

RESEARCH

Open Access



# sBioSITE enables sensitive identification of the cell surface proteome through direct enrichment of biotinylated peptides

Kishore Garapati<sup>1,2,3</sup>, Husheng Ding<sup>3</sup>, M. Cristine Charlesworth<sup>4</sup>, Yohan Kim<sup>5</sup>, Roman Zenka<sup>4</sup>, Mayank Saraswat<sup>3</sup>, Dong-Gi Mun<sup>3</sup>, Sandip Chavan<sup>3</sup>, Ashish Shingade<sup>2</sup>, Fabrice Lucien<sup>5,6</sup>, Jun Zhong<sup>3</sup>, Richard K. Kandasamy<sup>1,3,7</sup> and Akhilesh Pandey<sup>3,8\*</sup>

## Abstract

**Background** Cell surface proteins perform critical functions related to immune response, signal transduction, cell–cell interactions, and cell migration. Expression of specific cell surface proteins can determine cell-type identity, and can be altered in diseases including infections, cancer and genetic disorders. Identification of the cell surface proteome remains a challenge despite several enrichment methods exploiting their biochemical and biophysical properties.

**Methods** Here, we report a novel method for enrichment of proteins localized to cell surface. We developed this new approach designated surface Biotinylation Site Identification Technology (sBioSITE) by adapting our previously published method for direct identification of biotinylated peptides. In this strategy, the primary amine groups of lysines on proteins on the surface of live cells are first labeled with biotin, and subsequently, biotinylated peptides are enriched by anti-biotin antibodies and analyzed by liquid chromatography–tandem mass spectrometry (LC–MS/MS).

**Results** By direct detection of biotinylated lysines from PC-3, a prostate cancer cell line, using sBioSITE, we identified 5851 peptides biotinylated on the cell surface that were derived from 1409 proteins. Of these proteins, 533 were previously shown or predicted to be localized to the cell surface or secreted extracellularly. Several of the identified cell surface markers have known associations with prostate cancer and metastasis including CD59, 4F2 cell-surface antigen heavy chain (SLC3A2) and adhesion G protein-coupled receptor E5 (CD97). Importantly, we identified several biotinylated peptides derived from plectin and nucleolin, both of which are not annotated in surface proteome databases but have been shown to have aberrant surface localization in certain cancers highlighting the utility of this method.

**Conclusions** Detection of biotinylation sites on cell surface proteins using sBioSITE provides a reliable method for identifying cell surface proteins. This strategy complements existing methods for detection of cell surface expressed proteins especially in discovery-based proteomics approaches.

**Keywords** BioSITE, Surfaceome, Biotinylation

\*Correspondence:

Akhilesh Pandey

pandey.akhilesh@mayo.edu

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

The plasma membrane of eukaryotic cells is a dynamic barrier between the contents of the cell and its immediate environment and is home to a large number of proteins [1]. Surface proteins are important determinants of cell identity with roles in cell recognition, adhesion, migration, intercellular communication and signal transduction [2]. While the membrane proteome includes proteins that are localized to the plasma membrane by various mechanisms including single- and multi-pass transmembrane domains, lipid anchors and glycosylphosphatidylinositol (GPI) anchors, the surface proteome is a subset of these proteins defined by the presence of at least one amino acid exposed to the external surface of the cell [3]. The analysis of this subproteome has a wide range of applications in health and disease with a large number of biomarkers and therapeutic targets in neurodegenerative diseases, cancers, and autoimmune conditions being localized to the cell surface [4–6].

Although ~30% of genes are projected to encode plasma membrane proteins, these proteins are not detected as often in proteomic profiling studies [7, 8]. This underrepresentation has been attributed to the relatively lower abundance of these proteins as well as to several biochemical and sample preparation considerations. These include the hydrophobic nature and fewer tryptic cleavage sites in membrane transmembrane domains, abundant and complex post-translational modifications (e.g., glycosylation) and the dynamic nature of surface protein abundance and localization across cell states [1, 9, 10]. Several strategies have been used to overcome these difficulties through enrichment of membrane proteins by exploiting their unique properties. These include ultracentrifugation (based on differential density of subcellular fractions) [11], phase-separation (based on physical properties of lipids) [12], lectin- or antibody-mediated enrichment (based on affinity) [13] and chemical labeling followed by affinity purification (based on specific reactivity or extracellular chemical groups) [14]. Addition of a biotin tag to sugar chains of membrane glycoproteins or to extracellular lysine residues followed by pull-down with streptavidin analogs is often employed to study the cell surfaceome [15, 16]. Biotinylation reagents employed in such methods can be cleavable or non-cleavable, depending on the application and method of detection. While these methods successfully enrich membrane (or surface) proteins, they often suffer from contamination with cytoplasmic and organellar proteins to varying degrees. The specificity of several of the available methods is further reduced by the enrichment of intracellular interactors of transmembrane proteins. Additionally, they are often limited by the number of surface proteins that bear the reactive groups being targeted, e.g., the selective

enrichment of glycoproteins while precluding the enrichment of non-glycosylated cell surface proteins.

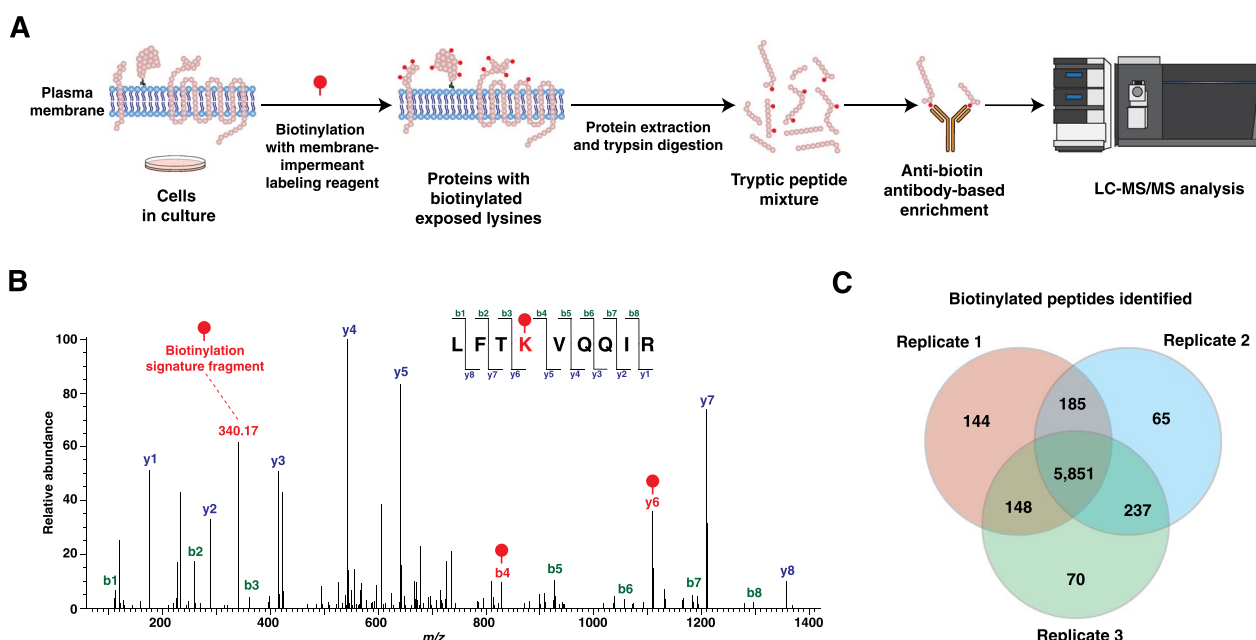
We report the development of surface Biotinylation Site Identification Technology (sBioSITE), a sensitive and reliable method for the enrichment of cell surface proteins by adapting the Biotinylation Site Identification Technology (BioSITE) for immunoprecipitation of biotinylated peptides [17]. We demonstrate that this new method results in confident identification of proteins by localization of extracellular lysine residues and signatures of the biotin tag in fragmentation mass spectra of peptides containing them. This method also enriches secreted proteins that are integral parts of the extracellular matrix or are extracellular binding partners of cell membrane molecules.

