activation of interferon-stimulated genes (ISGs) (**Figure 5B**). In the IL10RB KO line, we observed significantly increased levels of phosphorylated STAT1 compared to the wild-type (WT) line when stimulated with the Type I interferon IFN α 2 (**Figure 5C**, **Supplementary Figure 7D**). This was corroborated by intracellular flow cytometry, which showed elevated IRF9 levels following 12-hour stimulation with IFN α 2 (**Figure 5D**). To confirm the specificity of this effect, we generated a T-cell receptor alpha subunit knockout line (TCR- α KO), which was not predicted to

associate with the Type I interferon system (**Supplementary Figure 7E, 7F**). When stimulated with IFN α 2 the TCR- α KO line exhibited no changes in interferon signaling relative to WT cells when monitoring IRF9 levels using intracellular flow cytometry, underscoring the specificity of the IL10RB-IFNAR2 interaction (**Supplementary Figure 7G**). Finally, we evaluated the dynamics of ISG stimulation in the IL10RB KO line using time-course proteomics. Upon stimulation with 10 ng/mL IFN α 2, we monitored protein abundance changes in WT and IL10RB KO cells using DIA proteomics (**Figure 5E, Table S5**). Of 98 well-sampled ISGs, 35 exhibited increased stimulation rates in the IL10RB KO line, including the antiviral factor IFIT3, the microtubule aggregator IFI44, and the anti-apoptosis regulator XAF1 (**Figure 5F–I**). Additional proteins such as DDX60 also showed interferon-induced upregulation, further supporting the enhanced sensitivity of the IL10RB KO line to Type I interferon stimulation (**Supplementary Figure 7H**). However, it should be noted that interferon-induced GTP-binding protein MX2 exhibited a significant decrease in activation in the IL10RB KO line, possibly highlighting an alternative regulatory mechanism for this ISG in the IL10RB KO cell line (**Supplementary Figure 7I**)⁴⁹.

The functional association between IL10RB and the Type I interferon receptor further validated the machine learning classifier predictions. Based on these findings, we hypothesize that IL10RB, when co-expressed with components of the Type I interferon system, spatially localizes near IFNAR2. When the Type I interferon system is activated, IL10RB modulates its activity by competing with IFNAR1 for occupancy of Tyk2, the cytosolic Jak kinase known to bind to both IFNAR1 and IL10RB (**Figure 5J**). In the absence of IL10RB, freed Tyk2 will be more likely to bind to the Tyk2 binding domain on IFNAR1⁵⁰. We additionally hypothesize that the close localization of IFNAR2 to IL10RB during Interferon activation is resulting in a spatially restricted competition of IFNAR1 and IL10RB for Tyk2 occupancy.

Estimation of cell surface distribution of the cell surface proteome

We next investigated whether our proximity labeling dataset could be used to construct a virtual image of the surface distribution of individual proteins. Such visualization provides insights into the spatial organization of the cell membrane and the mechanisms governing surface protein distributions. Inspired by recent applications of DNA nanotechnology for mapping the spatial arrangement of a limited number of surface proteins⁵¹, we constructed a Kamada-Kawai spring network based on enrichment connections among 48 high-quality bait proteins on Jurkat cells. The network coordinates were transformed into a three-dimensional (3D) spherical representation, enabling us to project the normalized fold change of a protein of interest across the bait network onto this 3D model. This approach effectively models the distribution of proteins across the cell surface proteome using a metric we term the graph connectivity score. Applying this method to the entire Jurkat cell surface proteome, we estimated the surface coverage of each protein by calculating its relative abundance across the modeled cell surface (**Figure 6A**).

To quantify protein distribution, we developed the connectivity score, a metric that integrates a protein's enrichment across the bait network (**Table S7**). Coverage scores varied widely across the proteome, with proteins such as the pan-lymphocyte marker CD45 (coverage score = 0.719) exhibiting widespread distribution, while others, like the tetraspanin web-associated protein CD53 (coverage score = 0.227), displayed more restricted distributions. To determine whether our graph connectivity score was accurately predicting protein cell surface distribution we employed immunofluorescence microscopy to image the distribution of antigens on the cell surface. We first modeled the surface proteins CD45 (**Figure 6B, Movie S1**), CD81 (**Figure 6C, Movie S2**), and CD226 (**Figure 6D, Movie S3**), which represent a range of connectivity scores. Using confocal microscopy, we then imaged the distribution of these surface antigens and after staining cells with antibodies for these antigens, we measured median fluorescence intensity along the cell periphery to estimate their surface distributions. The microscopy data for CD45 (**Figure 6E, Supplementary Figure 8A**), CD81 (**Figure 6F, Supplementary**

Figure 8B), and CD226 (**Figure 6G, Supplementary Figure 8C**) closely aligned with the predicted surface distributions, confirming the reliability of our coverage score metric (**Supplementary Figure 8D**). It is worth noting that the graph connectivity score was independent of the protein abundance measured in surface proteomics experiments (**Figure 6H**).

As a practical application of this method, we explored the relationship between surface protein distribution and function. Cell adhesion-related proteins were broadly distributed across the surface, consistent with their role in mediating multiple attachment points. In contrast, receptors demonstrated a more restricted distribution (**Figure 6I**). Interestingly, cell junction proteins showed reduced dispersion relative to adhesion proteins, suggesting distinct organizational features. This pattern was conserved in Daudi cells, highlighting a shared aspect of surfaceome organization (**Supplementary Figure 8G**).

Discussion

In this work, we systematically mapped the surfaceome of two distinct cell types. By employing a low-resolution proximity labeling approach across 65 surface antigens on T-lymphocytes and 20 antigens on B-lymphoblasts, we assembled a high-resolution interactome for the surfaceome. While the cell membrane is a dynamic environment, and our dataset represents an averaged snapshot that does not fully capture temporal or state-dependent variability, it nonetheless provides valuable insights into surfaceome organization. From these interactomes, we identified functional modules on the cell surface that encompassed both known protein assemblies, such as the T-cell receptor (TCR) and HLA class I and II complexes, as well as novel clusters.

One particularly intriguing finding was the identification of clusters of maverick surface antigens, including proteins typically localized to intracellular organelles. In Jurkat cells, these were primarily mitochondrial proteins, such as the transcription factor TFAM. Reports of mitochondrial proteins, including the ATP synthase complex, maintaining functionality on the plasma membrane suggest a

broader phenomenon of mitochondrial-to-plasma membrane translocation^{37,38,40}. Recent studies describing vesicles derived from the mitochondrial inner membrane, which contain TFAM and other mitochondrial proteins, offer a plausible pathway for their delivery to the cell surface even under homeostatic conditions⁵². Additionally, questions remain regarding how mitochondrial matrix proteins lacking transmembrane domains are tethered to the cell surface, whether via direct interactions with canonical transmembrane proteins⁵³, lipid-anchored proteins, or other attachment mechanisms⁵⁴. Beyond mitochondrial proteins, the recent characterization of RNA-binding proteins on the cell surface, along with the clustering observed among them, suggests a potential conserved role for these proteins in cellular interactions and signaling²². These findings highlight an exciting avenue for future research into both mitochondrial and RNA-binding protein dynamics and their functions at the cell surface.

We developed a high-accuracy machine learning classifier to predict protein-protein associations on the cell surface. While this classifier is adaptable to other proximity labeling datasets, its application to different cell lines or tissues would require the assembly of a cell line-specific gold standard interaction dataset. Moreover, classifier accuracy is influenced by the size of the feature set, with more proximity labeling experiments enhancing performance. Applying this classifier to the Jurkat surfaceome yielded over 6,000 protein associations, encompassing both known and novel interactions. Among these, we identified a functional association between the beta subunit of the IL10 receptor (IL10RB) and a subunit of the type I interferon signaling system. Knocking out IL10RB led to increased interferon signaling upon stimulation, suggesting a role for IL10RB as a negative regulator of Type I interferon signaling. This observation aligns with the established anti-inflammatory role of IL-10/STAT3 signaling and extends it to include regulation of interferon responses, offering new perspectives on the interplay between these pathways⁵⁵.

Mapping surface antigens on two cell types enabled a comparison of both antigen presence and conserved spatial arrangement across cell types. Despite limited conservation of antigens themselves,

interaction networks were highly conserved. Particularly notable was the clustering of receptor ligands on Daudi cells, which included T-cell activation ligands such as ICOS ligand, LFA-3 (CD2 ligand), and TNFSF8 (CD30 ligand). This clustering may represent a regulatory mechanism whereby B cells spatially organize immune activation signals, potentially dispersing them in response to specific cues.

Finally, we leveraged the proximity labeling data to model the spatial distribution of the T cell surface proteome. This analysis revealed surface coverage patterns for antigens, distinguishing highly dispersed proteins, such as CD45, from those with more restricted distributions, such as CD226. Interestingly, our findings indicate that surface protein distribution appears largely independent of protein abundance. This observation suggests the involvement of active mechanisms in the sequestration or spatial regulation of surface proteins, rather than passive diffusion alone. These observations provide a framework for using cell surface PL to profile surface proteins that may be anchored in place, and/or confined within hypothesized cell-surface pickets⁵⁶.

It should be noted that this approach provides a protein-centric view of the cell surface without consideration of other biomolecules, such as lipids and glycans, that play critical roles in membrane structure and function. Nonetheless, given the expected density of approximately 30,000 proteins per μm^2 on mammalian cell membranes⁷, studying the cell surface proteome alone offers a rich and informative snapshot of cell membrane topology. 3D modeling in this format makes the resolution of mapping surface distribution directly dependent upon the amount of bait proteins used for proximity labeling. Depending on the application, a 3D network with as few as 10 bait proteins might be useful for mapping and comparing cell surfaceomes.

Overall, analyzing surface protein distributions using large-scale PL datasets opens new avenues for addressing longstanding questions in membrane biology. This methodology provides an evocative lens through which we can study how surface proteins are organized and distributed, yielding insights into the dynamic and functional architecture of the cell surface.

Limitations of this study

This study presents the first systematic maps of the cell surface proteome for Jurkat (T lymphocyte) and Daudi (B lymphoblast) cell lines. However, a direct comparison to the cell surface proteomes of primary human cells remains an open question. Current requirements for the workflow such as large numbers of cells and extensive antibody panels for proximity labeling pose significant challenges to its application in primary cell lines or small, heterogeneous cell populations. The cell membrane is a dynamic and fluid environment, and the protein associations and functional domains observed in this study represent those present in the majority of cells sampled. While this approach minimizes noise and highlights reproducible features of the surface interactome, it also obscures cellular heterogeneity. For example, variations in surface protein arrangement due to differences in cell cycle stages, activation states, or microenvironmental factors are not captured in this dataset. This limitation underscores the need for future studies to explore single-cell or subset-specific surfaceome mapping techniques to uncover the full diversity of cell surface organization.

Resource availability

Lead contact

Brendan Floyd - bmfloyd@stanford.edu

Carolyn Bertozzi - bertozzi@stanford.edu

Materials availability

Knockout cell lines generated in this study may be made available upon request following institutional approval of a material transfer agreement (MTA).

Data and code availability

Raw data for proteomics files are publicly available on ProteomeXchange with accession number PXD060303. Code used for analysis and visualization in this work is open source and freely available at https://github.com/brendanfloyd/cell_surface_map.

Acknowledgements

We thank Ciaran Seath (UF Scripps) for generous gifts of second and third generation Iridium photocatalysts. We also thank Nicholas Riley, Johannes Hevler, Mariko Morimoto, and David Roberts for helpful discussions. This work was supported in part by the National Institutes of Health grant R01CA200423 (C.R.B.) and R35GM151157 (R.A.F.). B.M.F is a Howard Hughes Medical Institute fellow of the Damon Runyon Cancer Research Foundation (DRG 2518-24). E.L.S is supported by the TIME Initiative Grant and a National Institutes of Health diversity supplement under parent grant R01CA200423.

Author contribution

Conceptualization and Methodology, B.M.F and C.R.B.; Software, B.M.F. and P.L.; Investigation, B.M.F, E.L.S, N.A.T, J.L.Y, P.L, B.M.G, R.A.F.; Formal Analysis and Visualization, B.M.F.; Writing – Original Draft, B.M.F. and C.R.B.; Writing – Review & Editing, B.M.F., E.L.S, N.A.T, J.L.Y, P.L, B.M.G, R.A.F., and C.R.B.; Funding Acquisition, B.M.F., R.A.F., and C.R.B.; Resources, C.R.B. and R.A.F.; Supervision, C.R.B.

Declaration of interests

C.R.B. is a co-founder and scientific advisory board member of Lycia Therapeutics, Palleon Pharmaceuticals, Enable Bioscience, Redwood Biosciences (a subsidiary of Catalent), OliLux Bio, InterVenn Bio, Firefly Bio, Neuravid Therapeutics, and Valora Therapeutics. B.M.G. is a co-founder of Btwo3 Therapeutics and Inograft Biosciences. R.A.F. is a stockholder of ORNA Therapeutics. R.A.F. is a board of directors member and stockholder of Chronus Health and Blue Planet Systems. The other authors declare no conflict of interests.

Materials and methods

Cell lines

The cell lines Jurkat E6-1 (ATCC #TIB-152), Daudi (ATCC #CCL-213), U-937 (ATCC CRL-1593.2), NCI-H82 (ATCC #HTB-175), K-562 (ATCC #CCL-243), Toledo (ATCC #CRL-2631), Raji (ATCC #CCL-86), COLO 205 (ATCC #CCL-222), ZR-75-1 (ATCC #CRL-1500), and A549 (ATCC #CCL-185) were cultured in T75 flasks

(Genesee Scientific #25-209) in media consisting of RPMI-1640 (Thermo Scientific #11875119), 10% fetal bovine serum (FBS) (Thermo Scientific #A5256801), and 1% Penicillin-Streptomycin (Fisher Scientific #sv30010). The cell lines HeLa (ATCC #CCL-2) and HEK-293T (ATCC #CRL-3216) lines were cultured in T75 flasks in media consisting of DMEM (Thermo Scientific #11995073), 10% FBS, and 1% Penicillin-Streptomycin. The NIH:OVCAR3 (ATCC #HTB-161) cell line was cultured in a T75 flask with media consisting of RPMI-1640, 20% FBS, 1% Penicillin-Streptomycin, and 0.01 mg/mL bovine insulin (Sigma-Aldrich #I0516). All cell lines were purchased from the American Type Culture Collection (ATCC) and incubated at 37°C with 5% CO₂.

Knock-Out Cell Line Generation

The CRISPR-Cas9 mediated TFAM, IL10RB, and TRAC knockout Jurkat cell pool was generated using Synthego Gene Knockout Kit v2 (Synthego Corporation). Cells were electroporated with Streptococcus pyogenes Cas9 (Sp-Cas9, Synthego Corporation) precomplexed to sgRNAs (9:1 sgRNA:SpCas9 ratio) targeting TFAM, IL10RB, and TRAC in the SE buffer (Lonza #V4XC-1032) using the Lonza 4D Nucleofector X on the CL-120 setting following Synthego's instructions. Cells were expanded for one week before and for the IL10RB knockout line, IL10RB-positive cells were removed by FACS following staining by anti-Hu IL10RB (S17009F, 1 µg/100 µL, Biolegend #396802) with secondary anti-mouse IgG Alexa Fluor 647 (2 µg/mL, Thermo Scientific #A-21235). sgRNA guides: IL10RB 5'- GGUACUGAGCUGUGAAAGUC -3'; 5'- UAUUUUCAUAGCAUUGGGAA -3'; 5'- AUUUCAAGAACAUUCUACAG -3'. TRAC 5'- CUCUCAGCUGGUACACGGCA -3', 5'- GAGAAUCAAUAUCGGUGAAU -3', 5'- ACAAACUGUGCUAGACAUG -3'. TFAM 5'- UGGCGUUUCUCCGAAGCAUG -3'; 5'- GGCAAGCGGGCCUACCUGAA -3'; 5'- UAGGGUCUACAUUCCAACCC -3'.

