

## **Broad and thematic remodeling of the surface glycoproteome on isogenic cells transformed with driving proliferative oncogenes**

Kevin K. Leung<sup>1,5</sup>, Gary M. Wilson<sup>2,5</sup>, Lisa L. Kirkemo<sup>1</sup>, Nicholas M. Riley<sup>2,4</sup>, Joshua J. Coon<sup>2,3</sup>, James A. Wells<sup>1\*</sup>

<sup>1</sup>Department of Pharmaceutical Chemistry, UCSF, San Francisco, CA, USA

Departments of Chemistry<sup>2</sup> and Biomolecular Chemistry<sup>3</sup>, University of Wisconsin-Madison, Madison, WI, 53706, USA

<sup>4</sup>Present address Department of Chemistry, Stanford University, Stanford, CA, 94305, USA

<sup>5</sup>These authors contributed equally

\*To whom correspondence should be addressed

## Abstract:

The cell surface proteome, the surfaceome, is the interface for engaging the extracellular space in normal and cancer cells. Here we apply quantitative proteomics of N-linked glycoproteins to reveal how a collection of some 700 surface proteins is dramatically remodeled in an isogenic breast epithelial cell line stably expressing any of six of the most prominent proliferative oncogenes, including the receptor tyrosine kinases, EGFR and HER2, and downstream signaling partners such as KRAS, BRAF, MEK and AKT. We find that each oncogene has somewhat different surfaceomes but the functions of these proteins are harmonized by common biological themes including up-regulation of nutrient transporters, down-regulation of adhesion molecules and tumor suppressing phosphatases, and alteration in immune modulators. Addition of a potent MEK inhibitor that blocks MAPK signaling brings each oncogene-induced surfaceome back to a common state reflecting their strong dependence on the MAPK pathway to propagate signaling. Using a recently developed glyco-proteomics method of activated ion electron transfer dissociation (AI-ETD) we found massive oncogene-induced changes in 142 N-linked glycans and differential increases in complex hybrid glycans especially for KRAS and HER2 oncogenes. Overall, these studies provide a broad systems level view of how specific driver oncogenes remodel the surface glycoproteome in a cell autologous fashion, and suggest possible surface targets, and combinations thereof, for drug and biomarker discovery.

## Keywords:

Oncogenes, Glycoproteomics, Surfaceome, MAPK signaling pathway

## **Significant statement**

The cell surface glycoproteome (surfaceome) mediates interactions between the cell and the extracellular environment, and is a major target for immunotherapy in cancer. Using state-of-the-art proteomics methods, we compared how six neighboring proliferative oncogenes cause large and bidirectional expression of some 700 surface proteins and the 142 different glycans that decorate them. While each oncogene induces large and somewhat unique glycoproteomes relative to non-transformed cells, we find common functional consequences that are massively reversed by small molecule inhibition of the MAPK pathway. This large-scale comparative study provides important insights for how oncogenes remodel isogenic cells in a cell autologous fashion, and suggest possible new opportunities for antibody drug discovery in more complex tumor settings.

## **Introduction:**

The cell surface proteome, or surfaceome, is the main interface for cellular signaling, nutrient homeostasis, cellular adhesion, and defines immunologic identity. To survive, cancer cells adjust to promote increased nutrient import, pro-growth signaling, and evasion of immunological surveillance among others (Hanahan and Weinberg, 2011). There are some 4,000 different membrane proteins encoded in the human genome (Gonzalez et al., 2010; Wallin and von Heijne, 1998), yet antibodies to only about two dozen cell surface targets have been approved for therapeutic intervention, prompting the need to discover novel tumor specific antigens (Carter and Lazar, 2018; Chames et al., 2009).

Recent surfaceome studies in an isogenic MCF10A breast epithelial cell line transformed with oncogenic KRAS have identified more than two dozen up-regulated surface proteins that function in a cell autologous fashion to promote increased cell proliferation, metastasis, metabolic activity and immunologic suppression (Martinko et al., 2018; Ye et al., 2016; Ying et al., 2012). Many of the most powerful oncogenes are linked to KRAS and the MAPK pathway including overactivation of receptor tyrosine kinases, such as EGFR and HER2, or mutations in BRAF or RAS (**Fig 1A**). It is well-known these neighboring oncogenes are mutually exclusive in human tumors. For example, lung cancer patients with oncogenic EGFR mutations seldom harbor oncogenic KRAS or BRAF mutations and vice versa (Cancer Genome Atlas Research Network et al., 2014). Also, recent evidence show that oncogene co-expression can induce synthetic lethality or oncogene-induced senescence further reinforcing that activation of one oncogene without the others is preferable in cancer (Cisowski et al., 2015; Unni et al., 2015).

There is also substantial evidence that cell surface glycosylation is altered in cancer. Incomplete or truncated synthesis, extended branching, core fucosylation, and sialylation of cell surface glycans are hallmarks of tumor cells, which alter physiological mechanisms of cell-cell adhesion, communication, and immune system recognition (Fernandes et al., 1991; Hakomori, 1996; Hudak et al., 2014; Xia et al., 2016; Xiao et al., 2016; Kannagi et al., 2010; Varki and Gagneux, 2012). Over the past decade, chemical glycoproteomics has revealed specific examples of altered glycosylation and heterogeneity of glycans on particular proteins (Palaniappan and Bertozzi, 2016). However, we do not know how expression of different oncogenes globally alters glycosylation on individual proteins at a proteome-wide scale. Very recently, hybrid type ETD methods, such as activated ion electron transfer dissociation (AI-ETD) and ETD with supplemental activation (EThcD), have emerged as forerunners for glycoproteomic profiling due to their ability to generate sequence informative tandem mass spectra that links both the peptide backbone and corresponding glycan modification (Liu et al., 2017; Riley et al., 2019; Yang et al., 2017). These techniques can provide granularity to the altered glycosylation states of cell surface proteins upon oncogenic transformation.

Here we address how neighboring driver oncogenes (KRAS<sup>G12V</sup>, HER2 overexpression, EGFR<sup>L858R</sup>, BRAF<sup>V600E</sup>, phosphomimetic MEK<sup>S218D/S222D</sup>, and myristoylated AKT) stably expressed in a common non-cancerous isogenic epithelial cell alters the surface glycoproteome. Using cell-surface capture (CSC)(Bausch-Fluck et al., 2015; Wollscheid et al., 2009) and AI-ETD glycoproteomics (Riley et al., 2019), we found each oncogene

induced common and unique sets of up and down-regulated surface proteins and associated glycans. These sets of protein revealed common biological themes including increased expression of nutrient transporters and decreased expression of adhesion molecules. These effects were massively reversed upon inhibition of the MAPK pathway emphasizing its central importance in remodeling the surfaceome. These oncogene-induced surface proteins highlight targets or combinations to consider for immunotherapy.

## Results:

### Phenotypic analysis of oncogene transformed MCF10A cells

Tumor biology is highly complex and varies depending on the cell type of origin, stage, oncogenes, epigenome, stroma, vascularization, the immune system and metabolism. We deliberately took a simplistic reductionist approach to compare how specific proliferative oncogenes alter the cell surface proteome in a cell autologous fashion by using isogenic cells stably transformed with different oncogenes. No single cell type, culturing condition, or context is representative of all or even one cancer type. For practical reasons we chose the spontaneously immortalized breast epithelial MCF10A as our parental cell line because it is often used as a neutral starting point for oncogene studies (Allen-Petersen et al., 2014; Martins et al., 2015; Qu et al., 2015; Soule et al., 1990). Although MCF10A certainly does not recapitulate the diversity of mammary cell biology, it is of epithelial origin like most common tumors; it is non tumorigenic, requires growth factors for survival and does not harbor gene amplifications or other chromosomal aberrations typical of advanced cancer cell lines. Our intent is to compare how neighboring driving oncogenes remodel their surface glycoproteomes in isogenic cells in a cell autologous fashion.

Lentivirus was used to stably transform MCF10A cells with six prevalent oncogenes that are neighbors in proliferative signaling: HER2 overexpression, EGFR<sup>L858R</sup>, KRAS<sup>G12V</sup>, BRAF<sup>V600E</sup>, the phosphomimetic MEK<sup>S218D/S222D</sup> (MEK<sup>DD</sup>) and myristoylated AKT (AKT<sup>myr</sup>) (**Figure 1A**). Remarkably, the morphologies of each of the transformed cells varied from each other when cultured in the absence of growth factors indicative of differences in proteomic landscape (**Figure 1B**). In the absence of growth factors, cells

transformed with HER2 and KRAS<sup>G12V</sup> grew to confluence, while cells harboring EGFR<sup>L858R</sup>, BRAF<sup>V600E</sup>, and MEK<sup>DD</sup> did not reach confluency, indicative of contact dependent growth inhibition. Cells transformed with AKT<sup>myr</sup>, which signals through a parallel pathway relative to the other five oncogenes, had the most dramatic morphology change, displaying vertically stacked clusters of cells. Unlike the parental cell line, all the MCF10A cells stably transformed with any of the six oncogenes proliferated in the absence of growth factors to various degrees (**Figure 1C**). The HER2 and KRAS<sup>G12V</sup> transformed cells proliferated most rapidly in the absence of growth factors, and even comparable or faster to the untransformed MCF10A cultured in the presence of growth factors. The HER2 and KRAS<sup>G12V</sup> cells also lifted off the plates much more readily than the others, suggesting reduced adhesion phenotype.

These oncogenes can drive multiple branched pathways yet it was previously shown that inhibition of the MAPK pathway with the potent and selective MEK inhibitor (PD032590, MEKi) significantly reverses the surfaceome changes of MCF10A cells transformed with KRAS<sup>G12V</sup> (Martinko et al., 2018). Indeed, MEKi substantially hampered growth for all cell lines either in the absence or presence of growth factors (**Figure 1D, Figure S1**). Overexpression of HER2 was most resistant to MEKi, followed by KRAS<sup>G12V</sup> and AKT<sup>myr</sup>, whereas cells containing EGFR<sup>L858R</sup>, BRAF<sup>V600E</sup>, or MEK<sup>DD</sup> were most sensitive to MEKi.

### **Differential expression of oncogene-induced surfaceomes in MCF10A cells**

We next probed how the cell surfaceome is altered in the oncogene transformed cells compared to the empty vector control. N-glycosylation is present on >85% of cell surface



proteins (Apweiler et al., 1999) and can be exploited to capture the N-glycosylated proteins using a biotin hydrazide enrichment method called cell surface capture (CSC) (Bausch-Fluck et al., 2018; Wollscheid et al., 2009). Here, we utilized a CSC protocol coupled with stable isotope labeling by amino acids in cell culture (SILAC) to compare the surfaceomes from the oncogene transformed MCF10A cells to the empty vector control (**Figure S2**) (Leung et al., 2019). We identified and quantified a total of 654 cell surface proteins across the six oncogenic cell lines (**Figure 2A**). Remarkably, the expression for 43% of the aggregate surface proteins (280 of 654) was altered by at least two-fold for the oncogene transformed cell lines relative to empty vector control reflecting significant remodeling of the surfaceomes. In each of the six datasets, we observed at least two-fold changes for 100 to 150 different surface proteins (**Figure 2A, Figure S3A-F**); these changes were evenly split between up or down regulated sets reflecting bi-directional remodeling. Many of the differentially expressed proteins overlapped, but each cell line had a substantial number of uniquely differentially regulated proteins, presumably resulting from slight differences in signaling between each oncogene. Although it is well known that correlations between protein and RNA levels are not often strongly correlated (Lundberg et al., 2010; Martinko et al., 2018; Schwanhäusser et al., 2011), we were prompted to determine whether the most common changes we observed. BCAM and NRCAM downregulation, in particular, were also observed in patient samples harboring the same oncogenic signatures in a number of cancers types of epithelial origin (**Figure 2B-C**). Utilizing 17 TCGA provisional dataset from The Cancer Genome Atlas with carcinoma (epithelial origin) annotation (Cancer Genome Atlas Research Network et al., 2013), activating oncogenic signature was defined as G12 or Q61 mutation in KRAS, L858 or amplification of EGFR, V600E mutation in

BRAF, or amplification of HER2. Transcriptional upregulation of MME, however, was not found in any of the dataset searched, suggesting regulation at the translational level or additional factors at play.