## Results

### sBioSITE-based direct identification of biotin-labeled cell surface proteins

Extracellular lysine residues of proteins expressed in PC-3, a prostate cancer cell line, were labeled using a non-cleavable, membrane-impermeant biotinylation reagent in triplicate. Following cell lysis and proteolytic digestion, biotinylated peptides were enriched using our previously published protocol which employs bead-immobilized anti-biotin antibodies [17]. Liquid chromatography–mass spectrometry (LC–MS/MS) analysis was performed to identify cell surface proteins (Fig. 1A). Lysine residues detected with biotin modification were inferred to have been present in the extracellular space during treatment. MS/MS fragment signatures of this modification, including diagnostic *b* and *y* ions as well as a signature fragment of the biotin tag ( $m/z = 340.17$ ), were used to confirm the identification of biotinylation (Fig. 1B). A total of 6700 biotinylated peptides were detected across the three replicates, and a majority of them, i.e., 5851 (87%) were detected and quantified in all replicates (Fig. 1C). These peptides were mapped to 1409 proteins (Additional file 3: Tables S1, S2) and showed consistent levels across the replicates (Additional file 1: Fig. S1).

A large number of the biotinylated proteins were detected by <10 biotinylated peptides (Fig. 2A). A total of 587 proteins were detected by a single biotinylated peptide each while 247 proteins were identified by two biotinylated peptides each. To estimate the relative contribution of each protein to the proteome exposed at the cell surface, biotinylation was quantified by the sum of the abundance of all biotinylated peptides derived from that protein. CD59, a GPI-anchored inhibitor of complement-mediated decay, was the single largest contributor of biotinylated peptides by abundance. Although identified by only 5 biotinylated peptides, these were the most



**Fig. 1** Surface proteome enrichment and direct detection of biotinylated peptides. **A** Experimental strategy. PC-3 cells were treated with a membrane-impermeant non-cleavable biotinylation reagent for preferential labeling of extracellular lysines. Proteins were harvested and digested prior to enrichment of biotinylated peptides using anti-biotin antibodies for LC-MS/MS analysis. **B** Annotated MS/MS spectrum of a representative biotinylated peptide mapped to the GPI-anchored cell surface protein 5'-nucleotidase (NT5E, CD73). The fragment ions in red are diagnostic of the label and localization of the site of biotinylation. **C** The number and overlap of biotinylated peptides identified from the three replicates

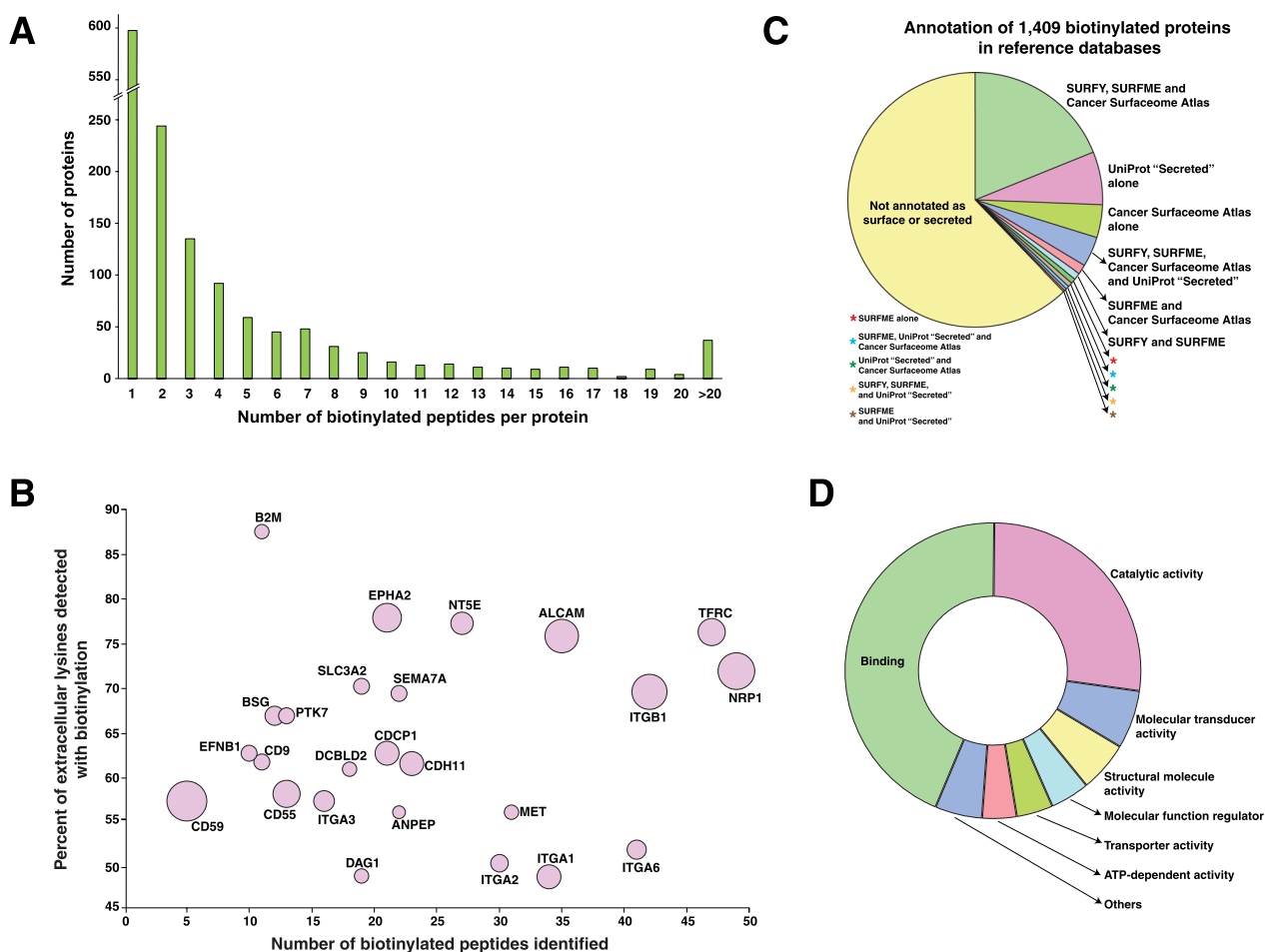
abundant peptides from any protein, indicating that this molecule is highly expressed at the surface. The relative contribution of the top 25 proteins to the abundance of biotinylated peptides is shown in Fig. 2B.