Flow cytometry cell surface staining

Cells were cultured in media overnight (37°C, 5% CO₂) in a T75 flask. Cells were moved to 15 mL conical tube and were counted using Trypan Blue (VWR #VWRVK940) and a Countess Cell Counter (Thermo

Scientific). Cells were then diluted to a concentration of 5 million cells/mL in FACS buffer (PBS pH 7.4 and 0.1% w/v bovine serum albumin). 200 μ L of the cell solution (or 1 million cells) was then transferred to a 96-well V bottom plate (Fisher Scientific #0720096). Cells were kept on ice and spun at 4°C for the remainder of the protocol. Cells were then spun at 1700 rpm for 3 min. Supernatant was aspirated and cells were washed 2X with 200 μ L FACS buffer (0.1% BSA in PBS). Cells were pelleted by spinning at 1700 rpm for 3 min after each wash. After aspirating following the second wash, cells were incubated in 100 μ L of FACS buffer containing the primary antibody diluted to 0.01 mg/mL. Cells were then incubated on ice for 30 min covered from light. Cells were then pelleted by spinning at 1700 rpm for 3 min. The supernatant was then aspirated and the cells were then washed twice with 200 μ L FACS buffer (0.1% BSA in PBS). After aspirating following the second wash, cells were incubated in 100 μ L of FACS buffer containing the secondary antibody diluted to 2 μ g/mL. Cells were then incubated on ice for 30 min covered from light. Cells were then pelleted by spinning at 1700 rpm for 3 min. The supernatant was then aspirated and cells were washed twice with 200 μ L FACS buffer (0.1% BSA in PBS). After aspirating, following the second wash cells were incubated in 100 μ L of FACS buffer containing Sytox red (Thermo Scientific #S34859) diluted by 1:1000 v/v to a final concentration of 4 nM. Cells were then pelleted by spinning at 1700 rpm for 3 min. The supernatant was then aspirated and cells were resuspended in 100 μ L of FACS buffer. For adherent cell lines, cells were resuspended in 100 μ L of FACS buffer containing 1 mM EDTA (Thermo Scientific #15575020). Fluorescence was then monitored using a MACSQuant Analyzer 10 (Miltenyi Biotec) and analysis was done using the FlowJo software package. Gating was performed on single cells and live cells.

Donor PBMC preparation and cell surface staining

A leukocyte filtration device was cut on both ends to allow drainage of blood. 30 mL of PBS and 20 mL of dead volume were pipetted through the distal end of the filters to fully eject leukocyte rich blood. This mixture of PBS and blood was spun at 500xg for 10 min, yielding a buffy coat. The buffy coat was spun

again to further remove any red blood cells. Buffy coats were resuspended in FACS buffer (PBS pH 7.4 and 0.1% w/v bovine serum albumin). Cells were incubated with Human TruStain FcX (BioLegend, 422302) blocking solution; 5µL of blocking solution was added for every 1e6 cells in 100µL. Cells were then stained with 500 ng of rabbit IgG isotype control (Novus Biologicals #NBP2-24893) or anti-TFAM (Proteintech #23996-1-AP) antibodies in 100 µL. Cells were incubated for 30 min on ice. They were washed once with FACS buffer and then stained with goat anti-rabbit IgG Alexa Fluor 647 and incubated for 30 min on ice. They were then washed once with FACS buffer and then stained with lineage antibodies CD3 FITC for T-cells (BioLegend #300452), CD19 PE-Cy7 for B-cells (BioLegend #302216), CD14 PerCP-Cy5.5 for monocytes (BioLegend #325622), and CD66b BV421 for neutrophils (BioLegend #392916) at a concentration of 1:100. Lineage antibody staining was performed on ice for 30 min. Cells were washed twice with FACS buffer and resuspended in DAPI containing buffer just prior to flow cytometry (BD LSRFortessa). Cells were gated first to yield singlets and then subsequently DAPI negative cells were gated to analyze total live cells. Afterwards, CD3 and CD19 single positive cells were gated; the double negative fraction was analyzed to look at CD14 and CD66b positive cells.

Interferon stimulation and reading out IRF9 abundance using intracellular flow cytometry

Jurkat E6-1 cells were cultured overnight at (37°C 5% CO₂) in a T75 flask. Cells were then transferred to 1.5 mL microcentrifuge tubes and pelleted at 300xg for 5 min and resuspended in media at 500,000 cells/mL. 500,000 cells (or 1 mL of the resuspended solution) were then added to a 6 well plate (Corning #3516). IFNα2 (Thermo Scientific #300-02AA-20UG) was then added at the specified concentration from a stock concentration of 1 µg/mL in PBS pH 7.4. For the blank negative control 10 µL of sterile PBS pH 7.4 was added to the media. Cells were then incubated overnight at 37°C with 5% CO₂. Cells were then transferred to 1.5 mL microcentrifuge tubes and pelleted at 300 rcf for 5 min. Cells were then counted and resuspended in FACS buffer to make a 5 million cell/mL solution. 200 µL of the cell solution (or 1 million cells) were then transferred to a 96-well V bottom plate. Cells were then spun at 1500 rpm for 3

min. The supernatant was then aspirated and cells were resuspended in 200 μ L of PBS. Cells were then spun at 1500 rpm for 3 min. The supernatant was then aspirated and cells were resuspended in 100 μ L of Zombie Violet (Biolegend #423113) diluted in PBS by 1:1,000. Resuspended cells were then incubated in the dark for 20 min at 22°C. Cells were then pelleted by spinning at 1500 rpm for 3 min and the supernatant was aspirated. Cells were resuspended in 200 μ L of FACS buffer and spun at 1500 rpm for 3 min. The supernatant was then aspirated and cells were fixed by adding 200 μ L Fixation Buffer (Biolegend #426803) and incubated in the dark for 20 min at 22°C. Cells were then spun at 1500 rpm for 5 min and the supernatant was aspirated. 10X Intracellular Staining Perm Wash Buffer (Biolegend #426803) was diluted to 1X in Deionized water. The fixed cells were then resuspended in 200 μ L of the diluted Intracellular Staining Perm Wash Buffer and centrifuged at 1500 rpm for 5 min. This washing step was repeated two more times. Anti-IRF9 mouse monoclonal primary antibody (Biolegend #660702) was then diluted to a final concentration of 0.01 mg/mL in 1X Intracellular Staining Perm Wash Buffer. After the final wash cells were resuspended in the diluted antibody solution and incubated for 30 min at 22°C. Cells were then washed twice with 200 μ L of 1X Intracellular Staining Perm Wash Buffer and centrifuged at 1500 rpm for 5 min and the supernatant was aspirated between each wash. Goat anti-mouse IgG Alexa Fluor 488 secondary antibody (Thermo Scientific #A32723) was then diluted to a final concentration of 2 μ g/mL in 1X Intracellular Staining Perm Wash Buffer. After the final wash cells were resuspended in the diluted secondary antibody solution and incubated for 30 min at 22°C. Cells were then washed twice with 200 μ L of 1X Intracellular Staining Perm Wash Buffer and centrifuged at 1500 rpm for 5 min between each wash. Cells were then resuspended in 100 μ L FACS buffer for flow cytometry analysis. Fluorescence was then monitored using a MACSQuant Analyzer 10 (Miltenyi Biotec) and analysis was done using the FlowJo software package. Gating was performed on single cells and live cells.

Interferon stimulation and reading out phosphorylated STAT1 abundance using Western blot

550 Jurkat E6-1 cells were cultured overnight at (37°C 5% CO₂) in a T75 flask. Cells were then transferred to
 551 1.5 mL microcentrifuge tubes and pelleted at 400xg for 4 min and resuspended in media at 2 million
 552 cells/mL. IFN α 2 (Thermo Scientific #300-02AA-20UG) was then added at 1 or 10 ng/mL from a stock
 553 concentration of 1 μ g/mL in PBS pH 7.4. For the blank negative control 10 μ L of sterile PBS pH 7.4 was
 554 added to the media. Cells were then incubated for 30 min at 37°C on a Thermomixer (Fisher Scientific
 555 #05-412-503) shaking at 300 rpm. Cells were then transferred to 1.5 mL microcentrifuge tubes and
 556 pelleted at 300xg for 5 min. Cells were then pelleted at 400xg for 4 min and the supernatant was
 557 aspirated. Cells were then resuspended in 1 mL of PBS and the cells were pelleted at 400xg for 4 min.
 558 The supernatant was aspirated and cells were resuspended in 100 μ L RIPA buffer containing 1X protease
 559 inhibitor and 1X phosphatase inhibitor. Cells were lysed at 4°C for 20 min. The samples were then spun
 560 at 16,000xg for 15 min at 4°C and the supernatant was moved to a new LoBind tube. The protein
 561 concentration was measured by BCA (Pierce BCA Assay, Thermo Scientific #23225). 10 μ g of protein in
 562 lysate was denatured with a 4X loading buffer which was supplemented with 10% v/v 1 M dithiothreitol
 563 (Thermo Scientific #15-508-013) and boiled at 95 °C for 10 min. 20 μ L of the reduced protein sample was
 564 then loaded on a 12-well SDS–PAGE gel (10% Bis-Tris gel, Bio-Rad #345-0123) and run at 200 V for 60
 565 min in 1X XT MOPS buffer (Bio-Rad #1610788). Protein was then transferred from the gel to a 0.45 μ m
 566 nitrocellulose membrane (Bio-Rad #1620115). After transfer, total protein was measured by staining
 567 with total protein stain (Revert 700 Total Protein Stain, LI-COR Biosciences #926-11021) and incubated
 568 for 3 min at 22°C. The membrane was then imaged using an OdysseyCLxImager (LI-COR Biosciences).
 569 The total protein stain was then removed using a basic buffer and washed three times with PBS pH 7.4.
 570 The membrane was then blocked with PBS Odyssey Blocking Buffer (LI-COR Biosciences #927-40000) for
 571 1 h at room temperature with gentle shaking. The membrane was stained with anti-pSTAT1 (BioLegend
 572 #666402) diluted by 1:500 to a final concentration of 1 μ g/mL in PBS Odyssey Blocking Buffer overnight
 573 at 4°C with gentle shaking, then washed four times with PBS with 0.1% Tween-20 (Bio-Rad #1706531),

PBS-T, for 5 min each. The membrane was then incubated with 800CW goat anti-mouse IgG (LI-COR Biosciences #926-32210) diluted by 1:1,000 to a concentration of 1 µg/mL in PBS Odyssey Blocking Buffer for 1 h at room temperature with gentle shaking. The membrane was washed four times with PBS-T, and then imaged using an OdysseyCLxImager (LI-COR Biosciences).

Surface antigen staining for confocal microscopy

Jurkat E6-1 cells were cultured overnight (37°C, 5% CO₂) in a T75 flask. Cells were moved to 15 mL conical tube and were counted using Trypan Blue (VWR #VWRVK940) and a Countess Cell Counter (Thermo Scientific). Cells were then diluted to a concentration of 5 million cells/mL in FACS buffer (PBS pH 7.4 and 0.1% w/v bovine serum albumin). 200 µL of the cell solution, or 1 million cells, were then transferred to a 96-well V bottom plate. Cells were kept on ice and spun at 4°C for the duration of the protocol. Cells were then spun at 1700 rpm for 3 min. The supernatant was then aspirated and cells were washed two times with 200 µL FACS buffer (0.1% BSA in PBS). Cells were pelleted by spinning at 1700 rpm for 3 min after each wash. After aspirating the supernatant following the second wash, cells were incubated in 100 µL of FACS buffer containing the primary antibody diluted to 0.01 mg/mL. Cells were then incubated on ice for 30 min covered from light. Cells were then pelleted by spinning at 1700 rpm for 3 min. The supernatant was then aspirated and cells were washed two times with 200 µL of FACS buffer. After aspirating following the second wash, cells were incubated in 100 µL of FACS buffer containing the secondary antibody goat anti-mouse IgG Alexa Fluor 488 (Thermo Scientific #A32723) diluted to 0.01 mg/mL. Cells incubated on ice for 30 min covered from light. Cells were then pelleted by spinning at 1700 rpm for 3 min. The supernatant was then aspirated and cells were washed two times with 200 µL of FACS buffer. After aspirating following the second wash, cells were incubated in 100 µL of FACS buffer containing diluted CellMask Orange Plasma Membrane Stain (Thermo Scientific #C10045) diluted by 1:10000 to a final concentration of 0.5 µg/mL in FACS and incubated on ice in the dark for 5 min. Cells were then pelleted by spinning at 1700 rpm for 3 min. The supernatant was aspirated and the

cells were washed two times with 200 μ L FACS buffer. For fixation, cells were resuspended in a solution of 4% v/v paraformaldehyde in PBS and incubated in the dark for 30 min at room temperature. Cells were then pelleted by spinning at 1700 rpm for 3 min. Cells were then washed one time with 200 μ L of PBS and pelleted again and the supernatant was aspirated. Cells were then resuspended in a solution of DAPI (Fisher Scientific #EN62248) diluted by 1:1000 to a final concentration of 5 μ g/mL in PBS and incubated in the dark for 5 min at room temperature. Cells were then pelleted and resuspended in 25 μ L of PBS. 25 μ L of the cell solution was then pipetted onto a glass slide and 25 μ L of mounting medium (Fisher Scientific #NC1601055) was then mixed with the cells. A coverslip was then placed over the solution and the cells were enclosed using clear nail polish (Electron Microscopy Sciences #72180). The slides were then left to dry for 16 h in the dark at room temperature.

Confocal microscopy

A Nikon A1R confocal microscope equipped with a Pan Fluor 60x oil immersion 1.30-numerical aperture objective (Nikon) was used. This instrument is equipped with a 405-nm violet laser, a 488-nm blue laser, a 561-nm green laser and a 639-nm red laser.

Wheat germ agglutinin-horseradish peroxidase (WGA-HRP) catalyzed cell surface biotinylation

Jurkat E6-1 were used at 10 million cells/labeling experiment and, unless otherwise noted, were pelleted at 1,000xg for 4 min at 4 °C. Cells were first removed from culture, washed twice with cold DPBS (Fisher Scientific #MT21031CV) and resuspended in cold DPBS at 10 million cells/mL and transferred in 1 mL aliquots to Protein LoBind tubes (Fisher Scientific #13-698-794). The cells were then pelleted to remove the supernatant and resuspended in 1 mL of cold DPBS containing 20 μ g of WGA-HRP (Vector Laboratories #PL-1026-2) or 20 μ g of unconjugated HRP (Thermo Fisher Scientific #31490). The cells were then incubated at 4°C for 1 h on a rotisserie, followed by centrifugation to remove the supernatant and then washed twice in cold DPBS. Each sample was resuspended in 1 mL of cold DPBS and added directly to a 15 mL conical tube containing 4 mL of reaction buffer, 250 μ M biotinyl tyramide (Sigma-

Aldrich #SML2135-250MG) and 1 mM H₂O₂ (Sigma-Aldrich #H1009-100ML) in cold DPBS, and incubated for 1 min at 4 °C. After 1 min, 5 mL quenching buffer, DPBS containing 5 mM Trolox (Sigma-Aldrich #238813-5G), 10 mM Sodium L-ascorbate (Sigma-Aldrich #11140-250G), 10 mM NaN₃ (Sigma-Aldrich #S2002-100G), was added to the cells. The cells were then pelleted by centrifugation (4 min at 1,000xg, 4°C) and washed once with 10 mL of quenching buffer.

After the final wash, the cells were pelleted and each sample was resuspended in 1 mL of RIPA buffer (Pierce RIPA Buffer, Thermo Scientific #89901) containing 1X protease inhibitors (cOmplete Protease Inhibitor Cocktail EDTA-Free, Roche #04693159001) and incubated for 20 min at 4 °C. The samples were then spun at 16,000xg for 15 min at 4 °C. The supernatant was then moved to a new tube and the protein concentration was measured by BCA (Pierce BCA Assay, Thermo Scientific #23225) and samples were stored at -80°C until streptavidin bead enrichment.