At a global level, there were greater similarities between particular oncogenes. For example, cells harboring KRAS<sup>G12V</sup> and HER2 clustered more closely together (Cluster 1), and those containing BRAF<sup>V600E</sup>, AKT<sup>myr</sup>, EGFR<sup>L858R</sup> and MEK<sup>DD</sup> clustered together (Cluster 2) as seen either in the upset plot (**Figure 2A**) or a heatmap with hierarchical clustering (**Figure 2D**). The Pearson correlation coefficients were also higher within rather than across the two Clusters (**Figure S4A**). One of the striking findings was the upregulation of HER2 expression in the KRAS<sup>G12V</sup> transformed cells but not in Cluster 2 transformed cells (**Figure S5**), which may help to explain the stronger similarity in oncogene-induced surfaceomes between the HER2 and KRAS<sup>G12V</sup> cell lines. This same analysis also showed striking compensating regulation, where HER2 is down-regulated in the EGFR oncogene expressing cell line.

Despite detailed differences at the individual target level, these harmonized into common biological processes when viewed by Gene Set Enrichment Analysis (GSEA) (**Figure 2E**). For example, glycosylation and carbohydrate metabolism were similarly altered features for all the oncogenes (**Figure 2F**), consistent with numerous reports that altered glycosylation correlates with the development and progression of cancer (Pinho and Reis, 2015; Varki et al., 2015). In addition, we see strong down-regulation of proteins involved in differentiation and adhesion reflecting cell attachment and migration (**Supplemental**

**Fig 6A**). There were also large changes in some cell surface phosphatases involved in down-regulation of receptor tyrosine kinases (**Figure S6B**). These specific heatmaps reinforce the general division between Cluster 1 and Cluster 2. Due to the complexity of this datatype, a data browser was made to view each individual gene sets identified ([https://wellslab.ucsf.edu/oncogene\\_surfaceome](https://wellslab.ucsf.edu/oncogene_surfaceome)).

### **MEK inhibition induces common surfaceome changes**

The MAPK pathway is a central driver of cell proliferation and has been a major therapeutic target in cancer. Indeed, MEKi significantly impedes the growth of all the oncogene transformed cells either in the absence (**Figure 1D**) or presence of growth factors (**Figure S2**). To determine how MEKi alters the surfaceome of the oncogene transformed cells, we compared the proteomic landscape of each oncogenic cell line in the presence and absence of MEKi. The proteomics dataset identified and quantified a total of 772 proteins, including an intersection of 492 proteins with the empty vector dataset discussed above, for a total of 934 proteins quantified between the two experiments. (**Figure 3A-C**).

In large measure MEKi reversed the effects of the oncogenes and remarkably induced a more common state between all the cell lines including the empty vector control. For example, in contrast to the uninhibited datasets where 43% of proteins were differentially regulated by more than two-fold in the oncogene expressing cells, only 17% of the proteins (129 of 772) were altered by more than two-fold in the presence of MEKi (**Figure 3A**). Of the proteins detected, 18 proteins were commonly changed across all oncogenes, as opposed to three in the absence of MEK inhibition (**Figure 3A, Figure S3G-M**). The

similarity can be seen by hierarchical clustering of the MEKi datasets showing common changes across all cell lines (**Figure 3B**). The Pearson correlation between the six oncogene and untransformed control were also higher in general (**Figure S4B**). MEKi in KRAS<sup>G12V</sup> and HER2 are still most closely correlated. GSEA of the MEKi data indicated a general common phenotypic reversal with down regulation of membrane transporters, metabolism, and upregulation of cell adhesion proteins consistent with a decrease in cancer associated phenotypes such as cellular proliferation and metastasis (**Figure 3C, Figure S6C-D**).

Integrating oncogenic transformation and MEK inhibition dataset, 20 protein targets symmetrically flip from being significantly up to down regulated or vice versa in at least three cell lines, suggesting that the expression of these proteins is strongly dependent on the MAPK signaling pathway (**Figure 3D**). One such target, PODXL, appears to be stringently regulated by the MAPK pathway. Utilizing the same informatics approaches described above, transcription of PODXL was also found to be strongly upregulated in several cancer types of epithelial origin (**Figure 3E**). Additionally, 75 targets that were markedly up or down regulated revert to a median level of expression upon treatment with MEKi in at least three cell lines (**Figure 3F**). Interestingly, there were a handful of oncogene-induced targets that are further up-regulated upon MEKi, such as MME, suggesting they are maintained by circuitous pathways outside of MAPK (MEKi independent). Overall, the six oncogenes cause profound changes to the surfaceome that alter common biological processes, which can be largely blunted by inhibition of the MAPK pathway.

## **Oncogenes induce large changes to the glycoproteome**

Glycosylation has long been known as a biomarker for cancer (Adamczyk et al., 2012; Mechref et al., 2012; Peracaula et al., 2008; Pinho and Reis, 2015) and our GSEA data shows systematic up regulation of proteins involved in glycosylation, especially in Cluster 2 (**Figure 2E**). We thus sought to identify the N-glycosylation modifications on specific membrane proteins for the six oncogene-transformed cell lines to compare among themselves and the empty vector control. For maximal coverage we enriched the glycoproteomes using the lectin Concanavalin A (ConA) and hydrophilic interaction liquid chromatography (HILIC), that provide a complementary means for capturing high mannose type and complex type glycans, respectively (Totten et al., 2017; Zhang et al., 2016). We processed the N-glycoproteomes in biological triplicate utilizing LC-MS/MS coupled with activated ion electron transfer dissociation (AI-ETD) (**Figure S7**) (Riley et al., 2017). AI-ETD fragmentation combines radical-driven dissociation and vibrational activation and was recently shown to afford robust fragmentation of intact glycopeptides (Riley et al., 2019). Spectral assignments were made using the Byonic search engine (Bern et al., 2012); glycopeptides that were not identified across each of the three biological replicates for each cell line were removed. (**Figure S8**)

Combining ConA and HILIC enrichments allowed for sampling of different glycan classes and the aggregated results of intact N-glycoproteomic analyses identified a total of 2,648 unique N-glycopeptides that passed conservative filtering described in Materials and Methods (**Fig 4, S9**). Of this total, 189 were redundant glycan-glycosite pairs, i.e.,

glycoforms that resulted from miscleaved or differentially oxidized peptides leaving a total of 2,459 unique glycoforms (**Figure 4A**). These glycoforms map to 785 N-glycosites found on a total of 480 glycoproteins. Thus, on average each glycoprotein had 1.6 N-linked glycosites and each site had on average 3.1 different glycans, reflecting significant heterogeneity in glycosylation. Similar, yet unique glycoforms indicated that the heterogeneity of protein N-glycosylation is driven largely by the biology and not artifacts of in-source fragmentation. Of the 785 glycosites, 324 are either known or predicted to be glycosylated in UniProt (UniProt Consortium, 2010), and over half are newly reported here.

In total we identified 142 different glycan structures. The glycans can be categorized into six structural classes based on their maturation state as they transition from ER to Golgi and then split off to either lysosomal, granular, or cell surface destinations (**Figure S10**)(Yarema and Bertozzi, 2001). The six glycan categories represent a gradation of maturation from the least mature high mannose, to paucimannose, phosphomannose, complex/hybrid, and finally fucosylated and sialylated (**Figure 4B**). We identified 1,636 mannose glycopeptides containing any of 8 high mannose-type glycans and 210 were found trimmed to contain any of 9 different paucimannose structures. The remaining 767 N-glycopeptides had more complex glycans including: 210 complex/hybrid glycopeptides with 28 different structures, 314 of the sialylated class with 54 different structures, and 226 of the fucosylated class with 42 different structures. We found only 17 phosphomannose containing glycopeptides each having the same structural type that are on proteins typically found in lysosomes (Kornfeld and Mellman, 1989; Rohn et al., 2000; Rouillé et al., 2000).

Thus, over 60% of the N-glycopeptides identified contained less mature high mannose or paucimannose structures originating from processing in the ER or early Golgi. It is not surprising to find these high mannose modifications associated with cell surface proteins as others have seen that high mannose N-glycosylation is abundant at the cell surface and especially associated with oncogene transformation (An et al., 2012; Balog et al., 2012; Holst et al., 2016).

We found significant heterogeneity in the number of different glycan structures at any given site (**Figure 4C**). Approximately 45% of the sites were observed with only one glycan structure, but the range of glycans on these sites was broad. Some glycosites, such as position N-234 on aminopeptidase N (ANPEP) had up to 39 different glycans detected. Additionally, the number of glycosites per protein varies considerably; about 70% of the glycoproteins have only one site of N-glycosylation but about 10% have over five sites. There is a general trend between the number of glycosites identified on a given protein and the number of unique glycans it has; however, some proteins show significantly increased heterogeneity of glycans relative to their number of total glycosites (**Figure 4D**). The most extreme glycan diversity was observed on ANPEP, on which 59 unique glycans were identified across 8 glycosites (**Figure 4E**). Interestingly, the glycosites with higher glycan diversity mapped to specific regions on the 3D structure of ANPEP. One face of the protein containing a cluster of four asparagines (at positions 128, 234, 265 and 818) had high glycan diversity ranging from 19 to 39 different glycan structures per glycosite, while the opposite face containing three clustered asparagines (at positions 625, 681 and 735) had

much lower diversity, ranging between 1 and 8 different glycan structures per site. This suggests topological bias of glycosylation on the folded protein structure.

We next applied quantitative analysis to compare the different oncogene transformed cell lines in terms of their glycoproteome landscapes and changes between them and the empty vector control. We observed hundreds of glycopeptides that were significantly and differentially expressed ( $q < 0.05$ ) upon oncogenic transformation relative to the empty vector control (**Figure S11**). In general, the volcano plots were symmetric reflecting bi-directional changes and the fold-changes ranged from 50-fold down to 50-fold upregulated representing significant remodeling of the glycoproteome. Changes in glycosylation were greater than changes in surface protein expression which range from 2-8 fold (**Figure S3**). We quantified about 600 glycopeptides in each data set, of which two-thirds are shared in all data sets (**Figure 5A**). There was high quantitative reproducibility of the glycopeptide measurements based on the close hierarchical clustering of the three biological replicates (**Figure 5B**), tight clustering by principal component analysis, and low (<30%) median CV from the empty vector control cell line (**Figure S12A-B**). MEK<sup>DD</sup> and BRAF<sup>V600E</sup> had the greatest glycoproteome similarity, while empty vector was the furthest removed from all the oncogenes.