We next assessed the robustness of the sBioSITE method for enrichment of proteins that are known to be expressed on the cell surface. We compared the list of biotinylated proteins to three resources: one, SURFY, a machine learning-based in silico surface proteome database [3], two, the Cancer Surfaceome Atlas, a database of genes encoding surface proteins that is based on experimental evidence, computational prediction and database annotation [18] and three, SURFME, a manually curated catalog of surface proteins [19]. Additionally, as many secreted proteins are part of the extracellular matrix which can also be labeled or can act as ligands that bind surface proteins, we made additional annotations for biotinylated proteins with “secreted” as their subcellular location in the UniProt resource [20]. As shown in Fig. 2C, several proteins were accounted for by multiple resources, with 38% of identified proteins annotated as surface or secreted molecules by at least one of these databases. To determine the broad functions of the enriched molecules in cellular processes, we carried out Gene Ontology analysis of biotinylated proteins. As shown in Fig. 2D, the majority of annotated molecules

had the molecular functions of “binding” (44%), “catalytic activity” (27%), “molecular transducer activity” (6%) and structural molecule activity (6%) [21].

**Different surface proteins are biotinylated to varying degrees**

As the biotinylation reagent employed for sBioSITE preferentially reacts with primary amines, we assessed the number of potential biotinylated sites. For proteins with annotations in the UniProt database of extracellular, transmembrane and intracellular regions, we analyzed the data to determine the proportion of extracellular lysines that were detected with biotinylation [20]. On average, ~20% of the extracellular lysine residues were labeled per protein with considerable variation in labeling efficiency. As shown in Fig. 3A, many proteins were identified with high degrees of biotinylation. For example, junctional adhesion molecule A (F11R), involved in the formation of epithelial tight junctions, was identified with eight out of nine (89%) extracellular lysines biotinylated. Presenilin-1 (PSEN1), a multi-pass membrane protein with nine transmembrane domains, was detected with biotinylation at the single lysine residue present in the extracellular region. On the other hand, several molecules involved in cell adhesion, including protocadherin-9 (PCDH9), cadherins 6 and 3 (CDH6,

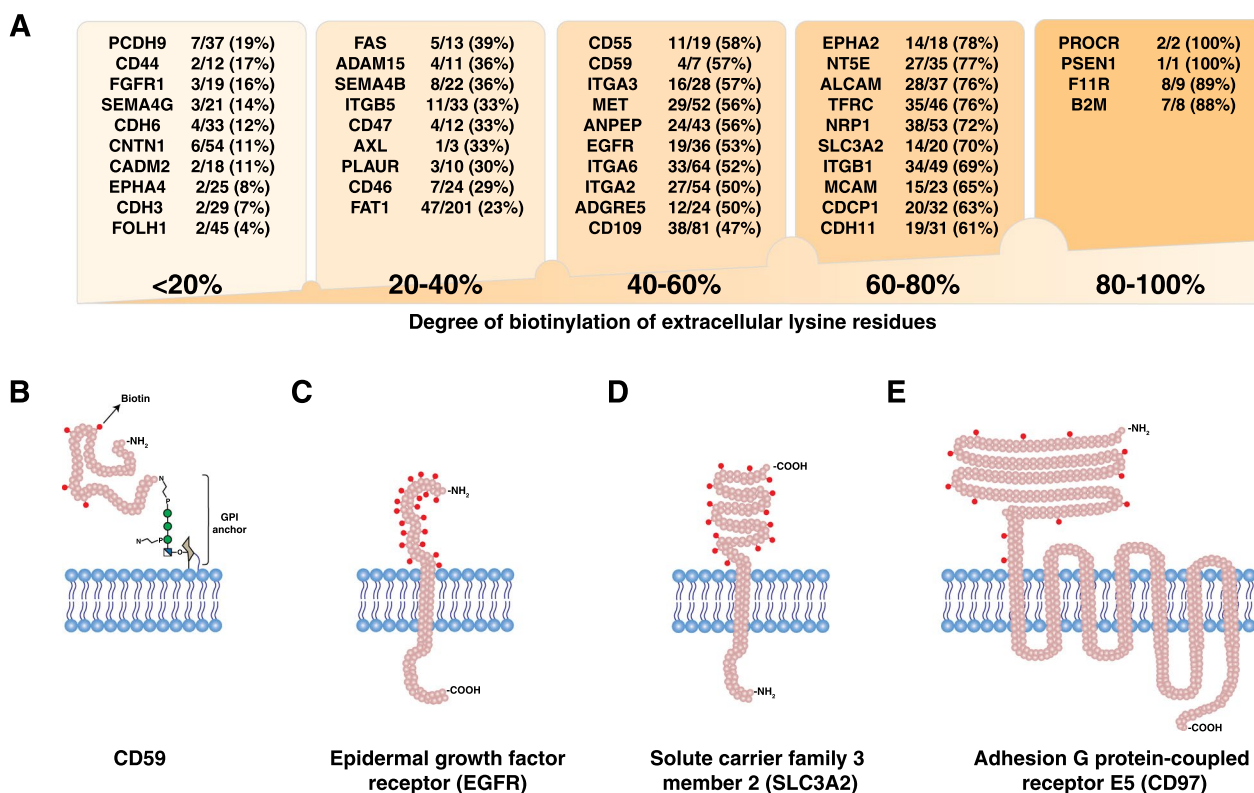


**Fig. 2** Biotinylation of cell surface proteins. **A** Bar chart showing the number of proteins identified with different numbers of biotinylated peptides. The y-axis represents the number of proteins identified by the number of biotinylated peptides shown on the x-axis. **B** Bubble plot representing the intensity of biotinylated peptides from the topmost abundant biotinylated proteins. The x-axis represents the number of biotinylated peptides identified per protein, and the y-axis represents the percent of biotinylation, as the proportion of known extracellular lysines in that protein that were detected with biotinylation. The diameter of each bubble is representative of the intensity of biotinylated peptides from that protein. **C** Pie chart showing the proportions of various proteins that were present in the resources referred to: SURFY (the in silico human surfaceome [3]), the Cancer Surfaceome Atlas [18], SURFME [19] and UniProt for annotations of secreted proteins [20]. **D** Relative proportions of the Gene Ontology annotations for molecular function of the identified biotinylated proteins

CDH3) and cell adhesion molecule 2 (CADM2) were detected with <20% of extracellular lysines biotinylated. Interestingly, prostate specific membrane antigen or glutamate carboxypeptidase 2 (PSMA/FOLH1) was identified with only two out of forty-five (4%) extracellular lysines biotinylated. A selection of proteins with different types of attachment to the plasma membrane are shown in Fig. 3B–D and E with representations of biotinylated lysine residues [22].