Horseradish peroxidase catalyzed cell surface biotinylation

Jurkat E6-1 or Daudi cells were used at 50 million cells/labeling experiment and, unless otherwise noted, were pelleted at 1,000xg for 4 min at 4 °C. Cells were first removed from culture, washed twice with cold DPBS (Fisher Scientific #MT21031CV) and resuspended in cold DPBS at 50 million cells/mL and transferred in 1 mL aliquots to Protein LoBind tubes (Fisher Scientific #13-698-794). The cells were then pelleted to remove the supernatant and resuspended in 1 mL of cold DPBS containing 25 µg of isotype control (Thermo Scientific #31903) or 25 µg of the respective targeting antibody. The cells were then incubated at 4°C for 30 min on a rotisserie, followed by centrifugation to remove the supernatant and then washed twice in cold DPBS. The cells were then resuspended in 1 mL of cold DPBS containing 25 µg goat α-Mouse IgG HRP (EMD Millipore #12-349) and incubated for 30 min at 4°C on a rotisserie. The cells were then pelleted to remove the supernatant and washed twice in cold DPBS. Each sample was resuspended in 1 mL of cold DPBS and added directly to a 15 mL conical tube containing 4 mL of reaction buffer, 250 µM biotinyl tyramide (Sigma-Aldrich #SML2135-250MG) and 1 mM H₂O₂ (Sigma-Aldrich

#H1009-100ML) in cold DPBS, and incubated for 1 min at 4 °C. After 1 min, 5 mL quenching buffer, DPBS containing 5 mM Trolox (Sigma-Aldrich #238813-5G), 10 mM Sodium L-ascorbate (Sigma-Aldrich #11140-250G), 10 mM NaN₃ (Sigma-Aldrich #S2002-100G), was added to the cells. The cells were then pelleted by centrifugation (4 min at 1,000xg, 4°C) and washed once with 10 mL of quenching buffer.

After the final wash, the cells were pelleted and each sample was resuspended in 1 mL of membrane permeabilization buffer (MEM-PER Plus Membrane Fractionation Kit, Thermo Scientific #89842) containing 1X protease inhibitors (cOmplete Protease Inhibitor Cocktail EDTA-Free, Roche #04693159001) and incubated for 20 min at 4 °C. The samples were then spun at 16,000xg for 15 min at 4 °C. The supernatant enriched in cytosolic proteins was removed, and the pellet was resuspended in 300 µL lysis buffer (Pierce RIPA Buffer, Thermo Scientific #89901) containing 1% SDS v/v and 1X protease inhibitors. The samples were sonicated in the lysis buffer to break up the membrane pellet once for 5 s at power level 6 using a probe sonicator and then heated for 5 min at 95 °C. The samples were then diluted to 1 mL with RIPA and sonicated twice for 5 s at power level 3 using a Misonix Sonicator 3000 probe sonicator. The protein concentration was measured by BCA (Pierce BCA Assay, Thermo Scientific #23225) and samples were stored at -80°C until streptavidin bead enrichment.

Iridium catalyzed cell surface biotinylation (µMap)

Jurkat E6-1 cells were used at 50 million cells/labeling experiment and, unless otherwise noted, were pelleted at 1,000xg for 4 min at 4 °C. Cells were first removed from culture, washed twice with cold DPBS (Fisher Scientific #MT21031CV) and resuspended in cold DPBS at 50 million cells/mL and transferred in 1 mL aliquots to Protein LoBind tubes (Fisher Scientific #13-698-794). The cells were then pelleted to remove the supernatant and resuspended in 1 mL of cold DPBS containing 25 µg of isotype control (Thermo Scientific #31903) or 25 µg of the respective targeting antibody. The cells were then incubated at 4°C for 30 min on a rotisserie, followed by centrifugation to remove the supernatant and then washed twice in cold DPBS. The cells were then resuspended in 1 mL of cold DPBS containing 25 µg

of iridium-conjugated goat α -mouse secondary IgG (Supplemental Methods). The samples were incubated on a rotisserie for 30 min at 4 °C and then pelleted to remove the supernatant. The cells were washed twice with 1 mL of cold DPBS and then resuspended in 1 mL of cold DPBS containing 250 μ M diazirine biotin (Supplemental Methods). The samples were placed on the Lumidox II photoreactor (Analytical #LUM2CON) and irradiated on a 24-well LED array (Analytical #LUM22418LA445) with 445 nm light at 195W for 10 min at 4 °. After irradiation, the cells were then pelleted to remove the supernatant and washed twice with 1 mL of cold DPBS.

After the final wash, the cells were pelleted and each sample was resuspended in 1 mL of membrane permeabilization buffer (MEM-PER Plus Membrane Fractionation Kit, Thermo Fisher Scientific #89842) containing 1X protease inhibitors (cOmplete Protease Inhibitor Cocktail EDTA-Free, Roche #04693159001) and incubated for 20 min at 4 °C. The samples were then spun at 16,000xg for 15 min at 4 °C. The supernatant enriched in cytosolic proteins was removed and the pellet was resuspended in 300 μ L of lysis buffer (Pierce RIPA buffer, Thermo Scientific #89901) containing 1% v/v SDS and 1X protease inhibitors. The samples were sonicated in the lysis buffer to break up the membrane pellet once for 5s at power level 6 using a Misonix Sonicator 3000 probe sonicator and then heated for 5 min at 95°C. The samples were then diluted to 1 mL with RIPA and sonicated twice for 5 s at power level 3 using a Misonix Sonicator 3000 probe sonicator. The protein concentration was measured by BCA and samples stored at -80°C until streptavidin bead enrichment.

Magnetic streptavidin bead enrichment

For bead enrichment membrane fractions were added to a Protein LoBind tube containing 125 μ L of streptavidin magnetic beads (Thermo Fisher Scientific #88817) that were pre-washed twice with 500 μ L of RIPA buffer. The samples were then incubated for 3 h at room temperature on a rotisserie and the beads were then pelleted on a magnetic rack (Thermo Scientific #12321D). The supernatant was removed, and the beads were washed three times with 500 μ L 1% v/v SDS (Thermo Scientific

#15553027) in DPBS, three times with 500 μ L 1M NaCl (Fisher Scientific #S6713) in DPBS, three times with 500 μ L 10% v/v EtOH (Gold Shield Distributors #412804) in DPBS, and once with 500 μ L RIPA buffer (Thermo Scientific #89901). The beads were incubated in each of the washes for 3 min before pelleting to remove the wash. After the final wash, the beads were resuspended in 25 μ L of 5% v/v SDS containing 20 mM DTT and 25 mM biotin (Sigma-Aldrich #B4501-1G) and heated to 95°C for 10 min. The samples were placed on the magnetic rack and the supernatant was collected and transferred to a new Protein LoBind tube and stored at -80 °C until quantitative proteomic sample preparation and analysis

Tandem mass spectrometry sample preparation

75 μ g of biotin-enriched membrane samples were reduced in a solution of 5% v/v SDS (Thermo Fisher Scientific #15553027) with 10 mM dithiothreitol (Bio-Rad #1610611) and incubated at 95°C for 10 min. Samples were then allowed to cool to room temperature and then alkylated using 40 mM 2-chloroacetamide (Sigma-Aldrich #C0267-500G) and incubated for 1 h at room temperature. The alkylation reaction was quenched by adding 85% (m/v) phosphoric acid (Millipore Sigma #W290017-1KG-K) to a final solution percentage of 1.2% phosphoric acid in the alkylated protein solution. 700 μ L (7X volume) of S-trap bind wash buffer, 100 mM triethylammonium bicarbonate buffer (TEAB) (Fisher Scientific #60044974), 90% (v/v) MeOH (Fisher Scientific #A452-4), was then added to the alkylated sample. Samples were then loaded onto a micro-S-trap-column (Protifi #C02-micro-80) in 175 μ L increments and spun at 4000xg for 20 s for each loading cycle. Sample caught on the column was then washed three times using the S-trap bind wash buffer with spins of 4000xg for 30 s to pass the buffer over the column. Mass spectrometry-grade trypsin (Promega #V5113) was then diluted in 50 mM TEAB buffer to a concentration of 0.15 mg/mL. 25 μ L of the diluted trypsin solution, or 3.75 μ g of trypsin, was loaded onto each column to make a final trypsin:protein ratio of 1:20. Samples were then incubated at 47°C for 90 min. The column was then washed using 50 mM TEAB followed by 0.2% (v/v) formic acid (Thermo Fisher #28905) in water with a spin at 1000xg for 60 s between each wash. Digested samples

were then eluted by adding 40 μL of a solution of 50% v/v acetonitrile (Fisher Scientific #A955-212), 0.2% v/v formic acid (Thermo Fisher #28905), and 50% v/v water then spun at 4000xg for 60 s. Samples were then dried using vacuum centrifugation (CentriVap Complete Vacuum Concentrator, Labconco #7315023) for 16 h. The samples were resuspended in 0.2% (v/v) formic acid diluted in water immediately prior to mass spectrometry data acquisition.

Tandem mass spectrometry data acquisition

All samples were resuspended in a 0.2% (v/v) formic acid diluted in water at 1 $\mu\text{g}/\mu\text{L}$, and 1 μg of each sample was added to a pooled sample tube for the six gas phase fractionation injections. 1 μL (1 μg) of the total peptide solution was injected in the column for each sample. Peptides were separated over a 25-cm EasySpray reversed phase LC column (75 μm inner diameter packed with 2 μm , 100 \AA , PepMap C18 particles, Thermo Fisher Scientific #ES802). The mobile phases (A: water with 0.2% (v/v) formic acid and B: acetonitrile with 0.2% (v/v) formic acid) were driven and controlled by a Dionex Ultimate 3000 RPLC nano system (Thermo Fisher Scientific). Gradient elution was performed at 300 nL/min. Mobile phase B was increased from 1 to 5% (v/v) over 6 min, followed by a gradual increase to 25% by 60 min, and ending with a ramp to 90% B at 71 min, and a wash at 90% B for 5 min. Flow was then ramped back to 1% B over the course of 1 min, and the column was re-equilibrated at 1% B for 15 min, for a total analysis of 90 min. Eluted peptides were analyzed on an Orbitrap Fusion Tribrid MS system (Thermo Fisher Scientific). Precursors were ionized using an EASY-Spray ionization source (Thermo Fisher Scientific) source held at +2.2 kV compared to ground, and the column was held at 40 $^{\circ}\text{C}$. The inlet capillary temperature was held at 275 $^{\circ}\text{C}$. For gas phase fractionation injections, a pooled sample was injected six times and acquired using staggered-window injections with m/z windows of 395 to 505 m/z, 495 to 605 m/z, 595 to 705 m/z, 695 to 805 m/z, 795 to 905 m/z, and 895 to 1005 m/z. For all staggered-window injections a 4 m/z precursor isolation window was used. MS/MS scans were collected using high-energy collisional dissociation at 33 normalized collision energy and mass analysis was performed

in the Orbitrap using a resolution of 30,000 while scanning from 120–2000 m/z. Individual samples were collected using single-injection data independent acquisitions (DIA) with survey scans of peptide precursors collected in the Orbitrap from 385–1015 m/z with an automatic gain control target of 400,000, a maximum injection time of 55 ms and a resolution of 60,000. An isolation window of 16 m/z was used to select precursor ions with the quadrupole. MS/MS scans were collected using high-energy collisional dissociation at 33 normalized collision energy and mass analysis was performed in the Orbitrap using a resolution of 30,000 while scanning from 120–2000 m/z.

Mass spectrometry database search and quantitation

Proteins were identified with DIA-NN⁵⁷ and MSFragger 4.0⁵⁸ software on the Fragpipe GUI version 21.0⁵⁹, searching against the UniProt human reviewed proteome in addition to common contaminants. Methionine oxidation (+15.995 Da) and N-terminal acetylation (+42.011 Da) were included as variable modifications, while carbamidomethylation of cysteine (+57.021 Da) was included as a fixed modification. Precursor ion search tolerance was set to 20 ppm and fragment ion mass tolerance was set to 20 ppm. Peptide and protein identifications were thresholded at a 1% false discovery rate using a target-decoy method. Proteins were quantified and normalized by label-free quantitation using the DIA-NN software. For quantitative comparisons, protein intensity values were log transformed, and missing values for proteins were imputed from a normal distribution with a width of 0.3 and a down shift value of 1.8 in Perseus (version 2.0)⁶⁰. A two-sided t-test was additionally done using Perseus with false discovery rate (FDR) correction performed using permutation-based FDR with 250 randomizations.

Pairwise protein subunit distance measurements

T-cell receptor subunit distances were measured in a pairwise manner across all TCR subunits from the center mass of each subunit from the cryo-EM structure of the TCR-CD3 complex (PDB 6JXR). For subunits with a stoichiometry greater than 1 (eg. 2 copies) the mean distance between protein pairs was taken.

Unsupervised clustering and functional domain analysis

A matrix of fold change values was generated including data from our 65 proximity labeling experiments in Jurkat cells and 20 proximity labeling experiments in Daudi cells. Fold change matrices were hierarchically clustered and the resulting dendrogram was cut at intervals to obtain a nested hierarchy of associations. Cuts closer to the root of the dendrogram result in larger groups of proteins and cuts closer to the tips define smaller groups. In both Jurkat and Daudi cell lines known protein assemblies cluster together at the smallest group. In Jurkat this is represented by the CD3-TCR complex and in Daudi cells the MHC class II HLA-DR complex.

Feature assembly for identifying surface protein associations

A matrix of fold change values containing the enrichment values for 65 proximity labeling experiments on Jurkat cells was created and filtered to contain only proteins quantified in at least 20 proximity labeling experiments. Fold changes were then normalized to lie between 0 and 1 by min-max normalization. The difference between fold change values for a given protein was then estimated on an all-by-all basis for each of 60,031 surface protein pairs. A matrix of similarity measures was then made from our fold change matrix by doing all-by-all pairwise estimations of Pearson correlation, Spearman correlation, Euclidean distance, Bray-Curtis dissimilarity, Manhattan distance, and Chebyshev distance for each surface protein pair. Euclidean distance, Bray Curtis score, Manhattan distance, and Chebyshev distance were inverted and normalized to fall between 0 and 1 using min-max normalization. The matrix of fold change differences and similarity measures were then concatenated to give a final feature matrix containing 4,202,170 data points.

Gold standard protein interaction training and test set assembly

Due to the limited size of datasets containing known human surface protein interactions we utilized several sources to construct our gold standard interaction set: CORUM mammalian complex dataset⁶¹, HuMap 2.0⁶², STRING.db⁶³ with interactions filtered to score above 800, and BioGRID⁶⁴. For protein

complex datasets, all complex members were converted to binary interactions where each subunit is considered an interactor to all other members of the complex. The compilation of these datasets led to a final binary gold standard interaction list of 1,185,470 protein-protein interactions. Of these interactors, 1,009 were found to contain two subunits contained in our Jurkat cell surface proteome. From this list of 1,009 gold-standard cell surface interactors we used 706 for our positive training interactions and 303 as our withheld positive test interactions. For negative training and test sets, 10,000 pairwise interactions not found in the gold standard set were randomly selected and split into training and test sets to make 7,000 negative training interactions and 3,000 negative test interactions.

Random forest classifier design and training

For the binary prediction of protein associations we used the scikit-learn⁶⁵ Random Forest Classifier. We then searched for optimal hyperparameters for the classifier with 5-fold cross-validation. This led to a classifier with an area under the receiver operator characteristic curve of 0.93 and an area under the precision-recall curve of 0.62 when applied to a withheld set of 303 gold-standard protein interactions. This model was applied to the entire feature matrix to give an association score for each surface protein pair with higher scores corresponding to higher likelihood of the proteins interacting or colocalizing. True positive rate, false positive rate, precision, recall, and false discovery rates were calculated from 303 positive interactions and 3,000 negative interactions.

3D network plot design and graph connectivity score estimation

To construct the 3D heatmap first a two dimensional Kawada-Kawaii spring network showing the enrichment connections of bait proteins for 49 high quality proximity labeling experiments was created. High quality refers to the enrichment of the targeted bait protein as one of the top hits in . The 2-dimensional bait network was then transformed to lay on a 3D sphere with radius = 1. This was done by finding a given node's position and estimating theta and phi to create a new 3D spherical coordinate using the following equation:

$$(\theta, \phi) = (\arctan2(\sqrt{x^2 + y^2}, z), \arctan(y, x))$$

Pole points at 0, 0, 1 and 0, 0, -1 were added to the transformed points. For a given protein of interest, the normalized fold change for a given bait experiment was displayed on the corresponding bait node. A solid shell was then added to the 3D network and the scipy griddata function was then used to interpolate fold change values for the space between bait nodes. The graph connectivity score was then estimated by integrating over the fold change values on the sphere using the following equation:

$$S = \frac{\sum_i v_i \sin(\theta_i)}{\sum_i \sin(\theta_i)}$$

Supplemental information

Document S1. Figures S1-S15

Table S1 Flow cytometry surface staining results for Jurkat and Daudi cells, related to Figures 1 and S2.