The UpSet plot in **Figure 5c** displays significant glycopeptide differential expression that is shared and unique to each cell line. MCF10A transformed with the KRAS<sup>G12V</sup> oncogene resulted in the largest set of uniquely changing glycopeptides. 154 of the 234 differentially expressed glycopeptides in the KRAS<sup>G12V</sup> cell line were unique to KRAS<sup>G12V</sup>



transformation. Some of these were highly protein specific. For example, 28 of the 154 glycopeptides uniquely differentially expressed by KRAS<sup>G12V</sup> were identified from ANPEP and all were upregulated upon oncogenic transformation as was the protein itself (**Figure S13**). Other glycopeptides from ANPEP were differentially expressed in different sets of cell lines. In fact, 51 of 69 glycopeptides from ANPEP were differentially expressed in at least one cell line; 12 ANPEP glycopeptides were significantly upregulated in HER2 and 5 were shared between HER2 and KRAS<sup>G12V</sup>.

BRAF<sup>V600E</sup>, EGFR<sup>L858R</sup>, and MEK<sup>DD</sup> shared the most overlap of significantly changing glycopeptides between any group of three cell lines (**Figure 5C**); 24 glycopeptides belonged to this intersection. 12 were overexpressed and 12 were under expressed upon oncogenic transformation. Laminin subunit alpha-3 (LAMA3) and N-acetylglucosamine-6-sulfatase (GNS) were most represented in the over expressed and under expressed group, respectively. Four glycopeptides from each protein were identified and significantly differentially expressed in these three cell lines. Six glycopeptides were significantly differentially expressed across all 6 cell lines, 3 of which were identified from galectin-3 binding protein (Gal-3BP) and all were downregulated glycopeptides modified with high mannose glycans (**Figure S14**).

We explored the glycan composition of differentially expressed glycopeptides to capture a broader view of differential glycosylation in the oncogene transformed cell lines and to see if general trends emerged. The heatmap in **Figure 5D** displays the differential glycome composition of glycopeptides changing more than 2-fold upon oncogenic transformation

compared to empty vector control. We, again, observe greatest similarity between BRAF<sup>V600E</sup>, EGFR<sup>L858R</sup>, and MEK<sup>DD</sup> cell lines, which have an increased proportion of high mannose glycans in upregulated glycopeptides. In contrast, HER2 and KRAS<sup>G12V</sup> expressed fewer upregulated high mannose-modified glycopeptides and showed an increased proportion of complex/hybrid type glycopeptides. Further inspection revealed that nearly all the upregulated glycopeptides with a complex/hybrid glycan from the cell lines harboring HER2 (12 of 12) and KRAS<sup>G12V</sup> (13 of 18) mapped to ANPEP. This protein was also upregulated on the KRAS<sup>G12V</sup> surfaceome (Martinko et al., 2018), displayed the highest degree of glycan heterogeneity within the glycoproteomic data, and has previously been implicated in tumorigenesis (Dong et al., 2000; Pasqualini et al., 2000).

## **Discussion:**

Oncogenesis is a complex phenomenon that involves aberrant changes in multiple biological processes to promote cancer cell survival (Hanahan and Weinberg, 2011). Here we study how the surfaceome remodels in a simplified cell autologous model by six prevalent and neighboring oncogenes that drive proliferation through the MAPK signaling node. Genetic studies have shown that these oncogenes typically exhibit mutual exclusivity in tumors from cancer patients (Cancer Genome Atlas Research Network et al., 2014). The surfaceome is a terminal manifestation of these signaling pathways. We find significant differences in detailed expression patterns consistent with differences in feedback loops and collateral signaling pathways between these oncogenes (Vaseva et al., 2018). However, we find these oncogene-induced surfaceome differences harmonize in similar functional outcomes overall, and consistent with observed mutual exclusivity.

We took a reductionist approach starting from an immortalized epithelial cell line stably transformed with each of the six different oncogenes to determine how the surfaceome is remodeled in a cell autologous fashion. This system is clearly an approximation and does not recapitulate the complexity of tumors that vary in cellular context, genetic variation, heterogeneity, oncogene expression, in the presence of host immune system and metabolism. Nonetheless, cell culture models are a practical reality allowing isogenic comparison between oncogenes, renewable access to materials that permits studies to be readily reproduced (Domcke et al., 2013; Rockwell, 1980; Wilding and Bodmer, 2014). We chose the spontaneously immortalized breast epithelial cell line, MCF10A, where the functional loss of p16<sup>INK4a</sup> allow cells to be cultured with a myriad of oncogenes without oncogene-induced senescence. We picked this over cancer cell lines which are already transformed or other artificially immortalized cell lines because they often contain genetic lesions that may cause genomic instability, leading to a more idiosyncratic background. Overall, we find that independent expression of each of these six oncogenes induced profound changes in the surfaceome, both in the proteins expressed and the glycans that decorate them.

### **Common phenotypes and biological themes induced by the six oncogenes**

Although the six oncogenes we studied (HER2 overexpression, EGFR<sup>L858R</sup>, KRAS<sup>G12V</sup>, BRAF<sup>V600E</sup>, MEK<sup>DD</sup>, and AKT<sup>myr</sup>) have many variants that could have different phenotypes, we chose well known representatives to begin to understand the similarities

and differences between them at a course-grain level. Transformation with each of these oncogenes led to rapid growth to varying degrees in the absence of growth factors and produced somewhat different cell morphologies characteristic of a cancer cell transformed phenotype. Each oncogene caused large changes in the surfaceome where about 40% of the detected N-glyco proteins had altered expression level by more than two-fold evenly split between up and down-regulated proteins reflecting bi-directional cell surface remodeling.

There were important differences between the oncogenes, and they clustered into two groups based on growth rates, surfaceomes and associated glycans. Cluster 1 containing HER2 overexpression and KRAS<sup>G12V</sup> was most aggressive in proliferation and reduced adhesion. Cluster 2 was less aggressive included EGFR<sup>L858R</sup>, BRAF<sup>V600E</sup>, MEK<sup>DD</sup> and AKT<sup>myr</sup>. We believe much of the similarity between HER2 and KRAS derives from KRAS induced expression of HER2, whereas EGFR suppresses expression of HER2 probably reflecting their signaling redundancy. MCF10A is derived from normal breast epithelial cells and among these oncogenes HER2 overexpression is most commonly seen in breast cancer patients at 13% compared to 2% for KRAS, 2.8 % for EGFR, 5% for AKT1, 1.7% for BRAF, and 0.6% for MEK1 (**Figure S15**). Each of these also shows remarkable mutual exclusivity relative to the others in breast cancer patients reflecting oncogene functional redundancies. Further *in silico* comparison of oncogenic mutational occurrence across all cancer types (Cerami et al., 2012; Gao et al., 2013) indicate a strong mutual exclusivity between BRAF and KRAS, and a slight mutual exclusivity with EGFR (**Figure S16**).

**Alterations in transporter expression:** One of the most pronounced oncogene-induced changes we observed is to proteins involved in solute transport that reverse upon MEK inhibition including: upregulation of SLC2A1, SLC6A15, SLC7A1, SLCO4A1, MF12/melanotransferrin, and down regulation of SLC22A5 and STEAP4 (**Figure 3D**). SLC6A15 acts as a preferential methionine amino acid transporter and shows dramatic upregulation in both KRAS and HER2 transformed cells. This is consistent with recent studies showing KRAS transformed cells exhibit an extreme sensitivity to methionine deprivation (De Sanctis et al., 2016). The Warburg Effect preferentially shifts cancer cells towards glycolysis, thereby promoting accelerated growth and division (Liberti and Locasale, 2016). We find SLC2A1, also known as glucose transporter-1 (GLUT1), is significantly upregulated in all oncogene transformed cells which would facilitate anaerobic glycolysis and increased growth for these cells. Our data show upregulation of SLCO4A1, a part of the organic anion transporter family that assists in transporting hormones such as prostaglandins, and vasopressin. Heightened expression of this hormone transporter has been previously seen in metastatic colorectal cancer (Buxhofer-Ausch et al., 2013). SLC7A1 has been shown to be upregulated across many cancer types, including colorectal cancer (Lu et al., 2013). Others have shown that SLC7A1 mRNA expression level decreases upon MAPK pathway inhibition (Calder et al., 2011), which is consistent with our proteomics data.

**General downregulation of receptor tyrosine phosphatases and tumor suppressors:**

Another important theme found within our data is a general down-regulation of surface proteins involved in receptor tyrosine phosphatases and other tumor suppressors such as:

PTPRF, PTPRS, UNC5B, and BCAM (**Figure 2A, Figure 3D**). The expression of PTPRF and PTPRS, in particular, have been associated with decreased metastasis through the inactivation of EGFR signaling, emphasizing their importance as tumor suppressor proteins in numerous cancer contexts (Davis et al., 2018; Tian et al., 2018). Both of these proteins show a marked decrease in expression across the majority of the oncogene transformed cells, perhaps as a means to promote growth and metastasis. Similarly, UNC5B has recently been shown to halt tumor progression in an *in vivo* model of bladder cancer through inducing cell cycle arrest in the G<sub>2</sub>/M phase (Kong et al., 2016). Lastly, BCAM has been previously shown to be a tumor suppressor in a model of hepatocellular carcinoma (Akiyama et al., 2013), and decreased expression in all oncogene transformed cells may signify the removal of tumor suppressive signaling.

**Alteration in adhesion molecules:** Metastasis is the leading cause of death for patients with cancer, and tumor cells acquire the ability to penetrate the surrounding tissues, thus leading to invasion (Martin et al., 2013). Many of these acquired functionalities are through changes in adhesion molecules on the cell surface, which play the role of mediating cell-cell interactions. Within our dataset, we identified five targets, LAMC2, LAMA3, LAMB3, PODXL, and MME, that are known to play important roles in mediating metastasis in various tumor types and we find to be upregulated across the majority of oncogenic mutants (**Figure 2A, Figure 3D**). LAMC2 has previously been shown to promote EMT and invasion in an *in vivo* model of lung adenocarcinoma (Moon et al., 2015). It has also been shown that colorectal cancer with high MAPK activity express heightened levels of LAMC2, supporting our proteomic results that show reversal upon MEK inhibition (Blaj

et al., 2017). PODXL had increased expression in all six mutant cell lines, and has been implicated in increasing the aggressiveness of breast and prostate cancer through the induction of both MAPK and PI3K signaling (Sizemore et al., 2007). PODXL is also a glycoprotein with extensive mucin-type O-glycosylation, the overexpression of which can lead to increased metastatic potential through increased cell cycle progression via the PI3K and MAPK axes (Nielsen and McNagny, 2009; Paszek et al., 2014; Woods et al., 2017). MME, known as matrix metalloprotease-12, was found upregulated in all six oncogenic mutants compared to the empty vector control. MME has been reported to participate in the break-down of extracellular matrix. It is involved in tumor invasion in some settings (Kerkelä et al., 2002) and can be used to promote oncolytic viral infection in tumors (Lavilla-Alonso et al., 2012). Interestingly, MME is upregulated across all oncogenes in both the oncogene and MEKi datasets, suggesting that MME expression is not regulated through the MAPK or PI3K axis. This heightened expression of MME in all oncogenic contexts presents an opportunity for future research into synthetic lethality studies with MME inhibition in combination with MEKi or other MAPK targets.