We wished to determine how the different classically known groups of cell surface proteins were biotinylated. Seventy-eight proteins from the Human Cell Differentiation Molecules, or CD proteins, (HCDM database [23]) were biotinylated. Of the 1409 proteins detected with

biotin, twenty-seven were annotated in the UniProt database as being GPI-anchored, and 458 proteins had annotations for transmembrane domains. Of these, 324 had annotations for single transmembrane domains (Additional file 2: Fig. S2A). Four hundred proteins identified were annotated with signal peptide sequences. Nineteen proteins were annotated with the Gene Ontology term of G-protein coupled receptor activity [20, 24], and an additional thirteen proteins were annotated as receptor tyrosine kinases in UniProt or HGNC [25]. Thirty-one proteins were from the solute carrier (SLC) group of proteins. On the other hand, in the subset of 876 biotinylated proteins that had no annotations for surface localization or secretion, 61 proteins had annotations for single



**Fig. 3** Extracellular lysines are labeled. **A** Representative list of biotinylated proteins grouped by degree of biotinylation, shown along with the ratio of biotinylated lysines to extracellular lysines. **B–E** Graphical representation of predicted topology of surface proteins and biotinylated lysine residues, for selected proteins as shown (not drawn to scale)

transmembrane domains and 44 for multiple transmembrane domains. Thirty-three of these unannotated proteins also had signal sequences (Additional file 2: Fig. S2B). A representative list of the top abundant biotinylated proteins is shown in Table 1 along with the number of extracellular lysines biotinylated in each of them. A complete list of all biotinylated proteins along with the number of biotinylated peptides and surface annotations is given in Additional file 3: Table S2.

### Discussion

Though cell surface proteins are of great importance in basic and translational research, enriching them for analysis is difficult owing to biochemical and analytical reasons. We have developed sBioSITE—a novel method to enrich surface proteins by labeling them with biotin followed by enrichment at the peptide level using anti-biotin antibodies, and detection of peptides with intact biotin tags. A unique feature of the sBioSITE method is the detection of intact biotinylated peptides, indicating extracellular localization of the labeled lysine residue. This gives the identified protein a degree of confidence not afforded by methods that do not allow the detection of intact labeled peptides. Our strategy combined the

use of a non-cleavable biotinylation reagent with enrichment using anti-biotin antibodies, enabling us to detect peptides with biotin tags intact. We profiled the surface proteome of PC-3, a prostatic adenocarcinoma cell line and demonstrated sensitive and reliable identification of cell surface proteins. A large number of the identified proteins are known to be localized to the cell surface or secreted into the extracellular space as matrix proteins [26].

### Markers of prostate cancer are abundantly enriched

Many of the biotinylated peptides detected with high abundance were derived from surface proteins that are associated with prostate cancer. A list of selected proteins is shown in Table 1 along with their relevance to prostate cancer. The protein detected with the most abundant biotinylated peptides, CD59, is a GPI-anchored protein which inhibits complement-mediated cell lysis. It has previously been shown to be highly expressed in prostate cancer in association with disease progression and adverse prognosis [27] (Fig. 3B). Another CD marker, CD166 (ALCAM) was detected with 35 biotinylated peptides. CD166 is a marker of prostate cancer progression with a role in metastasis to bone, and is known to be shed

**Table 1** Selected proteins detected with abundant biotinylated peptides and their known association with prostate cancer

	<b>Protein</b>	<b>Number of biotinylated extracellular lysine residues/ number of all extracellular lysines (percent)</b>	<b>Association with plasma membrane</b>	<b>Relevance in prostate cancer</b>
1	CD59	4/7 (57%)	GPI-anchored	Associated with disease progression and adverse prognosis [27]
2	Neuropilin-1 (NRP1)	38/53 (72%)	Single-pass type I membrane protein	Promotes progression through EGFR/AKT signaling axis [30]
3	Integrin beta-1 (ITGB1)	34/49 (69%)	Single-pass type I membrane protein	Involved in oncogenic TGF- $\beta$ signaling [38]
4	CD166 (ALCAM)	28/37 (76%)	Single-pass type I membrane protein	Marker of prostate cancer progression; regulates bone metastasis [28]
5	Ephrin type-A receptor 2 (EPHA2)	14/18 (78%)	Single-pass type I membrane protein	Involved in metastasis [50]
6	Complement decay-accelerating factor (CD55)	11/19 (58%)	GPI-anchored	Promotes cancer cell survival and tumor growth [51]
7	Transferrin receptor protein 1 (TFRC)	35/46 (76%)	Single-pass type II membrane protein	Induces proliferation, migration and invasion in cell lines; ferroptosis-related biomarker [36, 37]
8	CUB domain-containing protein 1 (CDCP1)	20/32 (63%)	Single-pass type I membrane protein	Highly expressed in castration-resistant prostate cancer; therapeutic target [52, 53]
9	Cadherin-11 (CDH11)	19/31 (61%)	Single-pass type I membrane protein	Increases migration and invasion [54]
10	5'-Nucleotidase (NT5E, CD73)	27/35 (77%)	GPI-anchored	Suppresses immune surveillance, prognostic factor [55]
11	Integrin alpha-6 (ITGA6, CD49F)	33/64 (52%)	Single-pass type I membrane protein	Marker of prostate cancer; necessary for self-renewal activity of prostate stem cells [39, 40]
12	Basigin (BSG, CD147)	8/12 (67%)	Single-pass type I membrane protein	Associated with progression; stimulates production of matrix metalloproteases [56, 57]
13	Integrin alpha-2 (ITGA2, CD49b)	27/54 (50%)	Single-pass type I membrane protein	Promotes prostate cancer cell growth within bone [41]
14	Inactive tyrosine-protein kinase 7 (PTK7)	16/24 (67%)	Single-pass type I membrane protein	Prognostic biomarker; predictor of lymph node metastasis [58]
15	4F2 cell-surface antigen heavy chain (SLC3A2)	14/20 (70%)	Single-pass type II membrane protein	Promotes progression via S-phase kinase-associated protein 2 (SKP2) [33]
16	Hepatocyte growth factor receptor (MET)	29/52 (56%)	Single-pass type I membrane protein	Highly expressed in bone metastasis [59, 60]
17	Glypican-4 (GPC4)	25/37 (68%)	GPI-anchored	Interacts with stromal cells; extracellular matrix remodeling, endocrine/paracrine signaling [61]
18	Melanoma cell adhesion molecule (MCAM)	15/23 (65%)	Single-pass type I membrane protein	Involved in metastasis by mediating E-selectin-dependent interaction with bone marrow endothelium [62, 63]
19	Adhesion G protein-coupled receptor E5 (ADGRE5, CD97)	12/24 (50%)	Multi-pass membrane protein	Mediates invasion of prostate cancer cells in association with lysophosphatidic acid receptor 1 (LPAR1) [64]
20	Epidermal growth factor receptor (EGFR)	19/36 (53%)	Single-pass type I membrane protein	Marker of dissemination to bones [29]