Table S2 Fold change enrichment values for Jurkat cell surface proximity labeling experiments, related to Figures 1, 3, 6, and S3.

Table S3 Hierarchical clustering of Jurkat cell surface proteins with functional groups assigned to clusters at different cuts, related to Figures 1 and S3

Table S4 Fold change enrichment values for Daudi cell surface proximity labeling experiments, related to Figures 2, 6, and S4.

Table S5 Hierarchical clustering of Daudi cell surface proteins with functional groups assigned to clusters at different cuts, related to Figures 2 and S4.

Table S6 Random forest classifier predictions of associations for Jurkat cell surface proteins, related to Figures 4, 5, S6, and S7.

Table S7 Time course DIA proteomics results for Jurkat WT and IL10RB KO lines stimulated with a Type I Interferon, related to Figures 5 and S7.

Table S8 Estimated graph connectivity scores for Jurkat and Daudi cells, related to Figures 6 and S8.

A

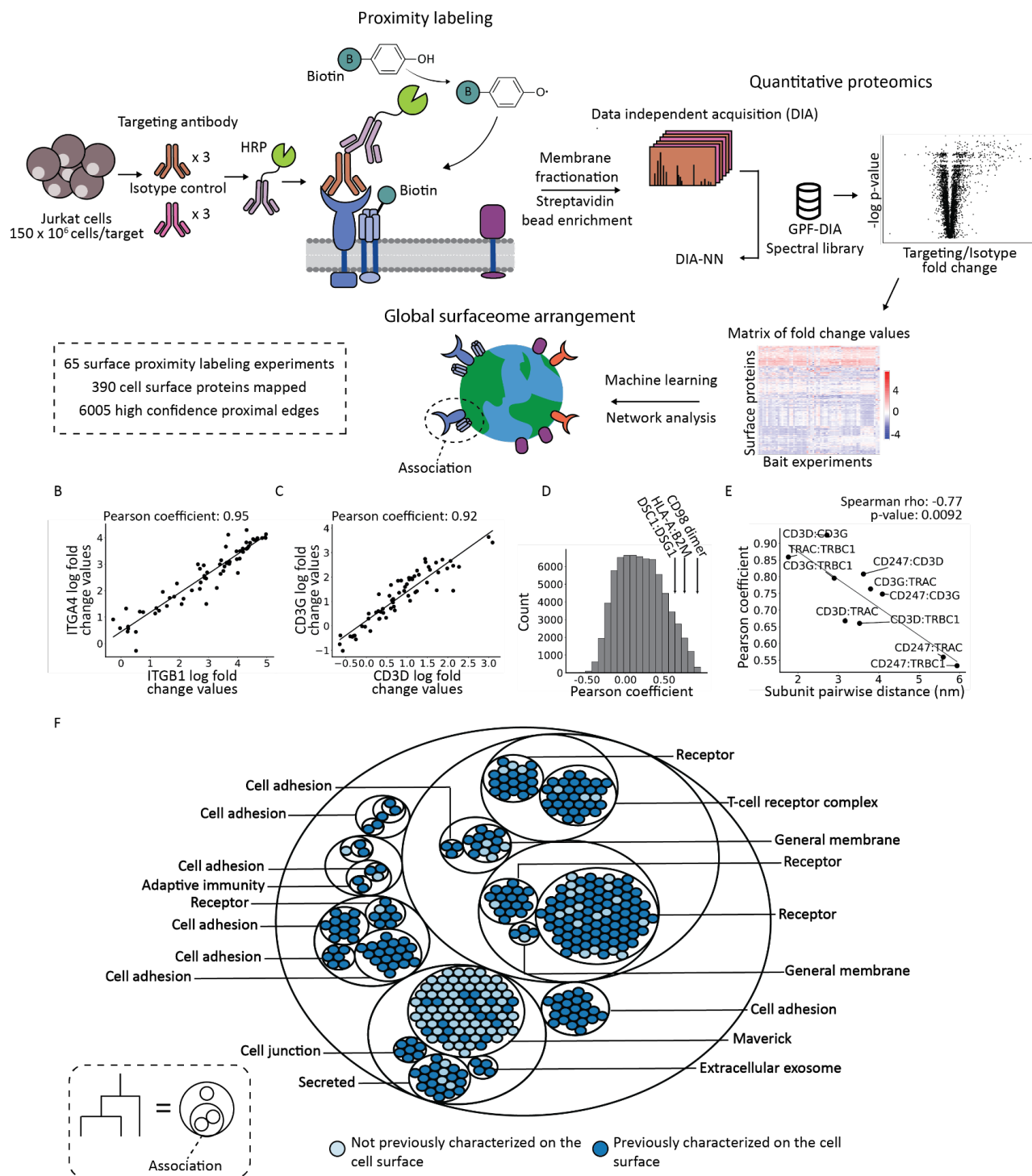


Figure 1. Systematic proximity labeling of the Jurkat cell surface proteome reveals functional clusters on the cell surface. (A) Overview of the approach to spatially map the Jurkat cell surface proteome. DIA-NN: DIA by neural networks; GPF-DIA: Gas phase fractionation data independent acquisition. (B) Scatterplot

841 of log fold change enrichment values and correlation coefficients for the known protein interactors (B)
842 ITGB1/ITGA4 and (C) CD3 complex members CD3D and CD3G. (D) Distribution of pairwise Pearson
843 correlation coefficient across all Jurkat surface protein pairs. Additional known protein interactions are
844 highlighted. (E) Scatterplot of T-cell receptor subunit pairwise distances compared to their estimated
845 Pearson correlation values from Jurkat surface proximity labeling experiments. (F) Circle plot of the
846 clustered Jurkat cell surface proteome labeled with enriched functional annotations for each cluster and
847 colors indicating cell surface characterization of individual surface proteins.

848

849

850

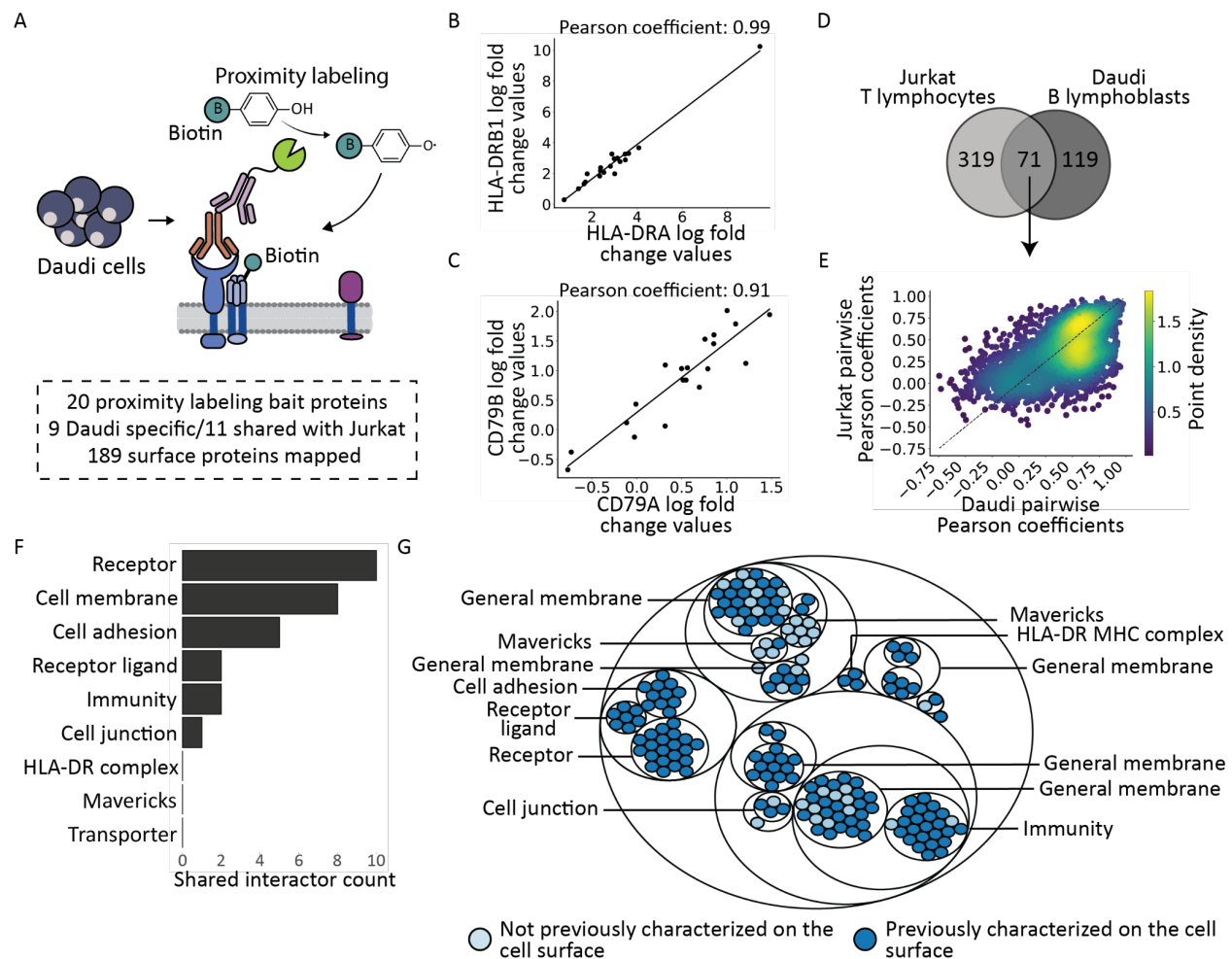


Figure 2. Cell surface protein interactions are conserved across different cell types. (A) Numerical summary of cell surface proximity labeling done on Daudi B-lymphoblast cells. (B) Scatterplot of log fold change enrichment values and correlation coefficients for the MHC class 2 subunits HLA-DRA and HLA-DRB1 (C) and B-cell receptor SRK subunits CD79A and CD79B. (D) Venn diagram showing overlap of cell surface proteins identified on Jurkat and Daudi cells. (E) Scatterplot of pairwise Pearson correlation coefficients in Daudi and Jurkat cell lines. The dashed line represents a line of perfect correspondence where $y = x$. (F) Functional annotation counts of conserved protein-protein interactions in Daudi and Jurkat cell lines. Conserved protein-protein interactions are defined as protein pairs with Pearson coefficients greater than 0.7 and difference between the two cell lines of less than 0.1. (G) Circle plot of

861 the clustered Daudi cell surface proteome labeled with enriched functional annotations for each cluster
862 and colors indicating cell surface characterization of individual surface proteins.

863

864

865

866

867

868

869

870

871

872

873

874

875

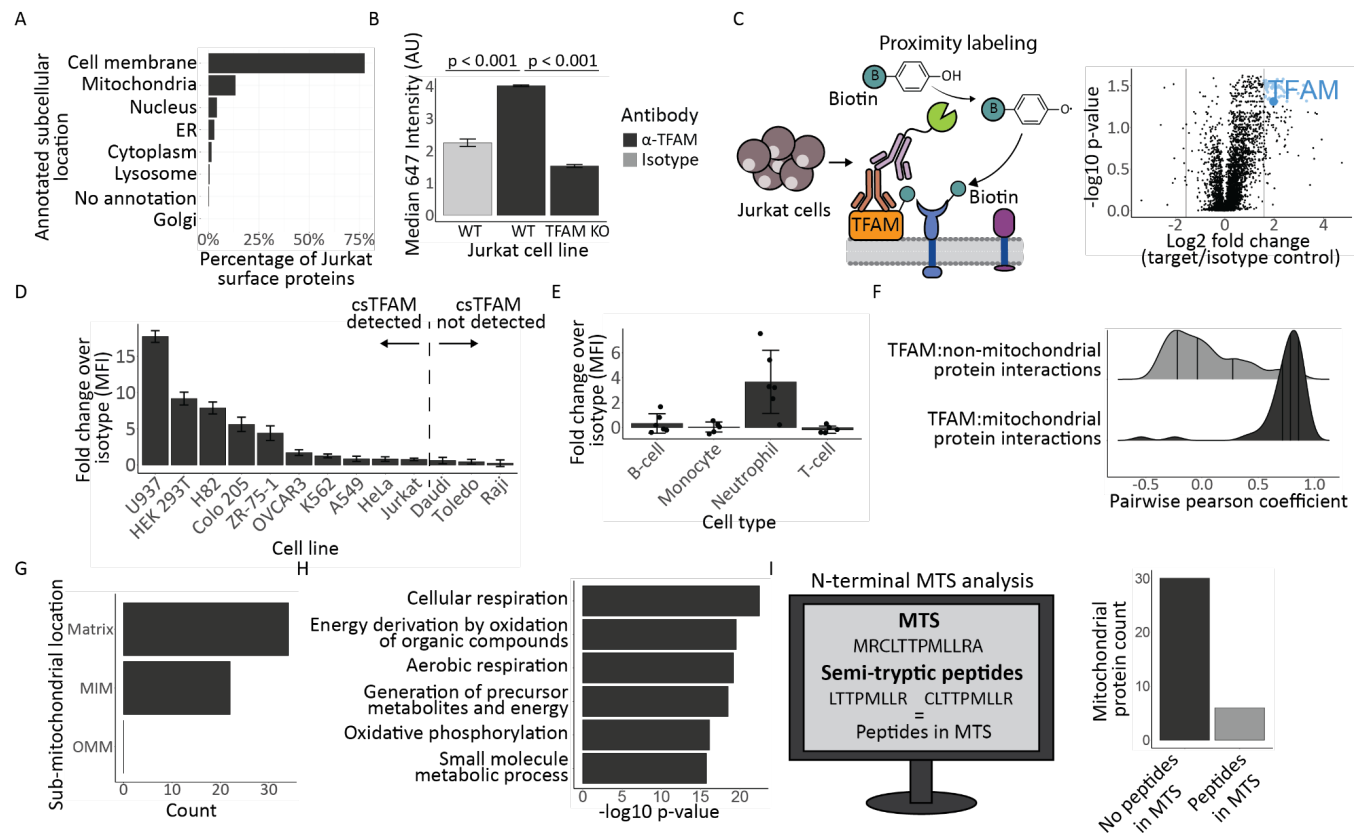


Figure 3. Detection and validation of maverick mitochondrial proteins on the cell surface of Jurkat cells. (A) Breakdown of annotated subcellular localizations for the Jurkat cell surface proteome. (B) Flow cytometry cell surface staining with an isotype control (grey) and anti-TFAM (black) antibody on Jurkat WT and TFAM KO cell lines. P-values were estimated using a Student's T-test. (C) Cartoon of the cell surface proximity labeling experiment using TFAM as a bait. Volcano plot showing the results of cell surface proximity labeling using an anti-TFAM antibody. TFAM is highlighted in dark blue and other enriched proteins are shown in light blue. (D) Anti-TFAM flow cytometry cell surface staining results across 13 different cell lines. Fold change was estimated by comparing to isotype control binding levels on each cell line. (E) Anti-TFAM cell surface staining results for different cell types from 6 different donors. Fold change was estimated by comparing to isotype control binding levels. (F) Distribution of TFAM pairwise Pearson values with other detected surface mitochondrial proteins (grey) and all non-mitochondrial surface proteins (black). (H) Sub-mitochondrial counts for mitochondrial proteins

detected on the cell surface. (H) Enriched GO biological process terms for cell surface mitochondrial proteins. (I) Description of the analysis done to identify native N-termini peptides from the mitochondrial transit sequence for mitochondrial proteins detected on the cell surface and bar plot showing the count of maverick mitochondrial proteins with peptides detected in the TPPRED3 predicted mitochondrial transit sequence (MTS).