**Immune Modulation:** Another prominent cause of cancer progression is the evasion of immune surveillance. This can be achieved by overexpressing proteins that have a net immunosuppressive effect or by downregulating proteins that increase immune activation. Here, we identified three differentially regulated proteins, NT5E, HLA-F, and DSE, that play important roles in immune functions (**Figure 3D**). NT5E, which was upregulated in all MAPK mutants, promotes immunosuppression through the production of adenosine from AMP, which can decrease the capacity of natural killer cells to produce immune

activating IFN molecules and prevents the clonal expansion of cytotoxic T cells in the surrounding tumor tissue (Gao et al., 2014). Conversely, HLA-F and DSE (also known as dermatan sulfate epimerase), were downregulated in all oncogene expressing cell lines, and both play a role in immune system activation and tumor-rejection, through activation of NK cells or cytotoxic T cells, respectively, in the tumor microenvironment (Dulberger et al., 2017; Liao et al., 2019). Both NT5E and HLA-F show reversed expression levels upon MEKi, suggesting a potential mechanism for immune re-activation after tumor development (**Figure 3D**).

**Down-regulation by shedding:** Proteolytic cleavage of proteins at the cell surface leads to ectodomain shedding and the existence of neo N-termini. The loss of ectodomains would lead to decreased peptide detection by LC-MS/MS. We identified four surface proteins reduced in detection, MSLN, TACSTD2, NRCAM, and GPC1, that are known to undergo proteolysis to enhance the growth and metastasis of cancer. (Kawahara et al., 2017; Morello et al., 2016; Stoyanova et al., 2012). In particular, NRCAM, which had lower surface levels in all six oncogenic mutants is a cell-cell adhesion molecule; it is known to be shed by proteolysis, and the shed form stimulates cell proliferation via the AKT pathway (Conacci-Sorrell et al., 2005). Perhaps unsurprisingly, NRCAM did not exhibit a renormalization of expression levels after MEK inhibition, due to its exclusive function through the AKT pathway.

**Changes in glycosylation machinery and protein N-glycosylation:** Alterations in glycosylation are common features of cancer cells (Pinho and Reis, 2015) which can result



from a variety of factors, including changes in expression of glycosyltransferases, availability of sugar nucleotide substrates, alteration in expression of substrate proteins, or changes to the tertiary structure of the protein substrate that disrupt transfer of the glycan. We find glycosyl transferases FUT10, EXT2, GALNT11, GCTN2, and ST6GALNAC2 are highly upregulated across all oncogenic cell lines. FUT10 is an  $\alpha$ -1,3-fucosyltransferase, and FUTs are critical for the production of Lewis<sup>X</sup> and Sialyl Lewis<sup>X</sup> antigens which are hallmarks of invasion (Blanas et al., 2018). In fact, the most aggressive of our oncogene cell lines (HER2 and KRAS) are significantly upregulated in complex hybrid type glycans for which these are members, and especially noteworthy for ANPEP. EXT2 forms a heterodimer with EXT1 in the Golgi functioning as a glycosyl transferase involved in catalyzing the formation of heparin, a substrate that contributes to structural stability of the extracellular matrix, thereby mediating processes such as adhesion, immune infiltration, and signaling (Nagarajan et al., 2018). Unsurprisingly, alterations in heparan sulfate formation has been shown to have emerging roles in oncogenesis, of which EXT2 is a primary facilitator (Knelson et al., 2014). GALNT11 is a protein involved in the initiation of mucin-type O-glycosylation a well known marker that over-expressed in cancer (Villacis et al., 2016) ((Libisch et al., 2014) (Hussain et al., 2016). MUC-1 is also a very important drug target for immunotherapy (Bafna et al., 2010; Posey et al., 2016; Rivalland et al., 2015; Vassilaros et al., 2013). Interestingly, a common substrate for this family of enzymes, PODXL, is also upregulated across all oncogenic cell lines (see above), suggesting a possible mechanism by which increased GALNT11 promotes pro-growth phenotypes in these cell lines. GCTN2 is a glycosyltransferase (Bierhuizen et al., 1993),

and specifically implicated in the transition from naïve to germinal center B cell (Meyer et al., 2018).

ST6GalNAc2 is a sialyltransferase that catalyzes attachment of GalNAc to core glycans and is paradoxically up-regulated in all the oncogene cell lines. A genome wide shRNA screen in mice for genes that suppress metastasis identified and validated ST6GalNAc2 as the strongest metastasis suppressor. ST6GalNAc2 is strongly down regulated in estrogen-receptor negative breast cancer tumors. Silencing of ST6GalNAc2 promotes binding of Galectin-3 and retention of tumor cells to sites of metastasis. Galectin-3 is a member of S-type lectins and known modulator of tumor progression (Liu and Rabinovich, 2005) by enhancing aggregation of tumor cells during metastasis and preventing anoikis (Bair et al., 2006; Grassadonia et al., 2002; Lin et al., 2015). Galectin-3 binds to core glycans branched glycan structures with terminal fucosylation on Galectin-3 binding protein (Gal-3BP) that ST6GalNAc blocks via sialylation (Lin et al., 2015). We found all glycopeptides from the eight glycosites Gal-3BP were down-regulated up to 50-fold in all cell lines (**Figure S14**).

Our glycoproteomic analyses show protein N-glycosylation is dramatically changed upon oncogenic transformation, and that distinct genetic drivers of oncogenesis promote unique changes to the glycoproteome. The majority of differentially expressed glycopeptides were unique to individual cell lines. We observed greatest similarity between BRAF<sup>V600E</sup>, EGFR<sup>L858R</sup>, and MEK<sup>DD</sup>, which has been consistent throughout different analyses of these cell lines (**Figure 5B**). There were significant differences between the oncogene Cluster 1 (KRAS/HER2) and Cluster 2 (EGFR/ BRAF/ MEK/AKT). Specifically, we found

differential increase in complex hybrid glycans and decrease in high mannose in Cluster 1 and just the opposite for Cluster 2.

There is abundant evidence to suggest that high-mannose type glycans are present at the cell surface, and furthermore, that cancerous cells display increased abundance of high-mannose glycans at their cell surface (Holst et al., 2016; Hua et al., 2014). A similar observation was reported in the glycomic comparison of transformed versus human embryonic stem cells (hESC), where high-mannose glycans were observed at significantly higher abundance on plasma membranes of hESCs (An et al., 2012). Together, these observations help to explain the similarities between cancerous cells and hESCs, such as the regulation of tumor suppressor genes and the ability to self-renew (Shackleton, 2010). This is consistent with our studies although we cannot exclusively rule out that some of the high mannose glycans we observe are coming from proteins in route to the membrane.

Our data suggest glycosylation is quite heterogeneous possibly reflecting incomplete maturation. The most stunning example in our data of heterogeneity in glycosylation was for ANPEP, an amino peptidase also upregulated at the protein level across most cell lines. The most extreme example was site N234 of ANPEP where we found 29 different glycan structures. When these are arranged in order of glycan maturation, we observed a close representation of its steady state progression through the secretion pathway (**Supplementary Figure 13**). ANPEP also showed the greatest representation within complex/hybrid glycopeptides that were upregulated in HER2 and KRAS<sup>G12V</sup> cell lines (**Figure 5D**). ANPEP was most upregulated in KRAS<sup>G12V</sup> in surfaceomic analyses and its

expression was not dramatically influenced by MEK inhibition. ANPEP expression in our surfaceomic and glycoproteomic data represents one way in which KRAS<sup>G12V</sup> diverges from other members of the MAPK pathway. These data, combined with ANPEP's previously identified role in angiogenesis(Pasqualini et al., 2000), underscore its importance in KRAS<sup>G12V</sup>-positive tumors. ANPEP glycosylation may serve as a very sensitive biomarker of glycan metabolism in cells given its highly diverse and heterogeneous patterns.

## **Conclusions:**

We provide a large-scale comparative study of how six neighboring proliferative oncogenes cause large cell autologous remodeling in the surface glycoproteome. While many of the changes are specific to given oncogenes at both the protein and glycopeptide level, we observe common biological themes suggesting functional redundancy of the specific cellular expression. This is consistent with observations that common tumors can express these different oncogenes although not together. We commonly observe upregulation of surface proteins involved in metabolite transport, glycosylation, and immune suppressors and down regulation of adhesion proteins and tumor suppressors which is consistent with increased cell growth and invasion that are well-known properties of cancer cells. These studies were deliberately conducted in isogenic cell lines to isolate the oncogene-induced changes. Even in a simplified autologous cell model we recapitulate many of the hallmarks of transformed cells driven by complex but functionally redundant changes to the surfaceome. We believe the work provides important insights into the

similarities and differences among the neighboring oncogenes and provide new opportunities to pursue antibody tools to follow in more complex tumor settings.

## **Acknowledgements:**

We wish to thank Professor Sourav Bandyopadhyay at UCSF for some of the lenti constructs and cell lines, Michael Westphall, Evgenia Shishkova, and Jean Lodge at University of Wisconsin for instrumentation and helpful discussions. JAW is grateful to funding from the National Institutes of Health grant R35GM122451 and Celgene corporation. KKL thanks Canadian Institutes of Health Research (CIHR) for post-doctoral funding. LLK is indebted to the Chemistry and Chemical Biology Training grant (T32 GM064337) at UCSF for funding. JJC wishes to thank NIH for generous funding from P41 (GM108538, The National Center for Quantitative Biology of Complex Systems). GMW was supported in part by the Biotechnology Training Program (5 T32 GM008349). NMR gratefully acknowledges support from National Institutes of Health grant K00CA21245403.

## **Materials and Methods**

All MCF10A cells were cultured according to established protocols unless otherwise stated. Surfaceome of each oncogene transformed cell line was compared to empty vector (EV) control using SILAC based quantification. For inhibition study, each oncogene transformed cell line was treated with 100nM MEK inhibitor PD032590 or DMSO control for three days and surfaceome changes were compared in each respective cell line using

SILAC based quantification. Glycoproteome of each oncogene transformed cell line was quantified using label-free quantification. Detailed materials and methods are included in SI appendix.

## Reference

Adamczyk, B., Tharmalingam, T., and Rudd, P.M. (2012). Glycans as cancer biomarkers.

*Biochim. Biophys. Acta* 1820, 1347–1353.

Akiyama, H., Iwahana, Y., Suda, M., Yoshimura, A., Kogai, H., Nagashima, A.,

Ohtsuka, H., Komiya, Y., and Tashiro, F. (2013). The FBI1/Akirin2 target gene, BCAM, acts as a suppressive oncogene. *PLoS One* 8, e78716.

Allen-Petersen, B.L., Carter, C.J., Ohm, A.M., and Reyland, M.E. (2014). Protein kinase C $\delta$  is required for ErbB2-driven mammary gland tumorigenesis and negatively correlates with prognosis in human breast cancer. *Oncogene* 33, 1306–1315.

An, H.J., Gip, P., Kim, J., Wu, S., Park, K.W., McVaugh, C.T., Schaffer, D. V, Bertozzi, C.R., and Lebrilla, C.B. (2012). Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol. Cell. Proteomics* 11, M111.010660.

Apweiler, R., Hermjakob, H., and Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* 1473, 4–8.

Bafna, S., Kaur, S., and Batra, S.K. (2010). Membrane-bound mucins: the mechanistic basis for alterations in the growth and survival of cancer cells. *Oncogene* 29, 2893–2904.

Bair, E.L., Nagle, R.B., Ulmer, T.A., Laferté, S., and Bowden, G.T. (2006). 90K/Mac-2 binding protein is expressed in prostate cancer and induces promatrilysin expression. *Prostate* 66, 283–293.

Balog, C.I.A., Stavenhagen, K., Fung, W.L.J., Koeleman, C.A., McDonnell, L.A., Verhoeven, A., Mesker, W.E., Tollenaar, R.A.E.M., Deelder, A.M., and Wührer, M.