from prostatic cancer cells by the action of the sheddase ADAM17, which was also detected with abundant biotinylated peptides [28]. Epidermal growth factor receptor (EGFR), a known marker of dissemination of prostate cancer to bones, was detected with 19 biotinylation sites [29] (Fig. 3C). Neuropilin-1 (NRP1), a single-pass type I membrane protein with a large extracellular domain was detected with 38 biotinylation sites. This protein, a co-receptor for various growth factors, is known to promote prostate cancer progression via modulating EGFR-dependent AKT pathway activation [30–32]. 4F2 cell-surface antigen heavy chain (SLC3A2), a single-pass type II membrane protein which is known to promote prostate cancer progression, was detected with 14 biotinylation sites [33] (Fig. 3D). Ephrin A2 (EPHA2), a tyrosine kinase receptor that is associated with aggressive prostate cancer and adverse prognosis, was detected with 14 biotinylation sites [34]. Adhesion G protein-coupled receptor (ADGRE5, CD97), a seven transmembrane domain-containing protein, was detected with 12 biotinylation sites, all within the large N-terminal extracellular region (Fig. 3E). Similarly, transferrin receptor protein 1 (TFRC), was identified with 42 biotinylation sites. Transferrin receptor levels have been previously shown to be elevated in prostate cancer patients in association with altered iron metabolism, and has been proposed as a ferroptosis-related biomarker [35–37]. Several integrins involved in signaling and metastasis, including integrins beta-1 (ITGB1), alpha-6 (ITGA6, CD49F) and alpha-2 (ITGA2, CD49b) were detected with high abundance [38–41]. It is important to note that the degree of detected biotinylation is influenced by a variety of factors including the topology of proteins and their post-translational modifications (PTMs) [42]. For example, prostate specific membrane antigen or glutamate carboxypeptidase 2 (FOLH1/PSMA) was identified by sBioSITE as a cell surface protein and it is a well characterized diagnostic and potential therapeutic target in prostate cancer [43]. It is a known surface protein with eleven N-glycosylation sites that could heavily influence the number of detected biotinylation sites.

#### High sensitivity of surface protein detection

A widely used method for enriching the surface proteome relies on the labeling of exposed sugar residues of glycoproteins [16]. Such a strategy would not identify surface proteins which do not have oligosaccharides available for labeling, including the 436 proteins that are not known to be glycosylated according to UniProt, but were detected by biotinylation in the current study (Additional file 3: Table S1). For example, matrix metalloproteinase-14 (MMP14) is a membrane protein with no known glycosylation sites; it was identified by 16 biotinylated

peptides. However, we acknowledge that at least some of the proteins without annotations for N-glycosylation in the UniProt database may actually be glycosylated. Additionally, several proteins have been predicted to be surface localized by machine learning in the *in silico* surfaceome resource [3]. In our data, we have experimental evidence for some of these predicted proteins to be localized to the cell surface. These include dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1 (RPN1), identified with 10 biotinylation sites, and cation-dependent mannose-6-phosphate receptor (M6PR), identified with 6 biotinylation sites.

Many biotinylated proteins were not annotated as cell surface or secreted molecules in the resources we referred to. Literature curation for some of these proteins revealed that they are known or believed to translocate to the surface or are secreted under some special circumstances. For example, plectin, a cytoskeletal protein involved in tethering the cytoskeleton to membrane complexes, had 114 biotinylated peptides. Although plectin is not annotated as a surface protein, it has been shown to have aberrant localization to the cell surface in some cancers [44]. Several histone proteins including H1.3, H1.2, H1.4 and H1.5 were also identified with biotinylated peptides, consistent with reports that histones are secreted in several cancers including prostate cancer [45, 46]. Nucleolin, a nucleolar protein, was identified by 18 biotinylated peptides; it has been shown to be expressed on the cell surface in association with the actin cytoskeleton in some cancers [47, 48]. Of the 876 proteins not annotated as surface or secreted proteins, 425 were detected with only one biotinylated peptide. As such, many proteins that are otherwise not known to localize to the cell surface but were biotinylated in our data may transiently translocate to the cell surface under specific circumstances. Further experiments in additional cell lines and conditions are needed to validate the surface or extracellular localization of such proteins. Additionally, in a few cases, proteins that are annotated as surface proteins were detected with biotinylation on a few lysines annotated to be “intracellular.” It is pertinent to acknowledge that our databases of the cell surface proteome are perhaps incomplete and accumulating experimental evidence will add more proteins to such lists. We also recognize that the method we report enriches both surface and secreted proteins without distinction and further data analysis may be required to separate the two classes of proteins. In addition, many surface proteins are also known to be secreted (Additional file 3: Table S1); therefore, the labeling of such proteins in the secreted form may lead to biotinylation of lysines that are annotated to be intracellular in their membrane-bound form. Another limitation of this method is that damage to the plasma membrane of even

a small proportion of cells during the labeling procedure could potentially lead to the biotinylation of a small number of intracellular proteins, or the biotinylation of intracellular lysines of membrane proteins.

### Conclusions

We have developed sBioSITE as a method for enrichment of biotinylated cell surface proteins with direct detection of labeled peptides. By applying this method for mass spectrometry-based detection of cell surface proteins and their identification by peptides with intact biotin residues, we demonstrate sensitive and confident identification of cell surface proteins. Though the adherent cells used in this study were washed before labeling, the identification of known secreted proteins points to the difficulty of differentially enriching surface proteins from live cells in the context of the extracellular matrix and the surface binding of secreted proteins. We anticipate that this method will be complementary to existing methods of cell surface proteome analysis and is best suited to comparative studies of surface proteomes of cultured cells in different experimental conditions to answer biological questions.