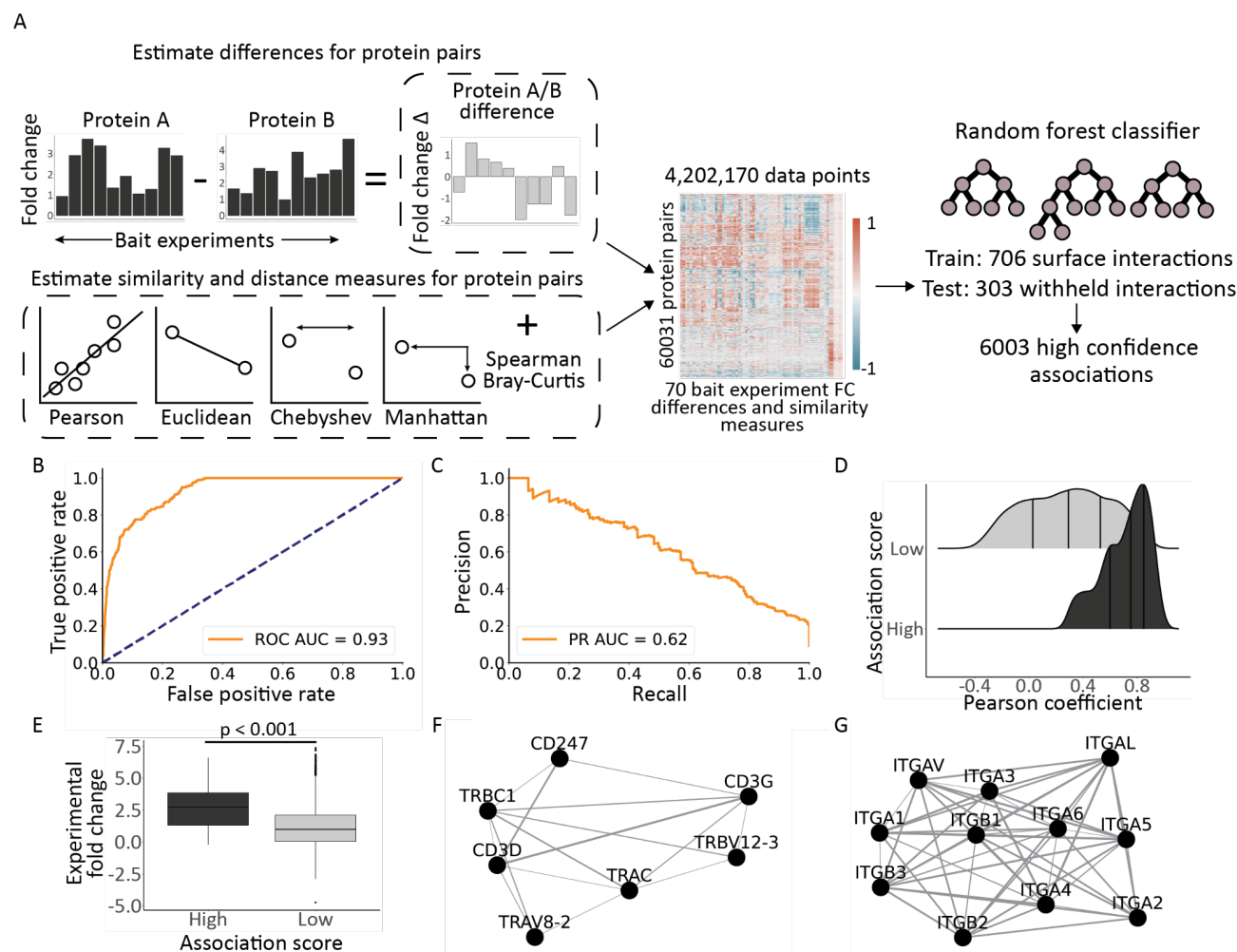


Figure 4. Design of a machine learning classifier for predicting cell surface protein associations. (A) Workflow for training a random forest classifier to identify cell surface associations. (B) Receiver operator characteristic (ROC) curve for 303 withheld gold standard protein pairs. Area under the ROC curve (ROC AUC) is 0.93. (C) Precision-recall curve for 303 withheld protein pairs. Area under the precision recall curve (PR AUC) is 0.62. (D) Distribution of Pearson correlation coefficients for protein pairs with high random forest classifier association predictions and low random forest classifier predictions. High association scores are defined as any protein pair with a score higher than 0.6. (E) Comparison of experimentally measured proximity labeling fold change (FC) enrichments for 49 bait experiments to the classifier predicted association scores of bait-prey interactions. Welch's T-test statistic 16.69, and p-value 8.045e-49. (F) Kawada-Kawai network of classifier predictions for T-cell

915 receptor subunits. Edge weight corresponds to association scores. (G) Kawada-Kawai network of
 916 classifier predictions for integrin alpha and beta proteins detected on the Jurkat cell surface. Edge
 917 weight corresponds to association scores.

918

919

920

921

922

923

924

925

926

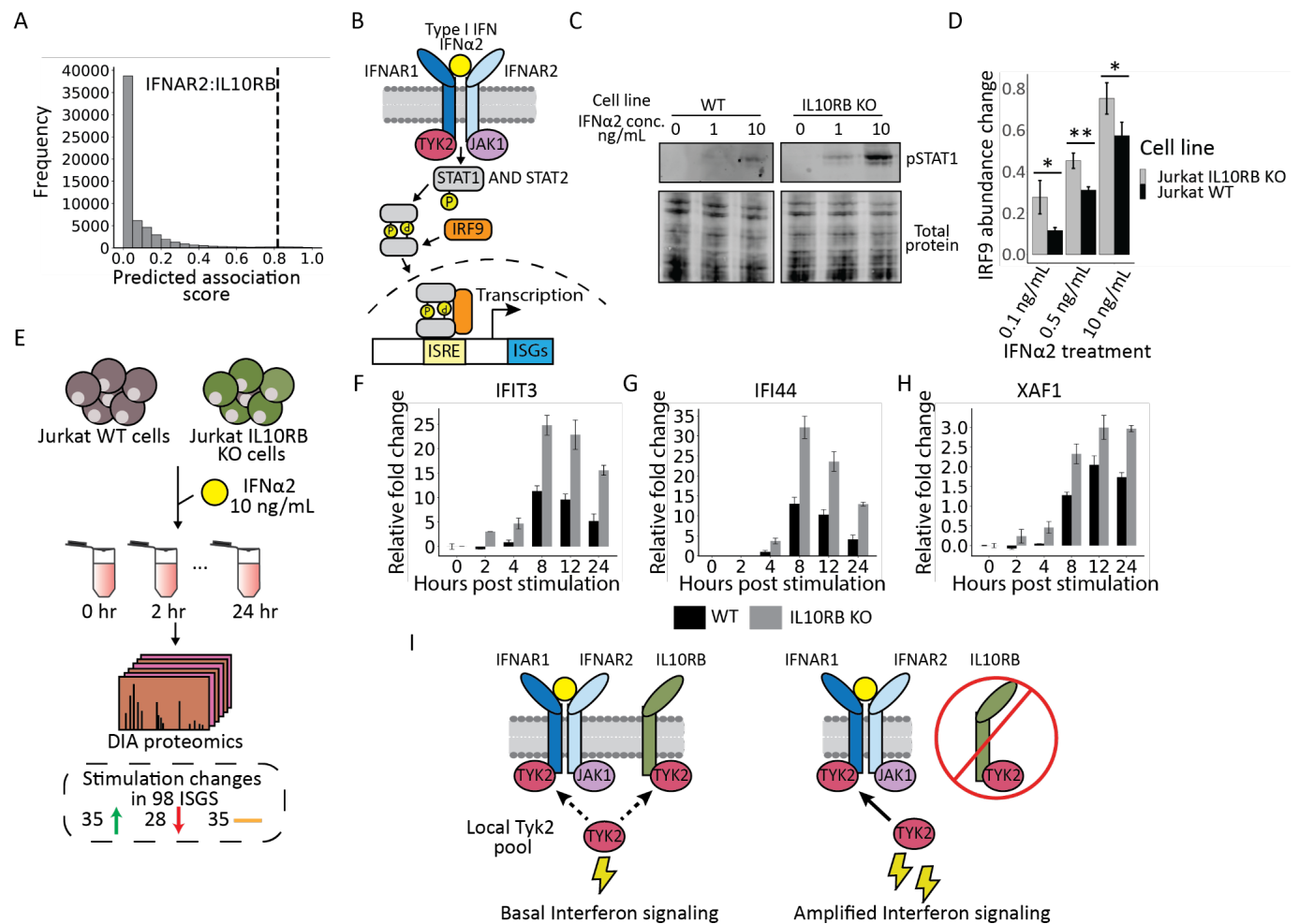


Figure 5. Prediction and characterization of IL10RB as a Type 1 Interferon regulator. (A) Distribution of classifier predicted association scores for all protein pairs with the prediction for IFNAR2 and IL10RB highlighted. (B) Cartoon of the Type 1 Interferon signaling transduction pathway. (C) Western blot of phosphorylated STAT1 (pSTAT1) abundance in IFNα2 stimulated Jurkat WT and IL10RB knockout lines at different interferon concentrations. (D) Intracellular flow cytometry measurement of cytosolic IRF9 abundance change in IFNα2 stimulated Jurkat wild type (black) and IL10RB knockout lines (grey) at various IFNα2 concentrations. Student's t-test p-values for at the different concentrations are 0.027, 0.0036, and 0.035 for 0.1, 0.5, and 10 ng/mL respectively. (E) Cartoon workflow for monitoring interferon stimulated gene protein stimulation using quantitative proteomics when Jurkat WT and IL10RB KO cells were treated with 10 ng/mL of IFNα2 for 0, 2, 4, 8, 12, and 24 h. (F-H) Abundance change

plots of selected ISGs (F) IFIT3, (G) IFI44, and (H) XAF1 in the Jurkat WT (black) and IL10RB KO (grey) cell lines. Relative fold change was estimated in comparison to the 0 h time point. (I) Proposed model for IL10RBs regulatory mechanism for Type I Interferon signaling by competing with IFNAR1 for binding of the Jak kinase family member Tyk2.

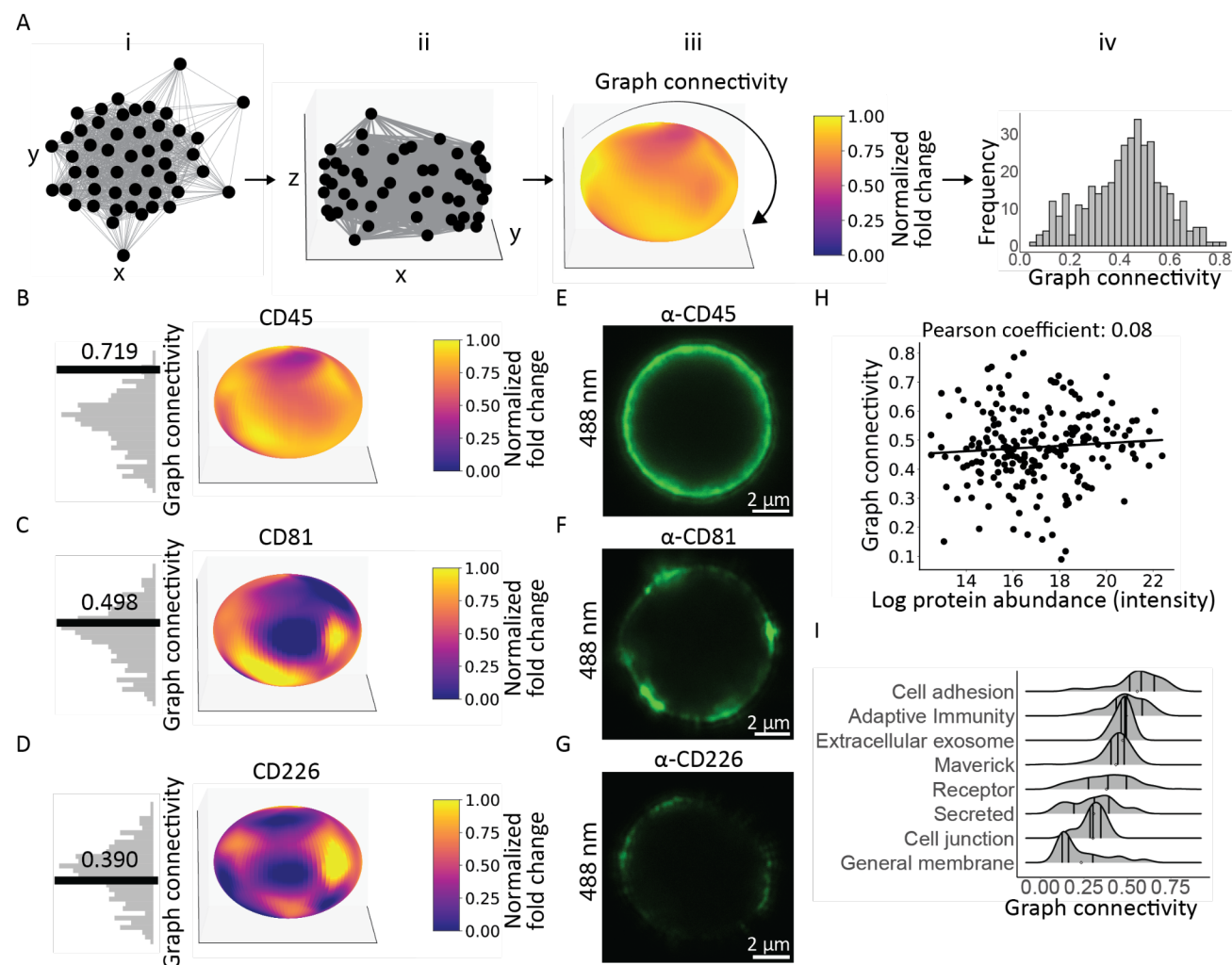


Figure 6. Three-dimensional network modeling to predict the surface distribution of the cell surface proteome. (A) Design walkthrough of 3D network heatmaps where (i) a 2D Kamada-Kawai network of high-quality proximity labeling bait proteins are (ii) transformed to lie on a 3D sphere where (iii) the normalized fold change of a given surface protein of interest is projected for each bait experiment (iv) and graph connectivity can be estimated for the entire cell surface proteome. (B) 3D network heatmap models of CD45, (D) CD81, and (E) CD226. (F-H) Representative confocal microscopy images of immunostained cell surface antigens CD45, CD81, and CD226. Antigens were visualized in the GFP channel using the dye Alexa Fluor 488. (H) Scatterplot comparing the protein abundance of Jurkat cell surface proteins measured using WGA-HRP for surface protein enrichment compared to their estimated

966 graph connectivity. (I) Distribution of estimated graph connectivity for cell surface proteins based on
967 their functional annotation.
968