- (2012). N-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. *Mol. Cell. Proteomics* *11*, 571–585.
- Bausch-Fluck, D., Hofmann, A., Bock, T., Frei, A.P., Cerciello, F., Jacobs, A., Moest, H., Omasits, U., Gundry, R.L., Yoon, C., et al. (2015). A Mass Spectrometric-Derived Cell Surface Protein Atlas. *PLoS One* *10*, e0121314.
- Bausch-Fluck, D., Goldmann, U., Müller, S., van Oostrum, M., Müller, M., Schubert, O.T., and Wollscheid, B. (2018). The in silico human surfaceome. *Proc. Natl. Acad. Sci.* *201808790*.
- Bern, M., Kil, Y.J., and Becker, C. (2012). Byonic: advanced peptide and protein identification software. *Curr. Protoc. Bioinforma. Chapter 13*, Unit13.20.
- Bierhuizen, M.F., Mattei, M.G., and Fukuda, M. (1993). Expression of the developmental I antigen by a cloned human cDNA encoding a member of a beta-1,6-N-acetylglucosaminyltransferase gene family. *Genes Dev.* *7*, 468–478.
- Blaj, C., Schmidt, E.M., Lamprecht, S., Hermeking, H., Jung, A., Kirchner, T., and Horst, D. (2017). Oncogenic Effects of High MAPK Activity in Colorectal Cancer Mark Progenitor Cells and Persist Irrespective of RAS Mutations. *Cancer Res.* *77*, 1763–1774.
- Blanas, A., Sahasrabudhe, N.M., Rodríguez, E., van Kooyk, Y., and van Vliet, S.J. (2018). Fucosylated Antigens in Cancer: An Alliance toward Tumor Progression, Metastasis, and Resistance to Chemotherapy. *Front. Oncol.* *8*, 39.
- Buxhofer-Ausch, V., Secky, L., Wlcek, K., Svoboda, M., Kounnis, V., Briasoulis, E., Tzakos, A.G., Jaeger, W., and Thalhammer, T. (2013). Tumor-specific expression of organic anion-transporting polypeptides: transporters as novel targets for cancer therapy. *J. Drug Deliv.* *2013*, 863539.



Calder, M.D., Watson, P.H., and Watson, A.J. (2011). Culture medium, gas atmosphere and MAPK inhibition affect regulation of RNA-binding protein targets during mouse preimplantation development. *Reproduction* *142*, 689–698.

Cancer Genome Atlas Research Network, Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* *45*, 1113–1120.

Cancer Genome Atlas Research Network, Collisson, E.A., Campbell, J.D., Brooks, A.N., Berger, A.H., Lee, W., Chmielecki, J., Beer, D.G., Cope, L., Creighton, C.J., et al. (2014). Comprehensive molecular profiling of lung adenocarcinoma. *Nature* *511*, 543–550.

Carter, P.J., and Lazar, G.A. (2018). Next generation antibody drugs: pursuit of the “high-hanging fruit”. *Nat. Rev. Drug Discov.* *17*, 197–223.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* *2*, 401–404.

Chames, P., Van Regenmortel, M., Weiss, E., and Baty, D. (2009). Therapeutic antibodies: successes, limitations and hopes for the future. *Br. J. Pharmacol.* *157*, 220–233.

Cisowski, J., Sayin, V.I., Liu, M., Karlsson, C., and Bergo, M.O. (2015). Oncogene-induced senescence underlies the mutual exclusive nature of oncogenic KRAS and BRAF. *Oncogene* 1–6.

- Conacci-Sorrell, M., Kaplan, A., Raveh, S., Gavert, N., Sakurai, T., and Ben-Ze'ev, A. (2005). The shed ectodomain of Nr-CAM stimulates cell proliferation and motility, and confers cell transformation. *Cancer Res.* *65*, 11605–11612.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* *26*, 1367–1372.
- Cox, J., Hein, M.Y., Lubner, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* *13*, 2513–2526.
- Davis, T.B., Yang, M., Schell, M.J., Wang, H., Ma, L., Pledger, W.J., and Yeatman, T.J. (2018). PTPRS Regulates Colorectal Cancer RAS Pathway Activity by Inactivating Erk and Preventing Its Nuclear Translocation. *Sci. Rep.* *8*, 9296.
- Domcke, S., Sinha, R., Levine, D.A., Sander, C., and Schultz, N. (2013). Evaluating cell lines as tumour models by comparison of genomic profiles. *Nat. Commun.* *4*, 2126.
- Dong, X., An, B., Salvucci Kierstead, L., Storkus, W.J., Amoscato, A.A., and Salter, R.D. (2000). Modification of the amino terminus of a class II epitope confers resistance to degradation by CD13 on dendritic cells and enhances presentation to T cells. *J. Immunol.* *164*, 129–135.
- Dulberger, C.L., McMurtrey, C.P., Hölzemer, A., Neu, K.E., Liu, V., Steinbach, A.M., Garcia-Beltran, W.F., Sulak, M., Jabri, B., Lynch, V.J., et al. (2017). Human Leukocyte Antigen F Presents Peptides and Regulates Immunity through Interactions with NK Cell Receptors. *Immunity* *46*, 1018-1029.e7.
- Fernandes, B., Sagman, U., Auger, M., Demetrio, M., and Dennis, J.W. (1991). Beta 1-6

branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. *Cancer Res.* *51*, 718–723.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* *6*, p11.

Gao, Z., Dong, K., and Zhang, H. (2014). The roles of CD73 in cancer. *Biomed Res. Int.* *2014*, 460654.

Gonzalez, R., Jennings, L.L., Knuth, M., Orth, A.P., Klock, H.E., Ou, W., Feuerhelm, J., Hull, M. V, Koesema, E., Wang, Y., et al. (2010). Screening the mammalian extracellular proteome for regulators of embryonic human stem cell pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 3552–3557.

Grassadonia, A., Tinari, N., Iurisci, I., Piccolo, E., Cumashi, A., Innominato, P., D'Egidio, M., Natoli, C., Piantelli, M., and Iacobelli, S. (2002). 90K (Mac-2 BP) and galectins in tumor progression and metastasis. *Glycoconj. J.* *19*, 551–556.

Hakomori, S. (1996). Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Res.* *56*, 5309–5318.

Hanahan, D., and Weinberg, R. a (2011). Hallmarks of cancer: the next generation. *Cell* *144*, 646–674.

Holst, S., Deuss, A.J.M., van Pelt, G.W., van Vliet, S.J., Garcia-Vallejo, J.J., Koeleman, C.A.M., Deelder, A.M., Mesker, W.E., Tollenaar, R.A., Rombouts, Y., et al. (2016). N-glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homebox 1 (CDX1)/Villin mRNA Expression. *Mol. Cell. Proteomics* *15*, 124–140.

Hua, S., Saunders, M., Dimapasoc, L.M., Jeong, S.H., Kim, B.J., Kim, S., So, M., Lee, K.-S., Kim, J.H., Lam, K.S., et al. (2014). Differentiation of cancer cell origin and molecular subtype by plasma membrane N-glycan profiling. *J. Proteome Res.* *13*, 961–968.

Hudak, J.E., Canham, S.M., and Bertozzi, C.R. (2014). Glycocalyx engineering reveals a Siglec-based mechanism for NK cell immunoevasion. *Nat. Chem. Biol.* *10*, 69–75.

Hussain, M.R.M., Hoessli, D.C., and Fang, M. (2016). N-acetylgalactosaminyltransferases in cancer. *Oncotarget* *7*, 54067–54081.

Kannagi, R., Sakuma, K., Miyazaki, K., Lim, K.-T., Yusa, A., Yin, J., and Izawa, M. (2010). Altered expression of glycan genes in cancers induced by epigenetic silencing and tumor hypoxia: clues in the ongoing search for new tumor markers. *Cancer Sci.* *101*, 586–593.

Kawahara, R., Granato, D.C., Yokoo, S., Domingues, R.R., Trindade, D.M., and Paes Leme, A.F. (2017). Mass spectrometry-based proteomics revealed Glypican-1 as a novel ADAM17 substrate. *J. Proteomics* *151*, 53–65.

Kerkelä, E., Ala-aho, R., Klemi, P., Grénman, S., Shapiro, S.D., Kähäri, V.-M., and Saarialho-Kere, U. (2002). Metalloelastase (MMP-12) expression by tumour cells in squamous cell carcinoma of the vulva correlates with invasiveness, while that by macrophages predicts better outcome. *J. Pathol.* *198*, 258–269.

Knelson, E.H., Nee, J.C., and Blobel, G.C. (2014). Heparan sulfate signaling in cancer. *Trends Biochem. Sci.* *39*, 277–288.

Kong, C., Zhan, B., Piao, C., Zhang, Z., Zhu, Y., and Li, Q. (2016). Overexpression of UNC5B in bladder cancer cells inhibits proliferation and reduces the volume of

transplantation tumors in nude mice. *BMC Cancer* *16*, 892.

Kornfeld, S., and Mellman, I. (1989). The biogenesis of lysosomes. *Annu. Rev. Cell Biol.* *5*, 483–525.

Lavilla-Alonso, S., Bauer, M.M.T., Abo-Ramadan, U., Ristimäki, A., Halavaara, J., Desmond, R.A., Wang, D., Escutenaire, S., Ahtiainen, L., Saksela, K., et al. (2012). Macrophage metalloelastase (MME) as adjuvant for intra-tumoral injection of oncolytic adenovirus and its influence on metastases development. *Cancer Gene Ther.* *19*, 126–134.

Leung, K.K., Nguyen, A., Shi, T., Tang, L., Ni, X., Escoubet, L., MacBeth, K.J., DiMartino, J., and Wells, J.A. (2019). Multiomics of azacitidine-treated AML cells reveals variable and convergent targets that remodel the cell-surface proteome. *Proc. Natl. Acad. Sci. U. S. A.* *116*, 695–700.

Liao, W.-C., Yen, H.-R., Liao, C.-K., Tseng, T.-J., Lan, C.-T., and Liu, C.-H. (2019). DSE regulates the malignant characters of hepatocellular carcinoma cells by modulating CCL5/CCR1 axis. *Am. J. Cancer Res.* *9*, 347–362.

Liberti, M. V, and Locasale, J.W. (2016). The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem. Sci.* *41*, 211–218.

Libisch, M.G., Casás, M., Chiribao, M., Moreno, P., Cayota, A., Osinaga, E., Oppezzo, P., and Robello, C. (2014). GALNT11 as a new molecular marker in chronic lymphocytic leukemia. *Gene* *533*, 270–279.

Lin, T.-W., Chang, H.-T., Chen, C.-H., Chen, C.-H., Lin, S.-W., Hsu, T.-L., and Wong, C.-H. (2015). Galectin-3 Binding Protein and Galectin-1 Interaction in Breast Cancer Cell Aggregation and Metastasis. *J. Am. Chem. Soc.* *137*, 9685–9693.

Liu, F.-T., and Rabinovich, G.A. (2005). Galectins as modulators of tumour progression. *Nat. Rev. Cancer* 5, 29–41.

Liu, M.-Q., Zeng, W.-F., Fang, P., Cao, W.-Q., Liu, C., Yan, G.-Q., Zhang, Y., Peng, C., Wu, J.-Q., Zhang, X.-J., et al. (2017). pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. *Nat. Commun.* 8, 438.