### Methods

#### Cell culture and labeling of surface proteins

PC-3 cells were cultured in RPMI 1640 medium with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin to ~80% confluence in 15 cm dishes, with 4 dishes used per replicate. The medium was aspirated and the cells were washed thrice with phosphate-buffered saline (PBS). The cells were then incubated in 10 ml of 0.4 mM of membrane-impermeant amine-reactive biotinylation agent sulfo-NHS-LC-biotin (APExBIO, Boston, MA) in ice-cold PBS for 30 min at 4 °C with gentle swirling of the dish once every 5 min. The solution was then discarded and the reaction quenched twice with 6 ml of 50 mM glycine in ice-cold PBS. The quenching solution was removed and the surface-labeled cells were gently scraped in 500 µl of 50 mM glycine in PBS and collected in a 1.5 ml microcentrifuge tube. The tube was centrifuged at 300×g for 5 min and supernatant was discarded. The cells were washed with ice-cold PBS twice and pelleted down by centrifugation at 300×g for 5 min and the supernatant removed.

#### Protein extraction and digestion

Cell lysis was carried out in 500 µl of modified radioimmunoprecipitation assay (mRIPA) buffer with probe sonication. The lysate was centrifuged at 14,000×g for 5 min and the supernatant was collected as the protein fraction, and the pellet discarded. Bicinchoninic acid (BCA) assay was performed to estimate protein amount.

6 mg of protein was taken from each replicate and prepared for digestion. Proteins were reduced with dithiothreitol (Sigma) at a final concentration of 10 mM for 45 min with mild shaking at 37 °C. The sample was then cooled to room temperature and alkylation was carried out with iodoacetamide (Sigma) at a final concentration of 40 mM by incubation at room temperature for 15 min in the dark. Sequencing grade trypsin was added to a final amount of 1:50 (trypsin: total protein), and the samples incubated overnight at 37 °C with mild shaking. The resulting peptides were desalted using C18 reversed phase columns (TopTips, GlyGen), dried in a vacuum centrifuge.

#### Enrichment of biotinylated peptides

Biotinylated peptides were enriched using agarose immobilized rabbit anti-Biotin antibody (Bethyl laboratories, A150109A). First, the antibody-coupled beads were washed thrice with PBS and BioSITE capture buffer (50 mM Tris, 150 mM NaCl, 0.5% Triton X-100) thrice. Peptides were dissolved in 600 µl of BioSITE capture buffer. After dissolving peptides, pH was adjusted to neutral (7.0 to 7.5) and colorimetric peptide assay (Pierce) was performed to estimate peptide concentration. Peptides were subsequently incubated with anti-biotin antibody-bound protein G beads for 2 h at 4 °C. The bead slurry was sequentially washed two times with BioSITE capture buffer, two times with 50 mM Tris, and two times with ultrapure water. Biotinylated peptides were eluted four times using elution buffer (80% acetonitrile and 0.2% trifluoroacetic acid in water). The eluent was further cleaned up using C18 reversed phase columns (TopTips, GlyGen) as previously described.

#### LC-MS/MS analysis

Previously published LC-MS/MS parameters were used with some modifications. Briefly, samples enriched for biotinylated peptides were analyzed by Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific). Peptides were separated by liquid chromatography on an EASY-Spray column (75 µm × 50 cm, PepMap RSCL C18, Thermo Fisher Scientific) packed with 2 µm C18 particles, maintained at 50 °C. 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B) were used as solvents. Peptides were trapped on a trap column (100 mm × 2 cm, Acclaim PepMap100 NanoTrap, Thermo Fisher Scientific) at a flow rate of 20 µl/min. LC separation was performed at a flow rate of 300 nl/min and the following gradient was used: equilibration at 5% solvent B from 0 to 10 min, 5–40% sol B from 10.1 to 125 min, 40–95% sol B from 125 to 137 min followed by equilibration for next run at 5% sol B for 13 min. Experiments were done in DDA mode with top 15



precursor ions isolated at a window of 1.2 ( $m/z$ ). Precursors with charge states ranging from +2 to +5 were considered for MS/MS events. Normalized collision energy was applied to fragment precursors at energies of 30%. Precursor ions were acquired in the Orbitrap mass analyzer in range of 340–1600  $m/z$  at a resolution of 120,000. Fragment ion spectra were detected in Orbitrap mass analyzer with a resolution of 15,000. Automatic gain control (AGC) target value for MS and MS/MS were  $10^6$  and  $10^5$  and maximum ion injection time were set as 50 ms and 250 ms respectively. Exclude isotopes feature was set to “on” and 20 s dynamic exclusion was applied.

### Mass spectrometry data analysis

Mass spectrometry raw data were searched against UniProt Human Reviewed protein sequences (20,432 entries, downloaded February 1, 2021) by Sequest HT on ProteomeDiscoverer (ver. 2.5., ThermoFisher) [20]. Fully tryptic cleavage specificity with 4 missed cleavages was used and precursor and fragment tolerance were set to 10 ppm and 0.02 Da respectively. Carbamidomethylation of cysteine was set as fixed modification. Oxidation of methionine and protein N-terminal acetylation were set as variable modifications. Biotinylation with NHS-LC-Biotin (+ 339.16 Da) was set as a variable modification at lysine (K). Target false discovery rate (FDR) target was set at 0.01. Minora feature detection was used for precursor ion quantitation calculated using precursor intensity.

### Data analysis

MS/MS spectra of all peptides identified with biotinylated lysine residues were filtered for those containing a signature ion of the biotin tag ( $m/z = 340.1698 \pm 10$  ppm) and 22 peptides without this signature were removed from analysis. Peptides identified in all replicates were mapped to proteins and considered for further analysis as described. Gene Ontology analysis of enriched proteins was carried out using PANTHER [21].

### Associated data

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [49] partner repository with the dataset identifier PXD044519.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12014-023-09445-6>.

**Additional file 1: Figure S1.** Scatter plots showing consistency of detection of biotinylated peptides. Intensity values of biotinylated peptides

from each replicate were log-transformed and plotted against each other in pairs.

**Additional file 2: Figure S2.** Features of identified proteins. A. Number of proteins among all biotinylated proteins grouped by distinguishing features B. Number of biotinylated proteins that have not been annotated as cell surface or secreted proteins with features as shown.

**Additional file 3: Table S1.** Biotinylated peptides detected in common in all replicates. **Table S2.** Biotinylated proteins detected in common in all replicates.

### Author contributions

KG, HD, MCC, YK, MS and SC performed the experiments. KG, RZ, DGM, SC, AS, FL, RKK and AP analyzed the data. KG, RKK and AP wrote the manuscript. KG and RKK prepared the figures. JZ, RKK and AP conceived the study and supervised it. All authors reviewed the manuscript.

### Funding

This work was supported in part by grants from NCI to A.P. (U01CA271410 and P30CA15083).