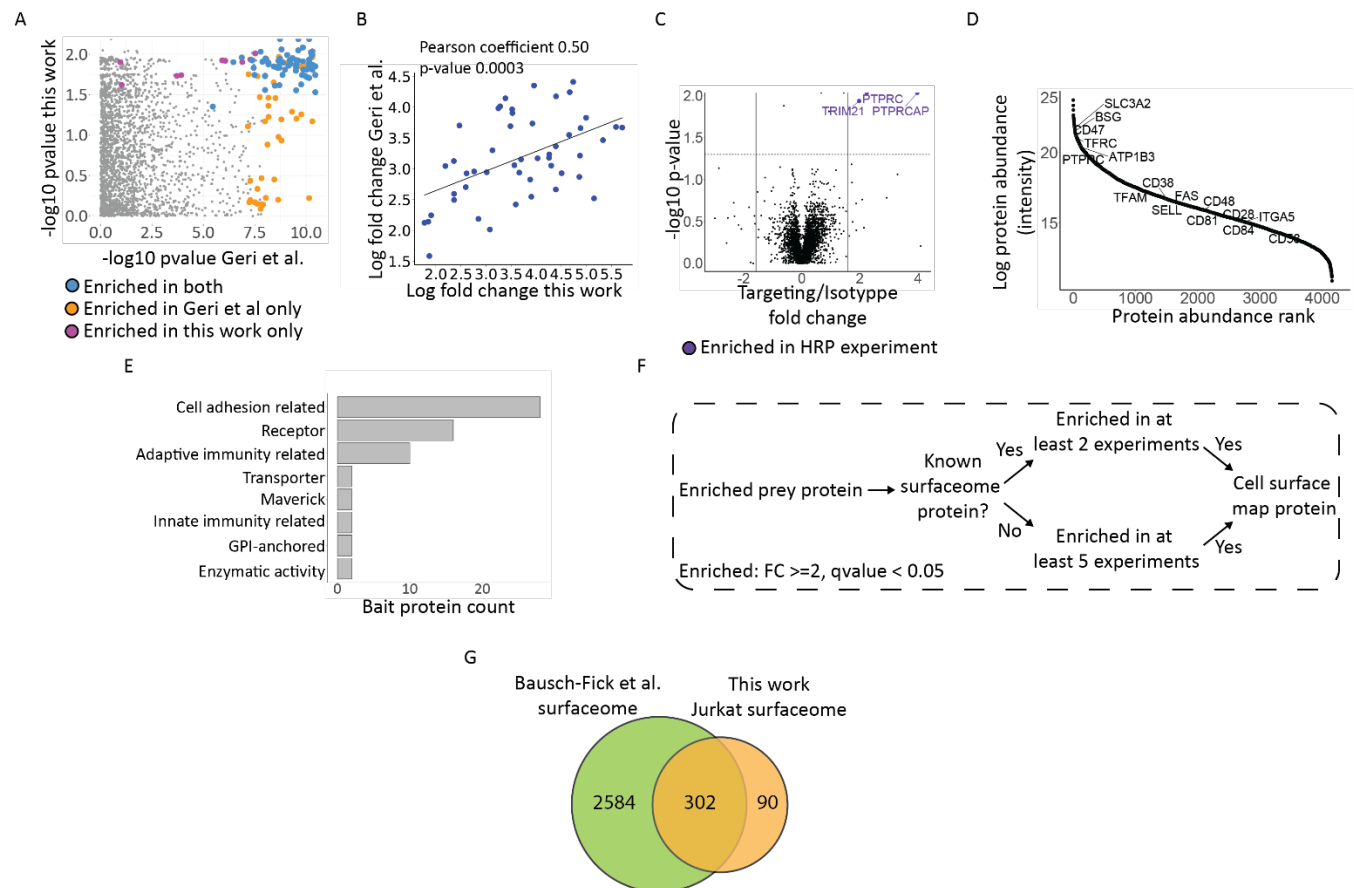


Figure S1 Validation of cell surface proximity labeling experiments collected in Jurkat cells. (A) Comparison of estimated FDR-corrected p-values from a CD45 targeting proximity labeling experiment to a previously published dataset. Proteins enriched (p-value < 0.05 and fold change ≥ 2). (B) Comparison of fold change values of enriched proteins from the CD45 proximity labeling experiment. (C) Volcano plot showing the results of a CD45-targeting μ Map proximity labeling experiment. All enriched proteins were also enriched in the HRP proximity labeling experiment. (D) Rank-abundance curve from a WGA-HRP surfaceomics experiment with proteins chosen as bait proteins for proximity labeling experiments highlighted. (E) Bar plot showing the different functional groups for the 65 bait proteins used for proximity labeling on Jurkat cells. (F) Workflow for determining if a protein is a cell surface protein expressed on Jurkat cells. (G) Venn diagram comparison of the estimated Jurkat surfaceome for this work compared to a published machine learning-predicted surfaceome.

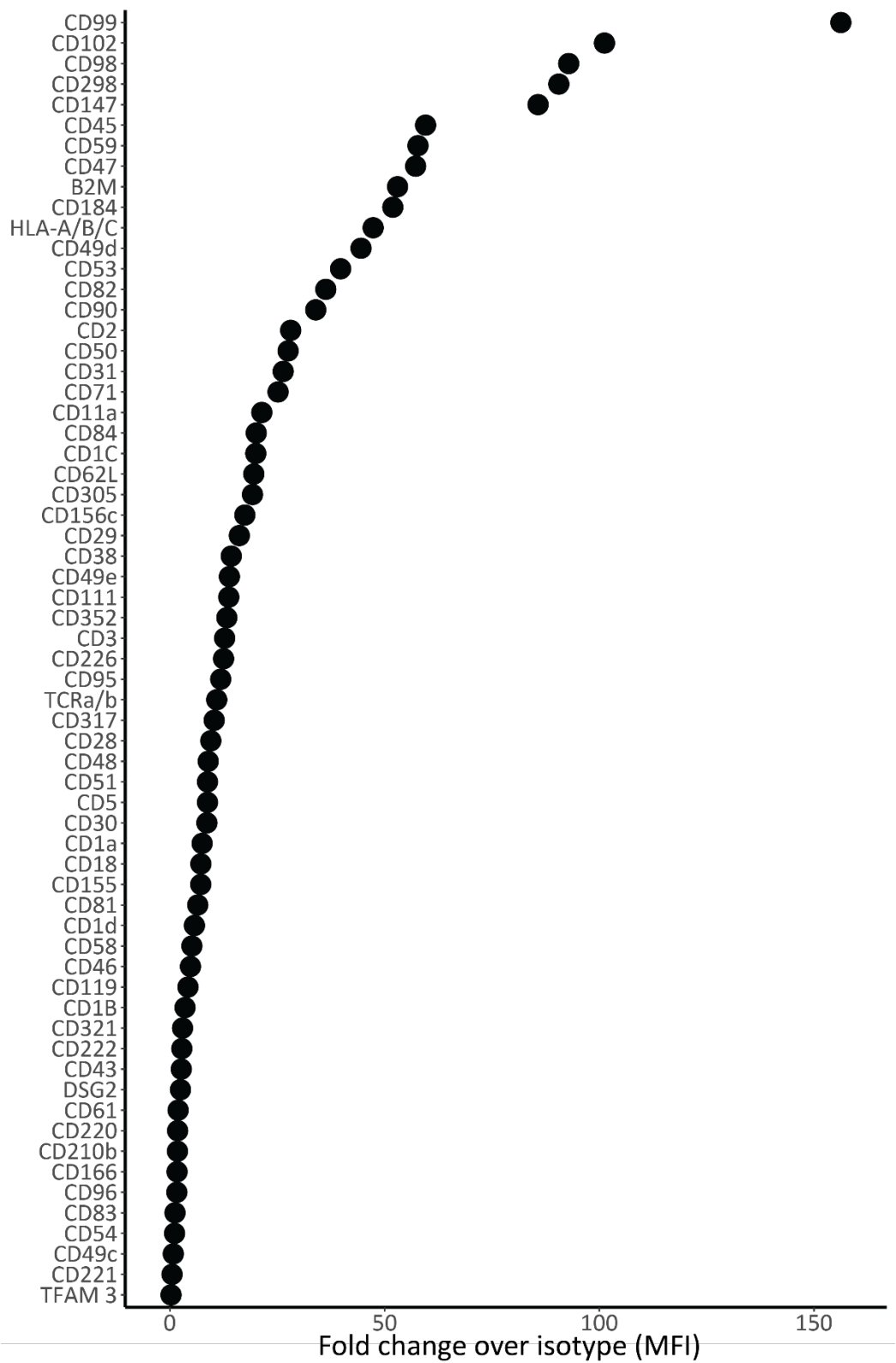


Figure S2. Flow cytometry surface staining results for all antibodies used for proximity labeling on the Jurkat cell surface.

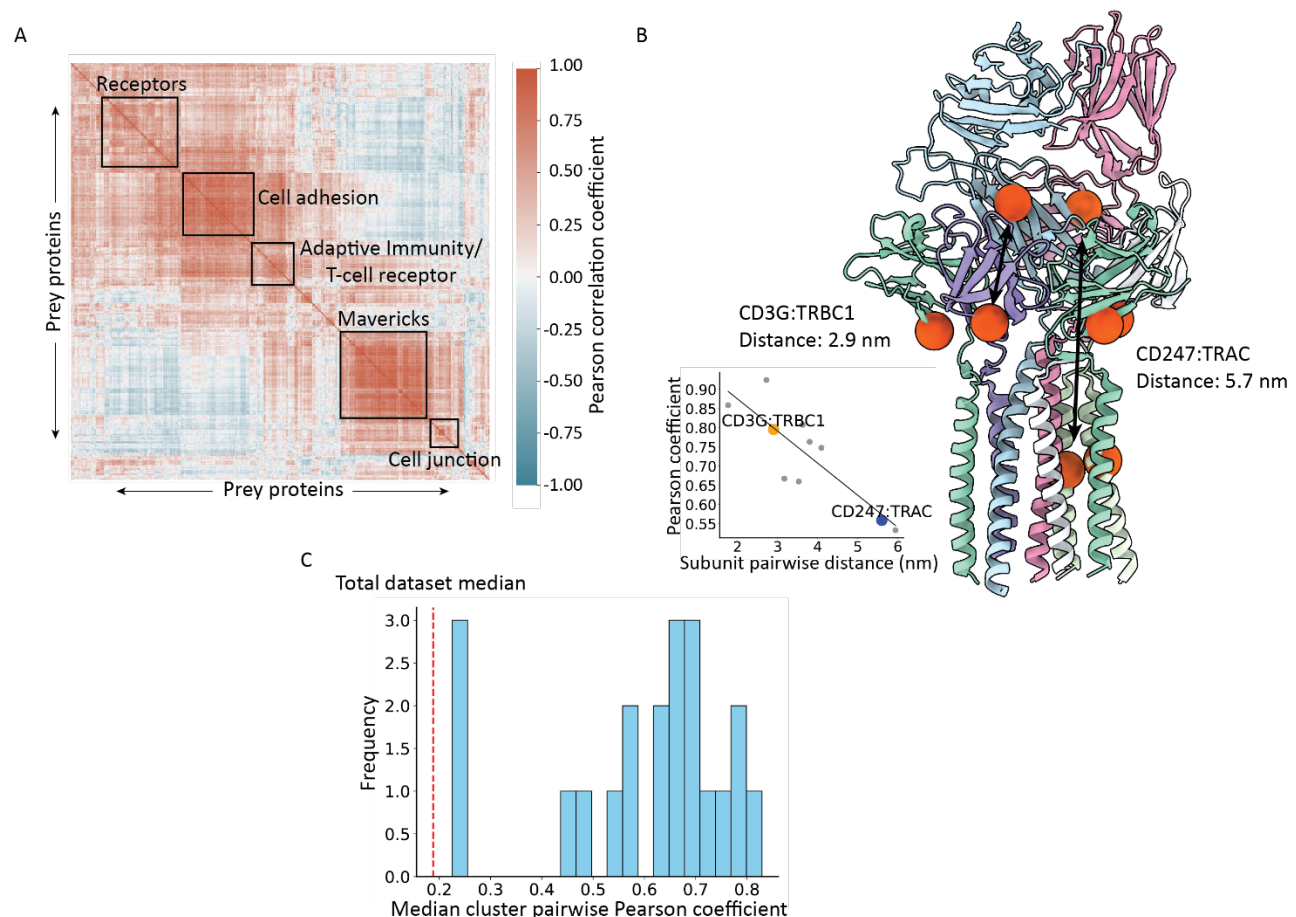


Figure S3 Additional data for the prey-centric analysis of the Jurkat cell surface proteome. (A) Hierarchically clustered heatmap of Pearson correlation values estimated from the fold change enrichment of all prey proteins on the Jurkat cell surface proteome. Enriched functional annotations are highlighted for the larger clusters of functional units. (B) Structure of the TCR-CD3 protein complex with the center of mass of each subunit highlighted as an orange sphere. Pairwise distance measurements were taken from the center of mass between two subunits and compared to the estimated pairwise Pearson correlation coefficient values. Two examples of pairwise measurements are highlighted on the structure and the scatterplot. (C) Distribution of median pairwise Pearson coefficients for the clusters shown in Figure 1F. The median value for pairwise Pearson coefficients for the entire Jurkat surfaceome is shown as a red line.

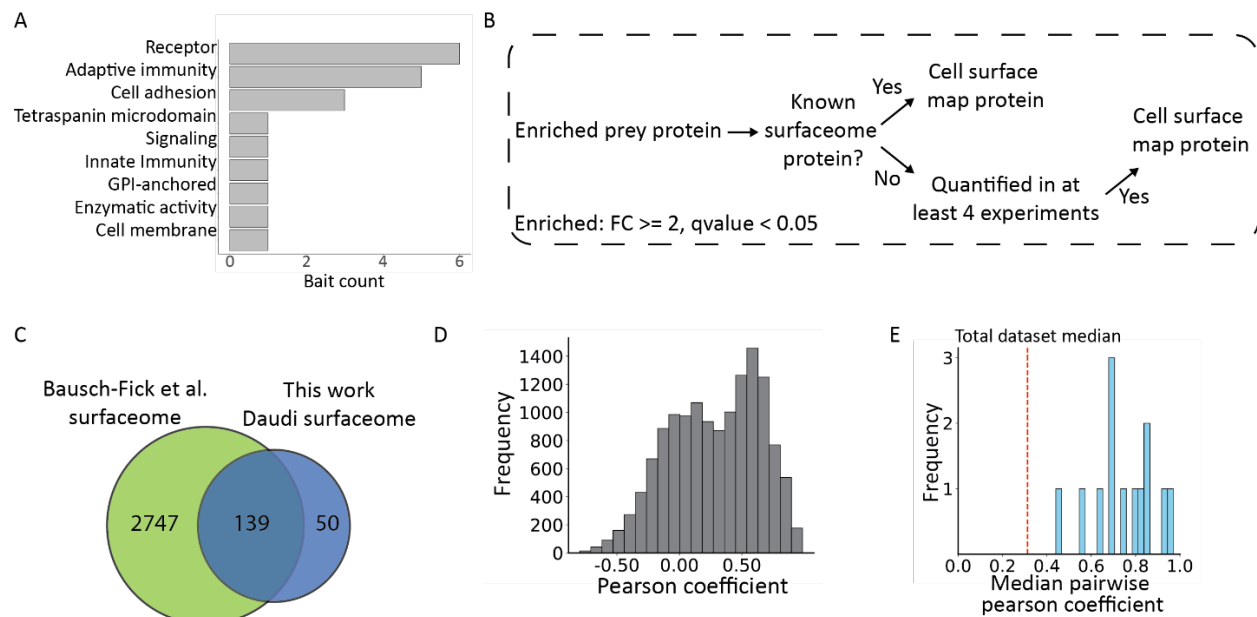


Figure S4 Characterization of cell surface protein characteristics on Daudi cell surfaces. (A) Bar plot showing the different functional annotations for the 20 bait proteins used for proximity labeling on Daudi cells. (B) Workflow for determining if a protein is a cell surface protein expressed on Daudi cells. (C) Venn diagram comparison of the estimated Daudi surfaceome for this work compared to a published machine learning-predicted surfaceome. (D) Distribution of all pairwise Pearson correlation coefficients for Daudi cell surface proteins. (E) Distribution of median pairwise Pearson coefficients for the clusters shown in Figure 2G. The median value for Pearson pairwise coefficients for the entire Daudi surfaceome is shown as a red line.