Lu, Y., Wang, W., Wang, J., Yang, C., Mao, H., Fu, X., Wu, Y., Cai, J., Han, J., Xu, Z., et al. (2013). Overexpression of arginine transporter CAT-1 is associated with accumulation of L-arginine and cell growth in human colorectal cancer tissue. *PLoS One* 8, e73866.

Lundberg, E., Fagerberg, L., Klevebring, D., Matic, I., Geiger, T., Cox, J., Algenäs, C., Lundeberg, J., Mann, M., and Uhlen, M. (2010). Defining the transcriptome and proteome in three functionally different human cell lines. *Mol. Syst. Biol.* 6, 450.

MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and MacCoss, M.J. (2010). Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966–968.

Martin, T.A., Ye, L., Sanders, A.J., Lane, J., and Jiang, W.G. (2013). Cancer invasion and metastasis: molecular and cellular perspective. In *Madame Curie Bioscience Database [Internet]*, (Landes Bioscience), p.

Martinko, A.J., Truillet, C., Julien, O., Diaz, J., Horlbeck, M.A., Whiteley, G., Blonder, J., Weissman, J.S., Bandyopadhyay, S., Evans, M., et al. (2018). Targeting RAS driven human cancer cells with antibodies to upregulated and essential cell-surface proteins.

Elife 7, e31098.

Martins, M.M., Zhou, A.Y., Corella, A., Horiuchi, D., Yau, C., Rakshandehroo, T.,

Gordan, J.D., Levin, R.S., Johnson, J., Jascur, J., et al. (2015). Linking tumor mutations to drug responses via a quantitative chemical-genetic interaction map. *Cancer Discov.* 5, 154–167.

Mechref, Y., Hu, Y., Garcia, A., and Hussein, A. (2012). Identifying cancer biomarkers by mass spectrometry-based glycomics. *Electrophoresis* 33, 1755–1767.

Meyer, S.J., Linder, A.T., Brandl, C., and Nitschke, L. (2018). B Cell Siglecs-News on Signaling and Its Interplay With Ligand Binding. *Front. Immunol.* 9, 2820.

Moon, Y.W., Rao, G., Kim, J.J., Shim, H.-S., Park, K.-S., An, S.S., Kim, B., Steeg, P.S., Sarfaraz, S., Changwoo Lee, L., et al. (2015). LAMC2 enhances the metastatic potential of lung adenocarcinoma. *Cell Death Differ.* 22, 1341–1352.

Morello, A., Sadelain, M., and Adusumilli, P.S. (2016). Mesothelin-Targeted CARs: Driving T Cells to Solid Tumors. *Cancer Discov.* 6, 133–146.

Nagarajan, A., Malvi, P., and Wajapeyee, N. (2018). Heparan Sulfate and Heparan Sulfate Proteoglycans in Cancer Initiation and Progression. *Front. Endocrinol. (Lausanne)*. 9, 483.

Nielsen, J.S., and McNagny, K.M. (2009). The role of podocalyxin in health and disease. *J. Am. Soc. Nephrol.* 20, 1669–1676.

Palaniappan, K.K., and Bertozzi, C.R. (2016). Chemical Glycoproteomics. *Chem. Rev.* 116, 14277–14306.

Pasqualini, R., Koivunen, E., Kain, R., Lahdenranta, J., Sakamoto, M., Stryhn, A.,

Ashmun, R.A., Shapiro, L.H., Arap, W., and Ruoslahti, E. (2000). Aminopeptidase N is a

receptor for tumor-homing peptides and a target for inhibiting angiogenesis. *Cancer Res.* *60*, 722–727.

Paszek, M.J., DuFort, C.C., Rossier, O., Bainer, R., Mouw, J.K., Godula, K., Hudak, J.E., Lakins, J.N., Wijekoon, A.C., Cassereau, L., et al. (2014). The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature* *511*, 319–325.

Peracaula, R., Barrabés, S., Sarrats, A., Rudd, P.M., and de Llorens, R. (2008). Altered glycosylation in tumours focused to cancer diagnosis. *Dis. Markers* *25*, 207–218.

Pinho, S.S., and Reis, C.A. (2015). Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* *15*, 540–555.

Posey, A.D., Schwab, R.D., Boesteanu, A.C., Steentoft, C., Mandel, U., Engels, B., Stone, J.D., Madsen, T.D., Schreiber, K., Haines, K.M., et al. (2016). Engineered CAR T Cells Targeting the Cancer-Associated Tn-Glycoform of the Membrane Mucin MUC1 Control Adenocarcinoma. *Immunity* *44*, 1444–1454.

Qu, Y., Han, B., Yu, Y., Yao, W., Bose, S., Karlan, B.Y., Giuliano, A.E., and Cui, X. (2015). Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells. *PLoS One* *10*, e0131285.

R Core Team (2017). R: A language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria.

Riley, N.M., Westphall, M.S., Hebert, A.S., and Coon, J.J. (2017). Implementation of Activated Ion Electron Transfer Dissociation on a Quadrupole-Orbitrap-Linear Ion Trap Hybrid Mass Spectrometer. *Anal. Chem.* *89*, 6358–6366.

Riley, N.M., Hebert, A.S., Westphall, M.S., and Coon, J.J. (2019). Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis. *Nat. Commun.* *10*, 1311.



Rivalland, G., Loveland, B., and Mitchell, P. (2015). Update on Mucin-1 immunotherapy in cancer: a clinical perspective. *Expert Opin. Biol. Ther.* *15*, 1773–1787.

Rockwell, S. (1980). In vivo-in vitro tumour cell lines: characteristics and limitations as models for human cancer. *Br. J. Cancer. Suppl.* *4*, 118–122.

Rohn, W.M., Rouillé, Y., Waguri, S., and Hoflack, B. (2000). Bi-directional trafficking between the trans-Golgi network and the endosomal/lysosomal system. *J. Cell Sci.* *113* (Pt 1), 2093–2101.

Rouillé, Y., Rohn, W., and Hoflack, B. (2000). Targeting of lysosomal proteins. *Semin. Cell Dev. Biol.* *11*, 165–171.

De Sanctis, G., Spinelli, M., Vanoni, M., and Sacco, E. (2016). K-Ras Activation Induces Differential Sensitivity to Sulfur Amino Acid Limitation and Deprivation and to Oxidative and Anti-Oxidative Stress in Mouse Fibroblasts. *PLoS One* *11*, e0163790.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* *473*, 337–342.

Sergushichev, A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *BioRxiv* 060012.

Shackleton, M. (2010). Normal stem cells and cancer stem cells: similar and different. *Semin. Cancer Biol.* *20*, 85–92.

Sizemore, S., Cicek, M., Sizemore, N., Ng, K.P., and Casey, G. (2007). Podocalyxin increases the aggressive phenotype of breast and prostate cancer cells in vitro through its interaction with ezrin. *Cancer Res.* *67*, 6183–6191.

Soule, H.D., Maloney, T.M., Wolman, S.R., Peterson, W.D., Brenz, R., McGrath, C.M.,

- Russo, J., Pauley, R.J., Jones, R.F., and Brooks, S.C. (1990). Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res.* *50*, 6075–6086.
- Stoyanova, T., Goldstein, A.S., Cai, H., Drake, J.M., Huang, J., and Witte, O.N. (2012). Regulated proteolysis of Trop2 drives epithelial hyperplasia and stem cell self-renewal via  $\beta$ -catenin signaling. *Genes Dev.* *26*, 2271–2285.
- Tian, X., Yang, C., Yang, L., Sun, Q., and Liu, N. (2018). PTPRF as a novel tumor suppressor through deactivation of ERK1/2 signaling in gastric adenocarcinoma. *Oncotargets. Ther.* *11*, 7795–7803.
- Totten, S.M., Feasley, C.L., Bermudez, A., and Pitteri, S.J. (2017). Parallel Comparison of N-Linked Glycopeptide Enrichment Techniques Reveals Extensive Glycoproteomic Analysis of Plasma Enabled by SAX-ERLIC. *J. Proteome Res.* *16*, 1249–1260.
- UniProt Consortium (2010). The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res.* *38*, D142-8.
- Unni, A.M., Lockwood, W.W., Zejnullahu, K., Lee-Lin, S.-Q., and Varmus, H. (2015). Evidence that synthetic lethality underlies the mutual exclusivity of oncogenic KRAS and EGFR mutations in lung adenocarcinoma. *Elife* *4*, 1–23.
- Varki, A., and Gagneux, P. (2012). Multifarious roles of sialic acids in immunity. *Ann. N. Y. Acad. Sci.* *1253*, 16–36.
- Varki, A., Kannagi, R., Toole, B., and Stanley, P. (2015). Glycosylation Changes in Cancer.
- Vaseva, A. V, Blake, D.R., Gilbert, T.S.K., Ng, S., Hostetter, G., Azam, S.H., Ozkan-Dagliyan, I., Gautam, P., Bryant, K.L., Pearce, K.H., et al. (2018). KRAS Suppression-

Induced Degradation of MYC Is Antagonized by a MEK5-ERK5 Compensatory Mechanism. *Cancer Cell* 34, 807-822.e7.

Vassilaros, S., Tsibanis, A., Tsikkinis, A., Pietersz, G.A., McKenzie, I.F.C., and Apostolopoulos, V. (2013). Up to 15-year clinical follow-up of a pilot Phase III immunotherapy study in stage II breast cancer patients using oxidized mannan-MUC1. *Immunotherapy* 5, 1177–1182.

Villacis, R.A.R., Miranda, P.M., Gomy, I., Santos, E.M.M., Carraro, D.M., Achatz, M.I., Rossi, B.M., and Rogatto, S.R. (2016). Contribution of rare germline copy number variations and common susceptibility loci in Lynch syndrome patients negative for mutations in the mismatch repair genes. *Int. J. Cancer* 138, 1928–1935.

Wallin, E., and von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* 7, 1029–1038.

Wilding, J.L., and Bodmer, W.F. (2014). Cancer cell lines for drug discovery and development. *Cancer Res.* 74, 2377–2384.

Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., Aebersold, R., and Watts, J.D. (2009). Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol.* 27, 378–386.

Woods, E.C., Kai, F., Barnes, J.M., Pedram, K., Pickup, M.W., Hollander, M.J., Weaver, V.M., and Bertozzi, C.R. (2017). A bulky glycocalyx fosters metastasis formation by promoting G1 cell cycle progression. *Elife* 6.

Xia, L., Schrupp, D.S., and Gildersleeve, J.C. (2016). Whole-Cell Cancer Vaccines Induce Large Antibody Responses to Carbohydrates and Glycoproteins. *Cell Chem. Biol.*

23, 1515–1525.

Xiao, H., Woods, E.C., Vukojcic, P., and Bertozzi, C.R. (2016). Precision glycoalkyl editing as a strategy for cancer immunotherapy. *Proc. Natl. Acad. Sci. U. S. A.* *113*, 10304–10309.