### Declarations

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Manipal Academy of Higher Education (MAHE), Manipal, Karnataka, India. <sup>2</sup>Institute of Bioinformatics, International Technology Park, Bangalore, Karnataka, India. <sup>3</sup>Department of Laboratory Medicine and Pathology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA. <sup>4</sup>Proteomics Core, Mayo Clinic, Rochester, MN, USA. <sup>5</sup>Department of Urology, Mayo Clinic, Rochester, MN, USA. <sup>6</sup>Department of Immunology, Mayo Clinic, Rochester, MN, USA. <sup>7</sup>Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, USA. <sup>8</sup>Center for Individualized Medicine, Mayo Clinic, Rochester, MN, USA.

Received: 21 August 2023 Accepted: 17 November 2023

Published online: 05 December 2023

### References

1. Tan S, Tan HT, Chung MC. Membrane proteins and membrane proteomics. *Proteomics*. 2008;8(19):3924–32.
2. Goodwin J, Laslett AL, Rugg-Gunn PJ. The application of cell surface markers to demarcate distinct human pluripotent states. *Exp Cell Res*. 2020;387(1): 111749.
3. Bausch-Fluck D, Goldmann U, Muller S, van Oostrum M, Muller M, Schubert OT, et al. The in silico human surfaceome. *Proc Natl Acad Sci USA*. 2018;115(46):E10988–97.
4. Arispe N, Doh M. Plasma membrane cholesterol controls the cytotoxicity of Alzheimer's disease Aβ(1–40) and (1–42) peptides. *FASEB J*. 2002;16(12):1526–36.
5. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. 1987;235(4785):177–82.
6. Brekke OH, Sandlie I. Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nat Rev Drug Discov*. 2003;2(1):52–62.
7. Wu CC, Yates JR III. The application of mass spectrometry to membrane proteomics. *Nat Biotechnol*. 2003;21(3):262–7.
8. Wallin E, von Heijne G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci*. 1998;7(4):1029–38.
9. Josic D, Clifton JG. Mammalian plasma membrane proteomics. *Proteomics*. 2007;7(16):3010–29.
10. Goshe MB, Blonder J, Smith RD. Affinity labeling of highly hydrophobic integral membrane proteins for proteome-wide analysis. *J Proteome Res*. 2003;2(2):153–61.