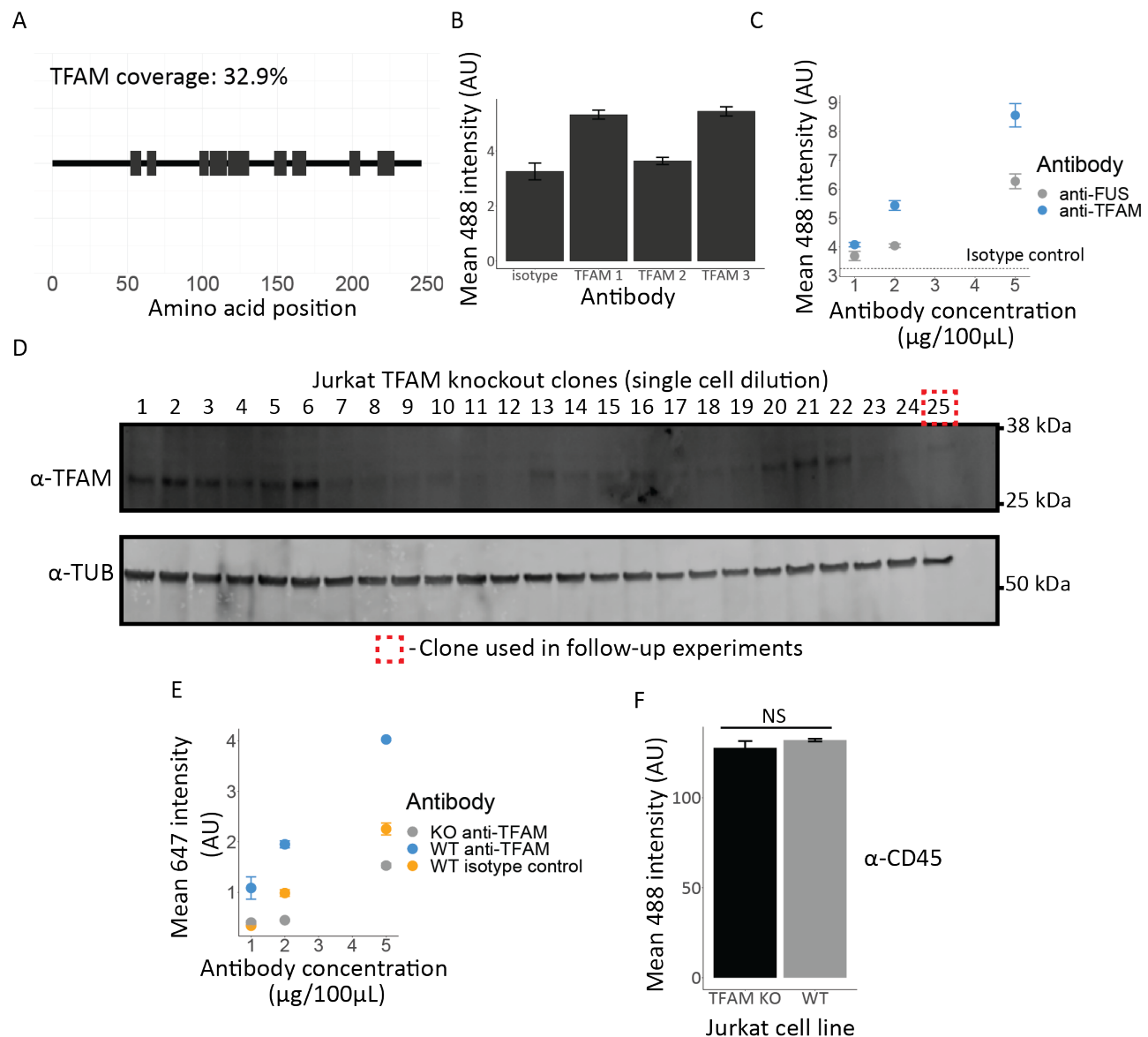


Figure S5 Validation and characterization of a maverick cell surface protein. (A) Peptide coverage of TFAM from proximity labeling experiments. The grey bars correspond to peptides detected by mass spectrometry. (B) Bar plot showing the flow cytometry results for proteins stained on the surface with 3 different anti-TFAM antibodies and an isotype control. 2 of the 3 anti-TFAM antibodies gave signal over the isotype control and anti-TFAM 3 (Proteintech) was used in all subsequent experiments. (C) Flow cytometry results from cell surface staining with an anti-TFAM and a previously validated anti-FUS antibody at 3 different concentrations. The isotype control binding is shown as a dashed line. (D) Western blot results following CRISPR-Cas9 knockout of TFAM in Jurkat cells and single-cell dilution to isolate clones with TFAM knocked out. Clone 25 was used in all subsequent experiments. (E) Flow cytometry results from cell surface staining with an anti-TFAM or isotype control antibody at 3 different concentrations on wildtype (WT) and TFAM knockout (KO) Jurkat cells. (F) Flow cytometry results from cell surface staining with an anti-CD45 antibody on Jurkat wildtype and TFAM knockout cells. Student's t-test p-value of 0.332.

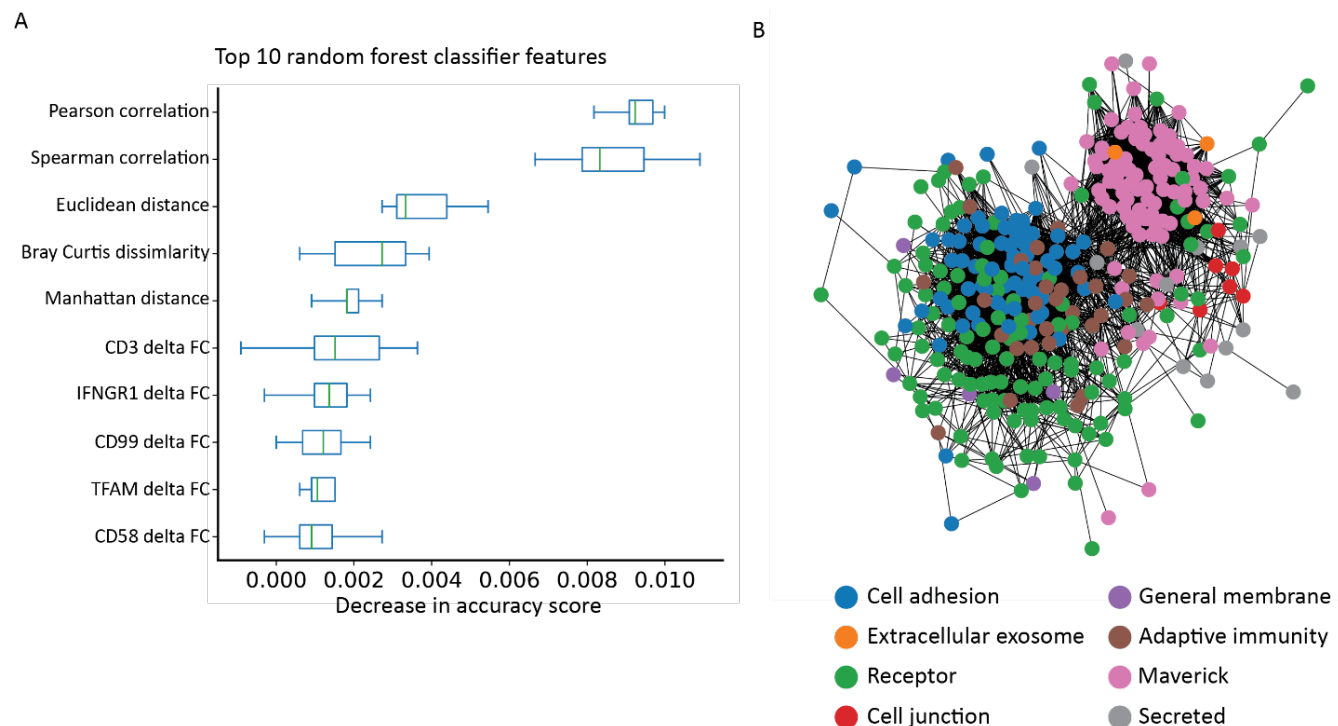


Figure S6. Additional information on the random forest classifier for predicting cell surface associations. (A) The top 10 most important random forest classifier features of the 70 used for predicting cell surface associations on Jurkat cells. For the boxplot the central line corresponds to the median and the whiskers extend to the most extreme data points within 10 times the interquartile range (IQR). (B) Kawada-Kawaii network where each node is a Jurkat cell surface protein, and the edges are drawn between nodes with an association score that passed a 10% false-discovery rate threshold. Nodes are colored according to the functional association of their assigned cluster shown in Figure 1F.

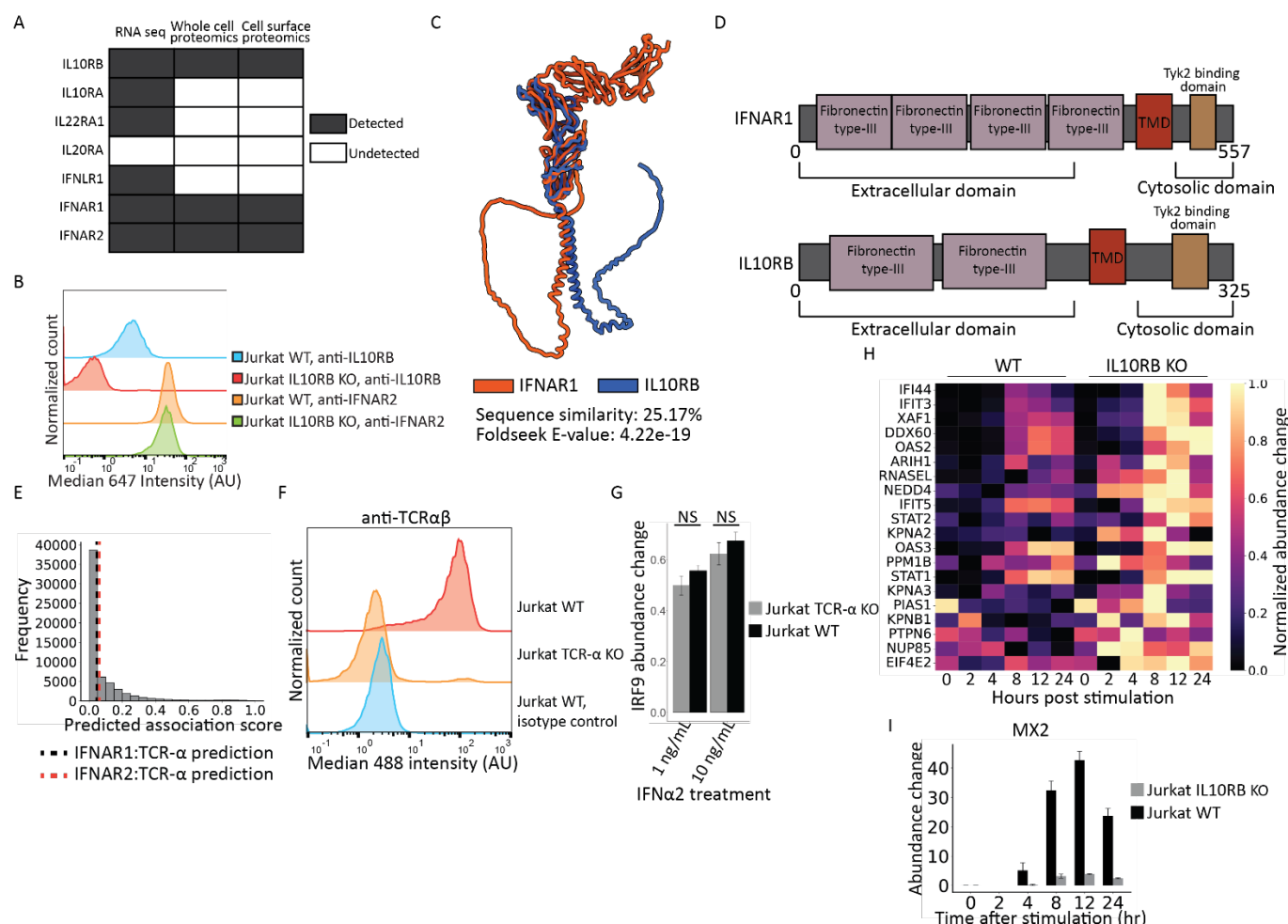


Figure S7 Characterization of IL10RB as a type I interferon regulator. (A) Identification of IL10RB interactors in Jurkat cells by RNA-seq, whole cell proteomics, and cell surface proteomics. RNA-seq data is from Felce et al. (2021). (B) Flow cytometry intensity distributions for Jurkat wild type (WT) and IL10RB knockout (KO) cells stained with an anti-IL10RB or anti-IFNAR2 antibody. A secondary antibody conjugated with Alexa Fluor 647 was used for fluorescent readout. (C) Aligned AlphaFold 3 predicted structures of IFNAR1 and IL10RB which share structural similarity as read out using Foldseek. (D) Protein domain architecture of IFNAR1 and IL10RB. (E) Distribution of predicted association scores for all 60,031 predictions with the prediction for TCR-α and the two subunits of the type I interferon protein complex highlighted. The association scores for TCR-α:IFNAR1 and TCR-α:IFNAR2 are 0.05 and 0.06 respectively. (F) Flow cytometry intensity distributions for Jurkat wild type (WT) and TCR-α knockout (KO) stained with an anti-TCR-αβ or isotype control antibody. (G) Intracellular flow cytometry measurement of cytosolic IRF9 abundance change in IFNα2 stimulated Jurkat wild type (black) and TCR-α knockout lines (grey) at IFNα2 concentrations of 1 and 10 ng/mL. Student's t-test p-values for at the different concentrations are 0.078 and 0.176 for 1 and 10 ng/mL respectively. (H) Heatmap showing the 20 interferon stimulated genes (ISGs) with the largest stimulation changes in the Jurkat IL10RB knockout (KO) line in comparison to the wild type line when stimulated with IFNα2 at various time points. (I) Abundance change plot of the ISG MX2 in the Jurkat WT (black) and IL10RB KO (grey) cell lines. Relative fold change was estimated in comparison to the 0 hour time point.

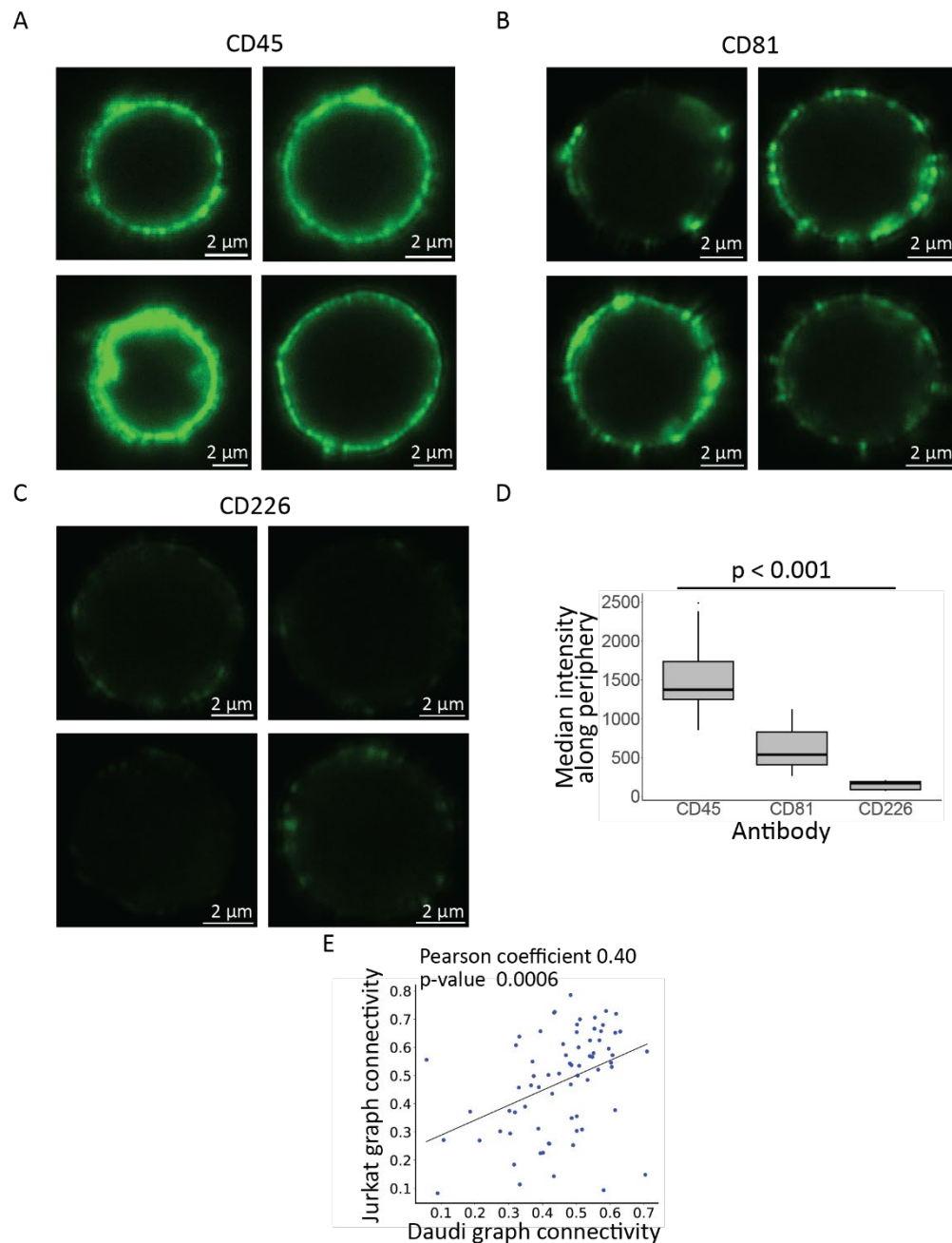


Figure S8 Additional data on the graph connectivity score and cell surface staining for confocal microscopy. (A) Additional representative images for cells stained with an anti-CD45 primary antibody and an Alexa Fluor 488 secondary antibody. (B) Additional representative images for cells stained with an anti-CD81 primary antibody and an Alexa Fluor 488 secondary antibody. (C) Additional representative images for cells stained with an anti-CD226 primary antibody and an Alexa Fluor 488 secondary antibody. (D) Box plot showing the median intensity along the cell periphery for cells stained with anti-

CD45, anti-CD81, and anti-CD226 antibodies. A one-way ANOVA was done to determine differences in the mean with an F-statistic of 32.328 and p-value of 1.549e-7. (E) Scatterplot showing the estimated graph connectivity for antigens both on Jurkat and Daudi cells. The line represents the line of best fit for the data.