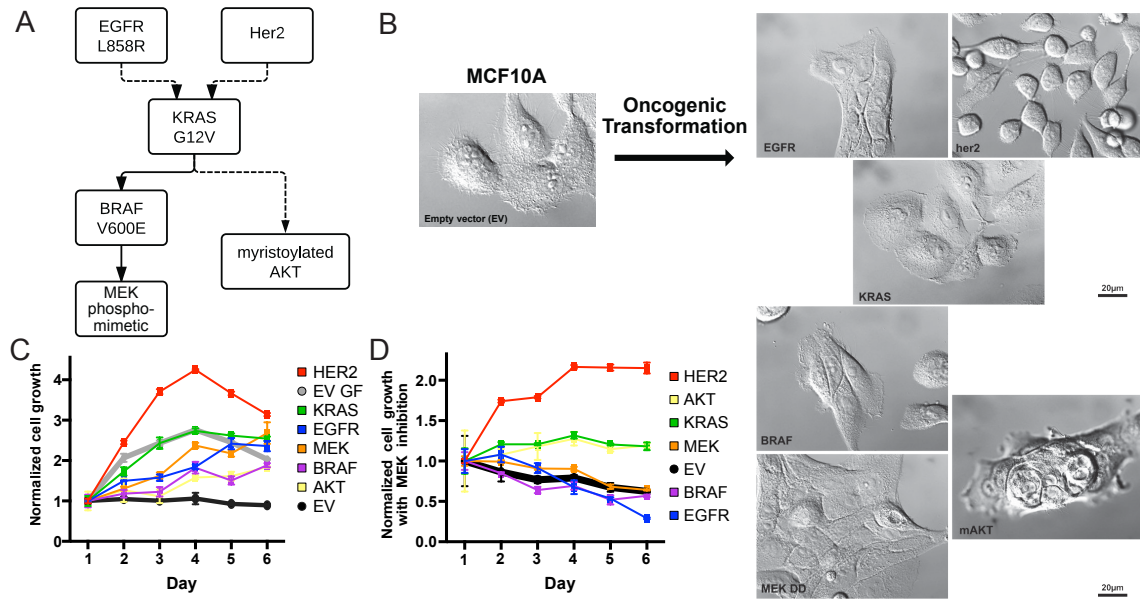
Yang, Y., Franc, V., and Heck, A.J.R. (2017). Glycoproteomics: A Balance between High-Throughput and In-Depth Analysis. *Trends Biotechnol.* *35*, 598–609.

Yarema, K.J., and Bertozzi, C.R. (2001). Characterizing glycosylation pathways. *Genome Biol.* *2*, REVIEWS0004.

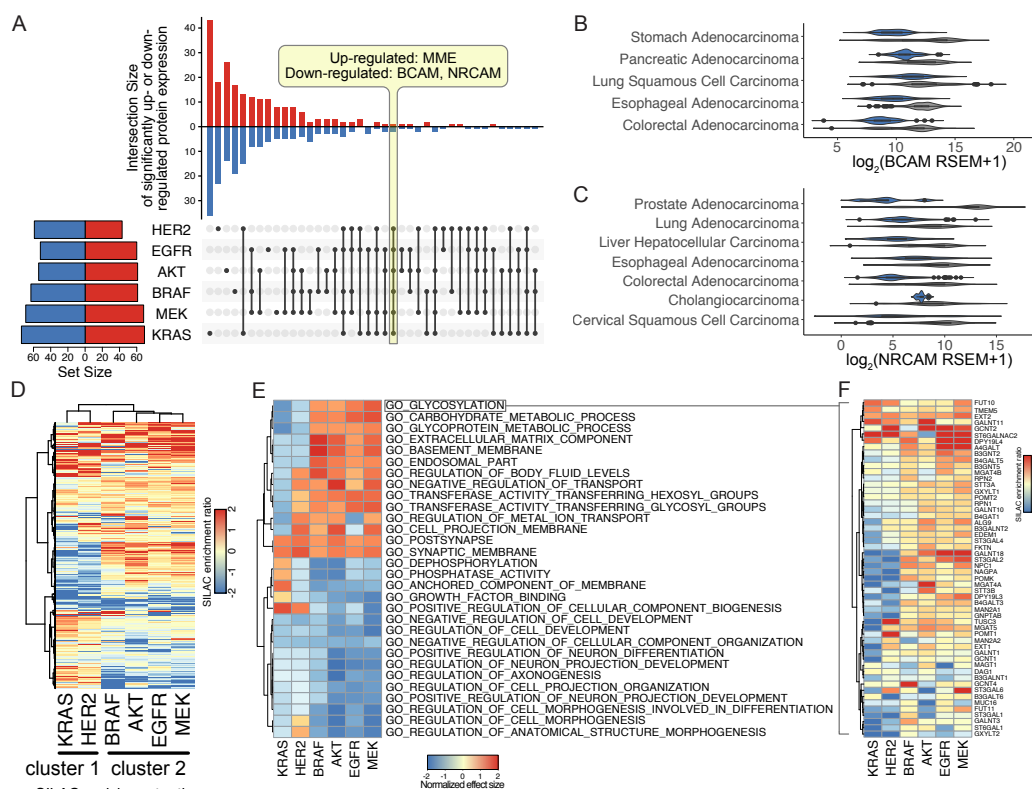
Ye, X., Chan, K.C., Waters, A.M., Bess, M., Harned, A., Wei, B., Loncarek, J., Luke, B.T., Orsburn, B.C., Hollinger, B.D., et al. (2016). Comparative proteomics of a model MCF10A-KRasG12V cell line reveals a distinct molecular signature of the KRasG12V cell surface. *Oncotarget* *7*, 86948–86971.

Ying, H., Kimmelman, A.C., Lyssiotis, C.A., Hua, S., Chu, G.C., Fletcher-Sananikone, E., Locasale, J.W., Son, J., Zhang, H., Coloff, J.L., et al. (2012). Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* *149*, 656–670.

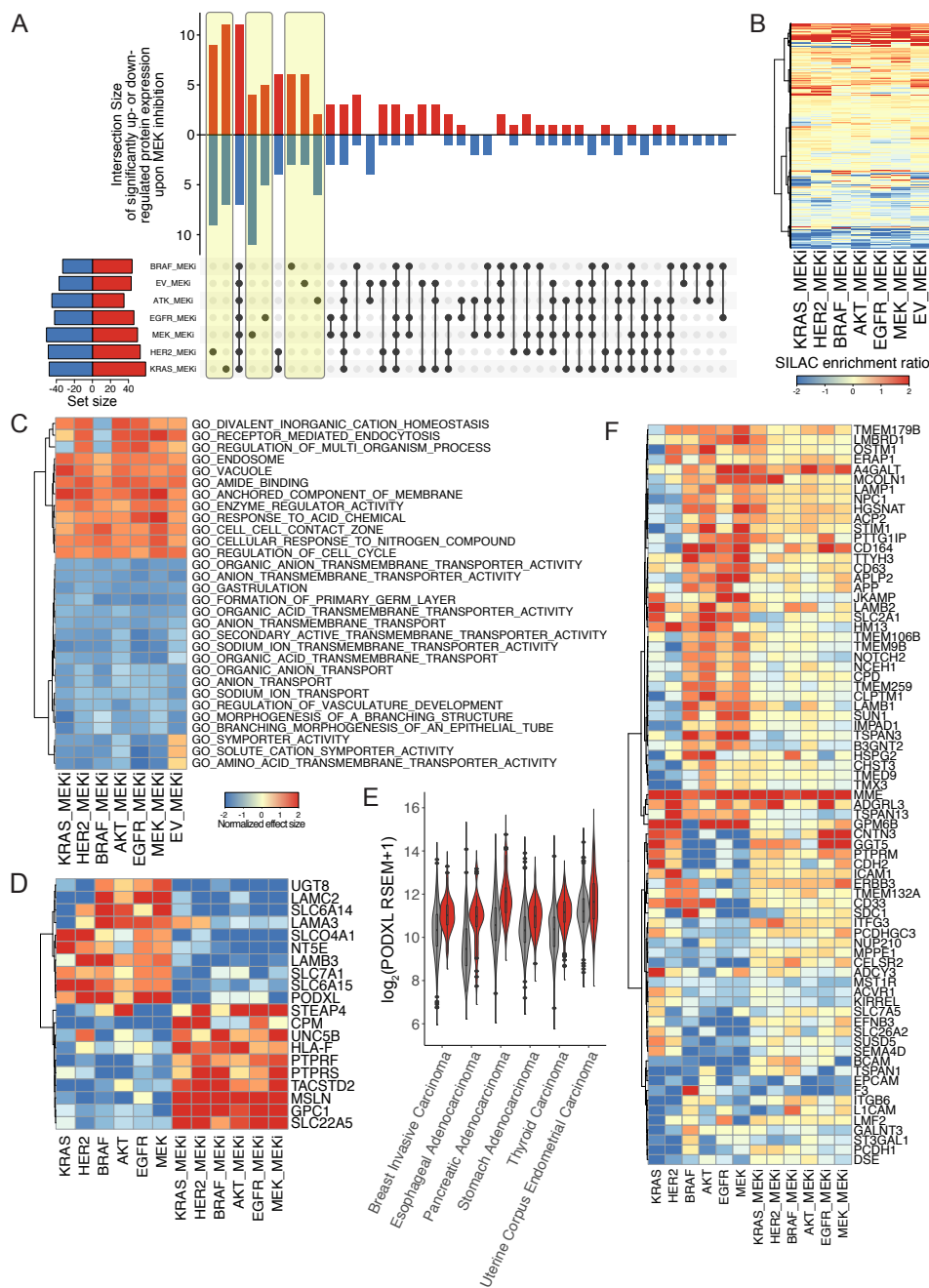
Zhang, C., Ye, Z., Xue, P., Shu, Q., Zhou, Y., Ji, Y., Fu, Y., Wang, J., and Yang, F. (2016). Evaluation of Different N-Glycopeptide Enrichment Methods for N-Glycosylation Sites Mapping in Mouse Brain. *J. Proteome Res.* *15*, 2960–2968.



**Figure 1: Growth rates and morphologies for MCF10A cells transformed with neighboring proliferative oncogenes that control the MAPK pathway.** (A) Simplified signaling schematic relationship of six proliferative oncogenes studied, EGFR<sup>L858R</sup>, HER2 over-expression, KRAS<sup>G12V</sup>, BRAF<sup>V600E</sup>, MEK<sup>DD</sup> and AKT<sup>myr</sup>. (B) Oncogenic transformation of MCF10A induces diverse cellular morphologies. Note, images are presented in the same order of the schematic in panel A. (C) MCF10A cells stably transformed with lenti-virus with the different oncogenes grow independent of growth factor. Grey and black lines indicate cellular growth of MCF10A EV control with and without growth factors, respectively. Cell growth (n = 3) was measured each day for 6 days by CellTiter-Glo luminescent cell viability assay and normalized to viability on day 1. (D) Suppression of growth for all cell lines by treatment with 100 nM MEK inhibitor (PD0325901) in the absence of growth factors. MCF10A cells transformed with HER2 appears to be least sensitive to MEKi, followed by KRAS<sup>G12V</sup> and AKT<sup>myr</sup>.



**Figure 2: Proliferative oncogenes cause large changes in the surfaceome that are diverse in detail but have common functional themes.** (A) Many differentially regulated proteins are unique to each cell line, and only three proteins were commonly up or down-regulated among all six cell lines. In the top bar graph, up-regulated proteins (red) are indicated by the upward bars, and down-regulated proteins (blue) are indicated by the downward bars. The specific overlapping groups are indicated by the black solid points below the bar graph. Total differentially regulated proteins for each cell line are indicated in the left bar graph. Up-regulated (red) and down-regulated (blue) proteins are defined as  $\log_2(\text{FC}) > 1$  and  $p \text{ value} < 0.05$ . (B-C) BCAM (B) and NRCAM (C) are significantly down regulated across various carcinoma with activating oncogenic signature in KRAS, HER2, BRAF and EGFR (blue) compared to no mutations in these genes (grey). Activating oncogenic signature was defined as G12 or Q61 mutation in KRAS, L858 or amplification of EGFR, V600E mutation in BRAF, or amplification of HER2. Genomic and expression data was obtained from 17 TCGA provisional dataset with carcinoma (epithelial origin) annotation. (D) Hierarchical clustering of surfaceome changes revealed similarities between HER2 and KRAS transformed cells. (E) Top 30 enriched gene sets identified by Gene Set Enrichment Analysis (GSEA) of the proteomics dataset using Gene Ontology terms shows clustering between HER2 and KRAS transformed cells. Positive normalized effect size (up-regulation) is shown in red and negative (down-regulated) normalized effect size is shown in blue. Proteins were pre-ranked by median SILAC peptide ratio and GSEA was performed using MySigDB C5 GO gene set collection. (F) Proteins involved in glycosylation (GO:0070085) are down regulated in the KRAS and HER2 cluster (blue) while the same proteins are up-regulated in the other oncogenes cluster. Heatmap for each Gene Set identified is also available at [wellslab.ucsf.edu/oncogene\\_surfaceome](http://wellslab.ucsf.edu/oncogene_surfaceome).

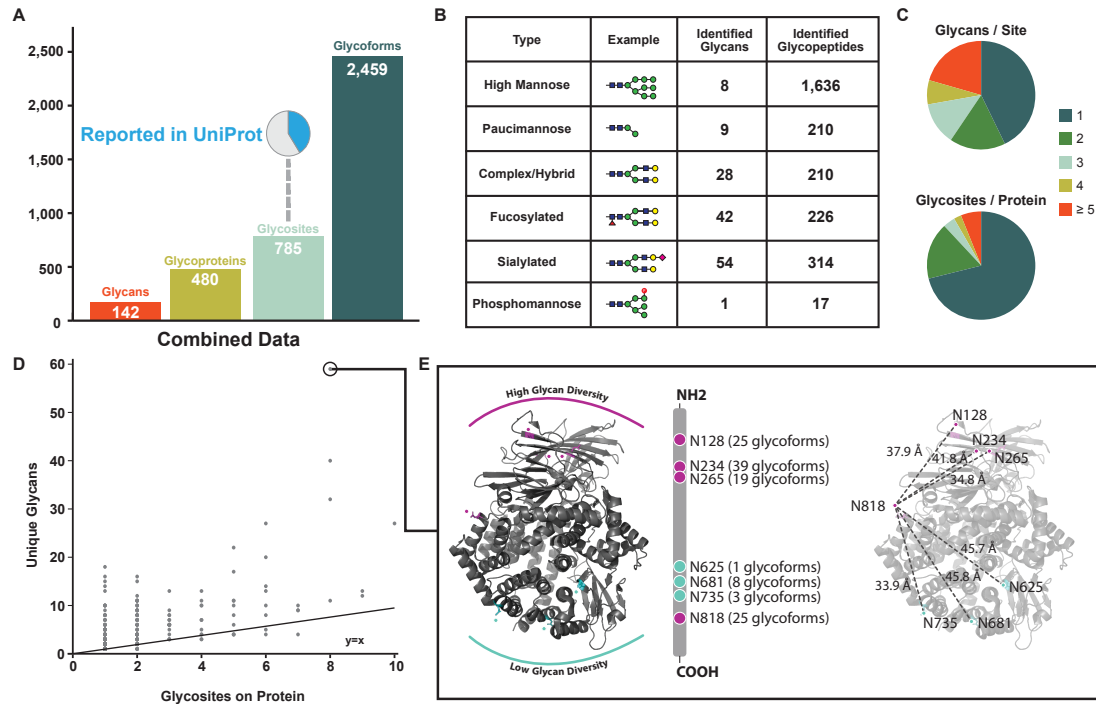


**Figure 3: Inhibition of MEK led to similar perturbations of the surfaceome regardless of oncogenic transformation.** (A) Majority of differentially regulated surface proteins were either completely unique (highlighted in yellow) or completely common (bar 3) to all seven cell lines. In the top bar graph, up-regulated genes (red) are indicated by the upward bars, and down-regulated genes (blue) are indicated by the downward bars. The specific overlapping groups are indicated by the black solid points below the bar graph. Total differentially regulated proteins for each cell line are indicated in the left

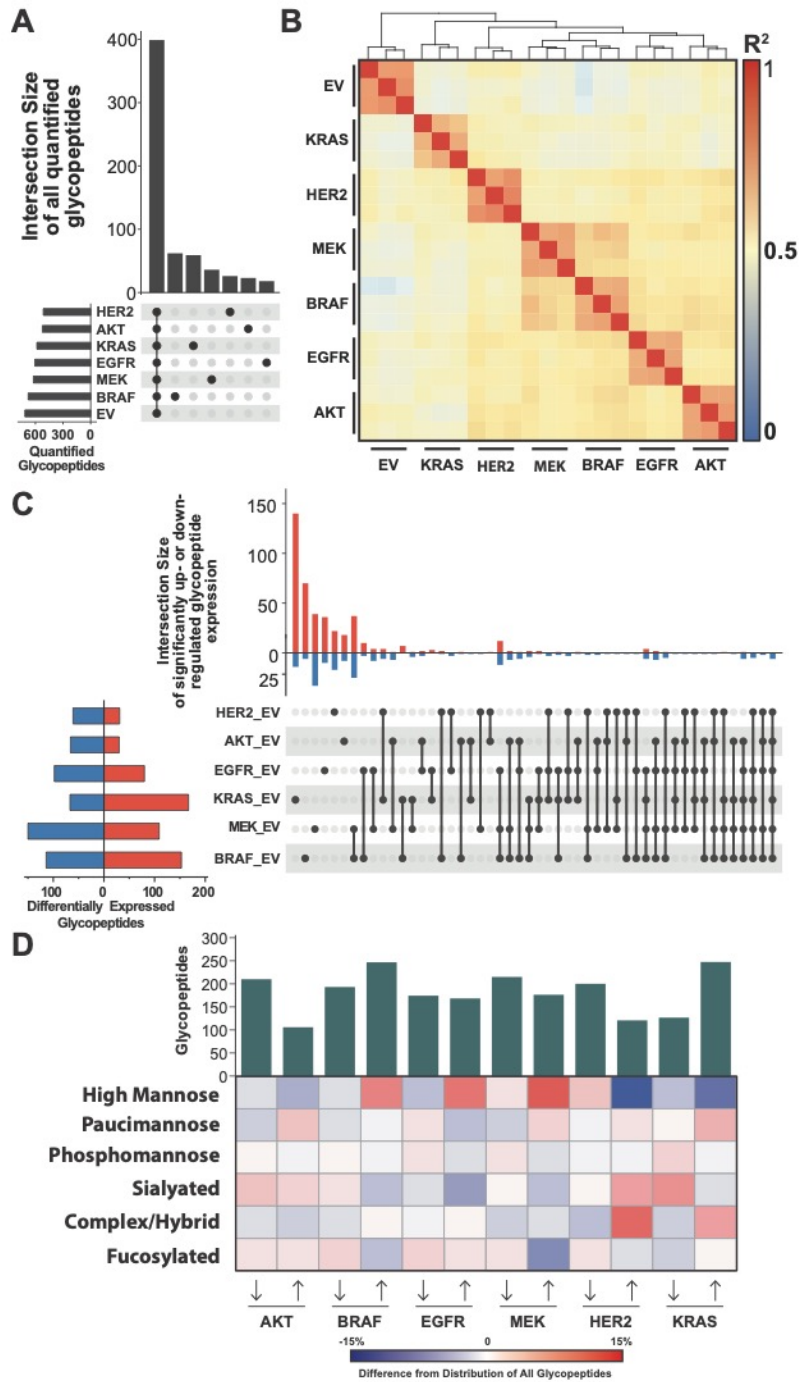


bar graph. Up-regulated (red) and down-regulated (blue) proteins are defined as  $\log_2(\text{FC}) > 1$  and  $p \text{ value} < 0.05$ . (B) Hierarchical clustering of surfaceome changes shows MEK inhibition treatment affected each cell line similarly. (C) Top 30 enriched gene sets identified by Gene Set Enrichment Analysis (GSEA) of the proteomics dataset using Gene Ontology terms show similarities between all cell lines, indicating unique changes could be functionally redundant. Positive normalized effect size (up-regulation) is shown in red and negative (down-regulated) normalized effect size is shown in blue. Proteins were pre-ranked by median SILAC ratio and GSEA was performed using MySigDB C5 GO gene set collection. Heatmap for each Gene Set identified is also available at [wellslab.ucsf.edu/oncogene\\_surfaceome](http://wellslab.ucsf.edu/oncogene_surfaceome). (D) MEK dependent regulation of protein expression induced by oncogenic transformation. Proteins shown were differentially regulated by oncogene (either  $\log_2(\text{FC}) > 1$  or  $\log_2(\text{FC}) < -1$  in oncogene vs EV) and were then reversed (flipped  $\log_2(\text{FC}) < -1$  and  $\log_2(\text{FC}) > 1$  in MEK inhibition vs respective cell line) in at least 3 cell lines. (E) PODXL is significantly upregulated across various carcinoma with activating oncogenic signature in KRAS, HER2, BRAF and EGFR (red) compared to no mutations in these genes (grey). Activating oncogenic signature was defined as G12 or Q61 mutation in KRAS, L858 or amplification of EGFR, V600E mutation in BRAF, or amplification of HER2. Genomic and expression data was obtained from 17 TCGA provisional dataset with carcinoma (epithelial origin) annotation. (F) MEK independent regulation of protein expression induced by oncogenic transformation. Proteins shown were differentially regulated by oncogene (either  $\log_2(\text{FC}) > 1$  or  $\log_2(\text{FC}) < -1$  in onco vs EV) and were then reverted to a median level but not reversed (flipped  $\log_2(\text{FC}) < -1$  and  $\log_2(\text{FC}) > 1$  in MEK inhibition vs respective cell line) in at least 3 cell lines.





**Figure 4: A global overview of the MCF10a glycoproteome identified from all cell lines.** (A) Unique glycans, glycoproteins, glycosites, and glycoforms identified across all six cell lines. A pie charts represents a high percentage of glycosites not previously reported in uniprot. (B) High mannose glycan type is the most common glycan identified. A total of six common glycan types are used to categorize glycans identified. (C) The varying degrees of microheterogeneity (top) and macroheterogeneity (bottom) of protein N-glycosylation. (D) A scatter plot of unique glycans versus number of glycosites for each protein provides an additional view of heterogeneity in protein N-glycosylation. The protein with the most heterogenous glycoform was Aminopeptidase N (ANPEP, UniProtKB: P15144). (E) Mapping glycan diversity on structure of ANPEP reveals spatial regulation of glycosylation patterns. Asparagine residues with high (magenta) and low (teal) numbers unique glycans are highlighted. The most C-terminus glycan, N818, has high glycan diversity distance and has similar distance to all other asparagine residues. (PDB ID: 4FYR)



**Figure 5: Quantitative glycopeptide measurements across mutant cell lines.** (A) An UpSet plot shows glycopeptide identifications that are unique to or shared between data sets. (B) Pairwise Pearson correlations from all replicate analyses illustrates clustering between MEK<sup>DD</sup> and BRAF<sup>V600E</sup> glycoproteome. (C) An UpSet plot displays the shared and unique glycopeptides that are differentially expressed upon oncogenic transformation. (D) A heatmap display the differences glycan type distribution for up and down regulated glycopeptides across cell lines.