1078

References

- 1079 1. Bausch-Fluck D, Goldmann U, Müller S, et al. The in silico human surfaceome. *Proc Natl Acad Sci*.
1080 2018;115(46):E10988-E10997. doi:10.1073/pnas.1808790115
- 1081 2. Tomas A, Futter CE, Eden ER. EGF receptor trafficking: consequences for signaling and cancer.
1082 *Trends Cell Biol*. 2014;24(1):26-34. doi:10.1016/j.tcb.2013.11.002
- 1083 3. Gartner ZJ, Prescher JA, Lavis LD. Unraveling cell-to-cell signaling networks with chemical biology.
1084 *Nat Chem Biol*. 2017;13(6):564-568. doi:10.1038/nchembio.2391
- 1085 4. Kabel AM. Tumor markers of breast cancer: New perspectives. *J Oncol Sci*. 2017;3(1):5-11.
1086 doi:10.1016/j.jons.2017.01.001
- 1087 5. Belov L, Zhou J, Christopherson RI. Cell Surface Markers in Colorectal Cancer Prognosis. *Int J Mol Sci*.
1088 2011;12(1):78-113. doi:10.3390/ijms12010078
- 1089 6. Dodson LF, Hawkins WG, Goedegebuure P. Potential Targets for Pancreatic Cancer
1090 Immunotherapeutics. *Immunotherapy*. Published online April 1, 2011. doi:10.2217/imt.11.10
- 1091 7. Jacobson K, Liu P, Lagerholm BC. The Lateral Organization and Mobility of Plasma Membrane
1092 Components. *Cell*. 2019;177(4):806-819. doi:10.1016/j.cell.2019.04.018
- 1093 8. Kalappurakkal JM, Anilkumar AA, Patra C, Zanten TS van, Sheetz MP, Mayor S. Plasma Membrane
1094 Nanodomains as an Integrator of Substrate Encoded Mechano-chemical Signals. *Biophys J*.
1095 2020;118(3):190a. doi:10.1016/j.bpj.2019.11.1152
- 1096 9. Qin W, Cho KF, Cavanagh PE, Ting AY. Deciphering molecular interactions by proximity labeling. *Nat*
1097 *Methods*. 2021;18(2):133-143. doi:10.1038/s41592-020-01010-5
- 1098 10. Kageler L, Perr J, Flynn RA. Tools to investigate the cell surface: Proximity as a central concept in
1099 glycoRNA biology. *Cell Chem Biol*. 2024;31(6):1132-1144. doi:10.1016/j.chembiol.2024.04.015
- 1100 11. Lam SS, Martell JD, Kamer KJ, et al. Directed evolution of APEX2 for electron microscopy and
1101 proximity labeling. *Nat Methods*. 2015;12(1):51-54. doi:10.1038/nmeth.3179
- 1102 12. Rees JS, Li XW, Perrett S, Lilley KS, Jackson AP. Selective Proteomic Proximity Labeling Assay Using
1103 Tyramide (SPPLAT): A Quantitative Method for the Proteomic Analysis of Localized Membrane-
1104 Bound Protein Clusters. *Curr Protoc Protein Sci*. 2015;80:19.27.1-19.27.18.
1105 doi:10.1002/0471140864.ps1927s80
- 1106 13. Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and
1107 interacting proteins in mammalian cells. *J Cell Biol*. 2012;196(6):801-810.
1108 doi:10.1083/jcb.201112098
- 1109 14. Branon TC, Bosch JA, Sanchez AD, et al. Efficient proximity labeling in living cells and organisms with
1110 TurboID. *Nat Biotechnol*. 2018;36(9):880-887. doi:10.1038/nbt.4201
- 1111 15. Geri JB, Oakley JV, Reyes-Robles T, et al. Microenvironment mapping via Dexter energy transfer on
1112 immune cells. *Science*. 2020;367(6482):1091-1097. doi:10.1126/science.aay4106
- 1113 16. Knutson SD, Buksh BF, Huth SW, Morgan DC, MacMillan DWC. Current advances in photocatalytic
1114 proximity labeling. *Cell Chem Biol*. 2024;31(6):1145-1161. doi:10.1016/j.chembiol.2024.03.012
- 1115 17. Lin Z, Schaefer K, Lui I, et al. Multiscale photocatalytic proximity labeling reveals cell surface
1116 neighbors on and between cells. *Science*. 2024;385(6706):eadl5763. doi:10.1126/science.adl5763
- 1117 18. Müller M, Gräbnitz F, Barandun N, et al. Light-mediated discovery of surfaceome nanoscale
1118 organization and intercellular receptor interaction networks. *Nat Commun*. 2021;12(1):7036.
1119 doi:10.1038/s41467-021-27280-x
- 1120 19. Sharma P, Varma R, Sarasij RC, et al. Nanoscale Organization of Multiple GPI-Anchored Proteins in
1121 Living Cell Membranes. *Cell*. 2004;116(4):577-589. doi:10.1016/S0092-8674(04)00167-9
- 1122 20. Torreno-Pina JA, Castro BM, Manzo C, Buschow SI, Cambi A, Garcia-Parajo MF. Enhanced receptor-
1123 clathrin interactions induced by N-glycan-mediated membrane micropatterning. *Proc Natl Acad Sci*.
1124 2014;111(30):11037-11042. doi:10.1073/pnas.1402041111

21. Wäldchen F, Schlegel J, Götz R, et al. Whole-cell imaging of plasma membrane receptors by 3D lattice light-sheet dSTORM. *Nat Commun.* 2020;11(1):887. doi:10.1038/s41467-020-14731-0
22. Perr J, Langen A, Almahayni K, et al. RNA binding proteins and glycoRNAs form domains on the cell surface for cell penetrating peptide entry. Published online September 5, 2023:2023.09.04.556039. doi:10.1101/2023.09.04.556039
23. Garcia-Parajo MF, Cambi A, Torreno-Pina JA, Thompson N, Jacobson K. Nanoclustering as a dominant feature of plasma membrane organization. *J Cell Sci.* 2014;127(23):4995-5005. doi:10.1242/jcs.146340
24. Huttlin EL, Bruckner RJ, Navarrete-Perea J, et al. Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell.* 2021;184(11):3022-3040.e28. doi:10.1016/j.cell.2021.04.011
25. Go CD, Knight JDR, Rajasekharan A, et al. A proximity-dependent biotinylation map of a human cell. *Nature.* 2021;595(7865):120-124. doi:10.1038/s41586-021-03592-2
26. Havugimana PC, Hart GT, Nepusz T, et al. A Census of Human Soluble Protein Complexes. *Cell.* 2012;150(5):1068-1081. doi:10.1016/j.cell.2012.08.011
27. Oakley JV, Buksh BF, Fernández DF, et al. Radius measurement via super-resolution microscopy enables the development of a variable radii proximity labeling platform. *Proc Natl Acad Sci.* 2022;119(32):e2203027119. doi:10.1073/pnas.2203027119
28. Searle BC, Swearingen KE, Barnes CA, et al. Generating high quality libraries for DIA MS with empirically corrected peptide predictions. *Nat Commun.* 2020;11(1):1548. doi:10.1038/s41467-020-15346-1
29. Tabb DL, Kaniyar MH, Bringas OGR, et al. Interrogating data-independent acquisition LC-MS/MS for affinity proteomics. *J Proteins Proteomics.* 2024;15(3):281-298. doi:10.1007/s42485-024-00166-4
30. Kirkemo LL, Elledge SK, Yang J, et al. Cell-surface tethered promiscuous biotinylators enable comparative small-scale surface proteomic analysis of human extracellular vesicles and cells. *eLife.* 2022;11:e73982. doi:10.7554/eLife.73982
31. Dong D, Zheng L, Lin J, et al. Structural basis of assembly of the human T cell receptor–CD3 complex. *Nature.* 2019;573(7775):546-552. doi:10.1038/s41586-019-1537-0
32. McWhite CD, Papoulas O, Drew K, et al. A Pan-plant Protein Complex Map Reveals Deep Conservation and Novel Assemblies. *Cell.* 2020;181(2):460-474.e14. doi:10.1016/j.cell.2020.02.049
33. Sae-Lee W, McCafferty CL, Verbeke EJ, et al. The protein organization of a red blood cell. *Cell Rep.* 2022;40(3). doi:10.1016/j.celrep.2022.111103
34. Sherman BT, Hao M, Qiu J, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 2022;50(W1):W216-W221. doi:10.1093/nar/gkac194
35. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44-57. doi:10.1038/nprot.2008.211
36. Kim JW, Cho YB, Lee S. Cell Surface GRP94 as a Novel Emerging Therapeutic Target for Monoclonal Antibody Cancer Therapy. *Cells.* 2021;10(3):670. doi:10.3390/cells10030670
37. Ma Z, Cao M, Liu Y, et al. Mitochondrial F1Fo-ATP synthase translocates to cell surface in hepatocytes and has high activity in tumor-like acidic and hypoxic environment. *Acta Biochim Biophys Sin.* 2010;42(8):530-537. doi:10.1093/abbs/gmq063
38. Kao YC, Chang YW, Lai CP, et al. Ectopic ATP synthase stimulates the secretion of extracellular vesicles in cancer cells. *Commun Biol.* 2023;6(1):1-17. doi:10.1038/s42003-023-05008-5
39. Savojardo C, Martelli PL, Fariselli P, Casadio R. TPpred3 detects and discriminates mitochondrial and chloroplastic targeting peptides in eukaryotic proteins. *Bioinformatics.* 2015;31(20):3269-3275. doi:10.1093/bioinformatics/btv367
40. Wang T, Ma F, Qian H li. Defueling the cancer: ATP synthase as an emerging target in cancer

- therapy. *Mol Ther - Oncolytics*. 2021;23:82-95. doi:10.1016/j.omto.2021.08.015
41. Chang YW, Wang CC, Yin CF, Wu CH, Huang HC, Juan HF. Quantitative phosphoproteomics reveals ectopic ATP synthase on mesenchymal stem cells to promote tumor progression *via* ERK/c-Fos pathway activation. *Mol Cell Proteomics*. 2022;21(6):100237. doi:10.1016/j.mcpro.2022.100237
42. Chang YW, Hsu CL, Tang CW, Chen XJ, Huang HC, Juan HF. Multiomics Reveals Ectopic ATP Synthase Blockade Induces Cancer Cell Death via a lncRNA-mediated Phospho-signaling Network. *Mol Cell Proteomics MCP*. 2020;19(11):1805-1825. doi:10.1074/mcp.RA120.002219
43. Denisko D, Hoffman MM. Classification and interaction in random forests. *Proc Natl Acad Sci*. 2018;115(8):1690-1692. doi:10.1073/pnas.1800256115
44. Riley JK, Takeda K, Akira S, Schreiber RD. Interleukin-10 Receptor Signaling through the JAK-STAT Pathway: REQUIREMENT FOR TWO DISTINCT RECEPTOR-DERIVED SIGNALS FOR ANTI-INFLAMMATORY ACTION *. *J Biol Chem*. 1999;274(23):16513-16521. doi:10.1074/jbc.274.23.16513
45. Hutchins AP, Diez D, Miranda-Saavedra D. The IL-10/STAT3-mediated anti-inflammatory response: recent developments and future challenges. *Brief Funct Genomics*. 2013;12(6):489-498. doi:10.1093/bfpg/elt028
46. Felce SL, Farnie G, Dustin ML, Felce JH. RNA-Seq analysis of early transcriptional responses to activation in the leukaemic Jurkat E6.1 T cell line. *Wellcome Open Res*. 2020;5:42. doi:10.12688/wellcomeopenres.15748.2
47. Ragimbeau J, Dondi E, Alcover A, Eid P, Uzé G, Pellegrini S. The tyrosine kinase Tyk2 controls IFNAR1 cell surface expression. *EMBO J*. 2003;22(3):537-547. doi:10.1093/emboj/cdg038
48. van Kempen M, Kim SS, Tumescheit C, et al. Fast and accurate protein structure search with Foldseek. *Nat Biotechnol*. 2024;42(2):243-246. doi:10.1038/s41587-023-01773-0
49. Forero A, Ozarkar S, Li H, et al. Differential Activation of the Transcription Factor IRF1 Underlies the Distinct Immune Responses Elicited by Type I and Type III Interferons. *Immunity*. 2019;51(3):451-464.e6. doi:10.1016/j.immuni.2019.07.007
50. Dondi E, Pattyn E, Lutfalla G, et al. Down-modulation of Type 1 Interferon Responses by Receptor Cross-competition for a Shared Jak Kinase*. *J Biol Chem*. 2001;276(50):47004-47012. doi:10.1074/jbc.M104316200
51. Karlsson F, Kallas T, Thiagarajan D, et al. Molecular pixelation: spatial proteomics of single cells by sequencing. *Nat Methods*. 2024;21(6):1044-1052. doi:10.1038/s41592-024-02268-9
52. Prashar A, Bussi C, Fearn A, et al. Lysosomes drive the piecemeal removal of mitochondrial inner membrane. *Nature*. 2024;632(8027):1110-1117. doi:10.1038/s41586-024-07835-w
53. Whitlock JM, Leikina E, Melikov K, et al. Cell surface-bound La protein regulates the cell fusion stage of osteoclastogenesis. *Nat Commun*. 2023;14(1):616. doi:10.1038/s41467-023-36168-x
54. Chai P, Perr J, Kageler L, et al. Cell surface ribonucleoproteins cluster with heparan sulfate to regulate growth factor signaling. Published online July 25, 2024:2024.07.25.605163. doi:10.1101/2024.07.25.605163
55. Carlini V, Noonan DM, Abdalalem E, et al. The multifaceted nature of IL-10: regulation, role in immunological homeostasis and its relevance to cancer, COVID-19 and post-COVID conditions. *Front Immunol*. 2023;14. doi:10.3389/fimmu.2023.1161067
56. Freeman SA, Vega A, Riedl M, et al. Transmembrane pickets connect cyto-and pericellular-skeletons forming barriers to receptor engagement. *Cell*. 2018;172(1-2):305-317.e10. doi:10.1016/j.cell.2017.12.023
57. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods*. 2020;17(1):41-44. doi:10.1038/s41592-019-0638-x
58. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods*.

1221 2017;14(5):513-520. doi:10.1038/nmeth.4256

1222 59. Yu F, Teo GC, Kong AT, et al. Analysis of DIA proteomics data using MSFragger-DIA and FragPipe

1223 computational platform. *Nat Commun.* 2023;14(1):4154. doi:10.1038/s41467-023-39869-5

1224 60. Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive

1225 analysis of (prote)omics data. *Nat Methods.* 2016;13(9):731-740. doi:10.1038/nmeth.3901

1226 61. Ruepp A, Brauner B, Dunger-Kaltenbach I, et al. CORUM: the comprehensive resource of

1227 mammalian protein complexes. *Nucleic Acids Res.* 2008;36(Database issue):D646-D650.

1228 doi:10.1093/nar/gkm936

1229 62. Drew K, Wallingford JB, Marcotte EM. hu.MAP 2.0: integration of over 15,000 proteomic

1230 experiments builds a global compendium of human multiprotein assemblies. *Mol Syst Biol.*

1231 2021;17(5):e10016. doi:10.15252/msb.202010016

1232 63. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein–protein association

1233 networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids*

1234 *Res.* 2023;51(D1):D638-D646. doi:10.1093/nar/gkac1000

1235 64. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for

1236 interaction datasets. *Nucleic Acids Res.* 2006;34(suppl_1):D535-D539. doi:10.1093/nar/gkj109

1237 65. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn*

1238 *Res.* 2011;12(85):2825-2830.