11. Huber LA, Pfaller K, Vietor I. Organelle proteomics: implications for subcellular fractionation in proteomics. *Circ Res*. 2003;92(9):962–8.
12. Bordier C. Phase separation of integral membrane proteins in Triton X-114 solution. *J Biol Chem*. 1981;256(4):1604–7.
13. Ghosh D, Krokhin O, Antonovici M, Ens W, Standing KG, Beavis RC, et al. Lectin affinity as an approach to the proteomic analysis of membrane glycoproteins. *J Proteome Res*. 2004;3(4):841–50.
14. Elia G. Biotinylation reagents for the study of cell surface proteins. *Proteomics*. 2008;8(19):4012–24.
15. Ferguson ID, Patino-Escobar B, Tuomivaara ST, Lin YT, Nix MA, Leung KK, et al. The surfaceome of multiple myeloma cells suggests potential immunotherapeutic strategies and protein markers of drug resistance. *Nat Commun*. 2022;13(1):4121.
16. Wollscheid B, Bausch-Fluck D, Henderson C, O'Brien R, Bibel M, Schiess R, et al. Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat Biotechnol*. 2009;27(4):378–86.
17. Kim DI, Cutler JA, Na CH, Reckel S, Renuse S, Madugundu AK, et al. BioSITE: a method for direct detection and quantitation of site-specific biotinylation. *J Proteome Res*. 2018;17(2):759–69.
18. Hu Z, Yuan J, Long M, Jiang J, Zhang Y, Zhang T, et al. The cancer surfaceome atlas integrates genomic, functional and drug response data to identify actionable targets. *Nat Cancer*. 2021;2(12):1406–22.
19. Governa V, Talbot H, Goncalves de Oliveira K, Cerezo-Magana M, Bang-Rudensam A, Johansson MC, et al. Landscape of surfaceome and endocytome in human glioma is divergent and depends on cellular spatial organization. *Proc Natl Acad Sci USA*. 2022;119(9): e2114456119.
20. UniProt C. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;49(D1):D480–9.
21. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, et al. PANTHER version 11: expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. 2017;45(D1):D183–9.
22. Omasits U, Ahrens CH, Muller S, Wollscheid B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*. 2014;30(6):884–6.
23. Zola H, Swart B, Banham A, Barry S, Beare A, Bensussan A, et al. CD molecules 2006—human cell differentiation molecules. *J Immunol Methods*. 2007;319(1–2):1–5.
24. Isberg V, Mordalski S, Munk C, Rataj K, Harpsøe K, Hauser AS, et al. GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res*. 2016;44(D1):D356–64.
25. Seal RL, Braschi B, Gray K, Jones TEM, Tweedie S, Haim-Vilmsky L, et al. Genenames.org: the HGNC resources in 2023. *Nucleic Acids Res*. 2023;51(D1):D1003–9.
26. Kaighn ME, Narayan KS, Ohnuki Y, Lechner JF, Jones LW. Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest Urol*. 1979;17(1):16–23.
27. Xu C, Jung M, Burkhardt M, Stephan C, Schnorr D, Loening S, et al. Increased CD59 protein expression predicts a PSA relapse in patients after radical prostatectomy. *Prostate*. 2005;62(3):224–32.
28. Hansen AG, Arnold SA, Jiang M, Palmer TD, Ketova T, Merkel A, et al. ALCAM/CD166 is a TGF-beta-responsive marker and functional regulator of prostate cancer metastasis to bone. *Cancer Res*. 2014;74(5):1404–15.
29. Nastaly P, Stoupić S, Popeda M, Smentoch J, Schlomm T, Morrissey C, et al. EGFR as a stable marker of prostate cancer dissemination to bones. *Br J Cancer*. 2020;123(12):1767–74.
30. Zhang P, Chen L, Zhou F, He Z, Wang G, Luo Y. NRP1 promotes prostate cancer progression via modulating EGFR-dependent AKT pathway activation. *Cell Death Dis*. 2023;14(2):159.
31. Tse BWC, Volpert M, Rattner E, Stylianou N, Nouri M, McGowan K, et al. Neuropilin-1 is upregulated in the adaptive response of prostate tumors to androgen-targeted therapies and is prognostic of metastatic progression and patient mortality. *Oncogene*. 2017;36(24):3417–27.
32. Glinka Y, Stoilova S, Mohammed N, Prud'homme GJ. Neuropilin-1 exerts co-receptor function for TGF-beta-1 on the membrane of cancer cells and enhances responses to both latent and active TGF-beta. *Carcinogenesis*. 2011;32(4):613–21.
33. Maimaiti M, Sakamoto S, Sugiura M, Kanesaka M, Fujimoto A, Matsusaka K, et al. The heavy chain of 4F2 antigen promote prostate cancer progression via SKP-2. *Sci Rep*. 2021;11(1):11478.
34. Kurose H, Ueda K, Kondo R, Ogasawara S, Kusano H, Sanada S, et al. Elevated expression of EPHA2 is associated with poor prognosis after radical prostatectomy in prostate cancer. *Anticancer Res*. 2019;39(11):6249–57.
35. Kuvibidila S, Gauthier T, Warriar RP, Rayford W. Increased levels of serum transferrin receptor and serum transferrin receptor/log ferritin ratios in men with prostate cancer and the implications for body-iron stores. *J Lab Clin Med*. 2004;144(4):176–82.
36. Liang J, Liao Y, Wang P, Yang K, Wang Y, Wang K, et al. Ferroptosis landscape in prostate cancer from molecular and metabolic perspective. *Cell Death Discov*. 2023;9(1):128.
37. Liu H, Gao L, Xie T, Li J, Zhai TS, Xu Y. Identification and validation of a prognostic signature for prostate cancer based on ferroptosis-related genes. *Front Oncol*. 2021;11: 623313.
38. Pellinen T, Blom S, Sanchez S, Valimaki K, Mpindi JP, Azegrouz H, et al. ITGB1-dependent upregulation of caveolin-1 switches TGFbeta signaling from tumour-suppressive to oncogenic in prostate cancer. *Sci Rep*. 2018;8(1):2338.
39. Lawson DA, Xin L, Lukacs RU, Cheng D, Witte ON. Isolation and functional characterization of murine prostate stem cells. *Proc Natl Acad Sci USA*. 2007;104(1):181–6.
40. Mulholland DJ, Xin L, Morim A, Lawson D, Witte O, Wu H. Lin-Sca-1+CD49high stem/progenitors are tumor-initiating cells in the Pten-null prostate cancer model. *Cancer Res*. 2009;69(22):8555–62.
41. Hall CL, Dai J, van Golen KL, Keller ET, Long MW. Type I collagen receptor (alpha 2 beta 1) signaling promotes the growth of human prostate cancer cells within the bone. *Cancer Res*. 2006;66(17):8648–54.
42. Suckau D, Mak M, Przybylski M. Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping. *Proc Natl Acad Sci USA*. 1992;89(12):5630–4.
43. Chang SS. Overview of prostate-specific membrane antigen. *Rev Urol*. 2004;6(Suppl 10):13–8.
44. Perez SM, Brinton LT, Kelly KA. Plectin in cancer: from biomarker to therapeutic target. *Cells*. 2021;10(9):2246.
45. Jeon JM, Kwon OK, Na AY, Sung EJ, Cho IJ, Kim M, et al. Secretome profiling of PC3/nKR cells, a novel highly migrating prostate cancer subline derived from PC3 cells. *PLoS ONE*. 2019;14(8): e0220807.
46. Chen R, Kang R, Fan XG, Tang D. Release and activity of histone in diseases. *Cell Death Dis*. 2014;5(8):e1370.
47. Soundararajan S, Wang L, Sridharan V, Chen W, Courtenay-Luck N, Jones D, et al. Plasma membrane nucleolin is a receptor for the anticancer aptamer AS1411 in MV4-11 leukemia cells. *Mol Pharmacol*. 2009;76(5):984–91.
48. Hovanessian AG, Puvion-Dutilleul F, Nisole S, Svab J, Perret E, Deng JS, et al. The cell-surface-expressed nucleolin is associated with the actin cytoskeleton. *Exp Cell Res*. 2000;261(2):312–28.
49. Perez-Riverol Y, Bai J, Bandla C, Garcia-Seisdedos D, Hewapathirana S, Kamatchinathan S, et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*. 2022;50(D1):D543–52.
50. Taddei ML, Parri M, Angelucci A, Bianchini F, Marconi C, Giannoni E, et al. EphA2 induces metastatic growth regulating amoeboid motility and clonogenic potential in prostate carcinoma cells. *Mol Cancer Res*. 2011;9(2):149–60.
51. Loberg RD, Day LL, Dunn R, Kalikin LM, Pienta KJ. Inhibition of decay-accelerating factor (CD55) attenuates prostate cancer growth and survival in vivo. *Neoplasia*. 2006;8(11):69–78.
52. Zhao N, Chopra S, Trepka K, Wang YH, Sakhamuri S, Hooshdaran N, et al. CUB domain-containing protein 1 (CDCP1) is a target for radioligand therapy in castration-resistant prostate cancer, including PSMA null disease. *Clin Cancer Res*. 2022;28(14):3066–75.
53. Ji D, Shang G, Wei E, Jia Y, Wang C, Zhang Q, et al. Targeting CDCP1 gene transcription coactivated by BRD4 and CBP/p300 in castration-resistant prostate cancer. *Oncogene*. 2022;41(23):3251–62.
54. Huang CF, Lira C, Chu K, Bilan MA, Lee YC, Ye X, et al. Cadherin-11 increases migration and invasion of prostate cancer cells and enhances their interaction with osteoblasts. *Cancer Res*. 2010;70(11):4580–9.
55. Leclerc BG, Charlebois R, Chouinard G, Allard B, Pommey S, Saad F, et al. CD73 expression is an independent prognostic factor in prostate cancer. *Clin Cancer Res*. 2016;22(1):158–66.

56. Zhong WD, Liang YX, Lin SX, Li L, He HC, Bi XC, et al. Expression of CD147 is associated with prostate cancer progression. *Int J Cancer*. 2012;130(2):300–8.
57. Guo H, Li R, Zucker S, Toole BP. EMMPRIN (CD147), an inducer of matrix metalloproteinase synthesis, also binds interstitial collagenase to the tumor cell surface. *Cancer Res*. 2000;60(4):888–91.
58. Zhang H, Wang A, Qi S, Cheng S, Yao B, Xu Y. Protein tyrosine kinase 7 (PTK7) as a predictor of lymph node metastases and a novel prognostic biomarker in patients with prostate cancer. *Int J Mol Sci*. 2014;15(7):11665–77.
59. Humphrey PA, Zhu X, Zarnegar R, Swanson PE, Ratliff TL, Vollmer RT, et al. Hepatocyte growth factor and its receptor (c-MET) in prostatic carcinoma. *Am J Pathol*. 1995;147(2):386–96.
60. Verhoef EI, Klijn K, De Herdt MJ, van der Steen B, Hoogland AM, Sledens HF, et al. MET expression during prostate cancer progression. *Oncotarget*. 2016;7(21):31029–36.
61. Quach ND, Kaur SP, Eggert MW, Ingram L, Ghosh D, Sheth S, et al. Paradoxical role of glypican-1 in prostate cancer cell and tumor growth. *Sci Rep*. 2019;9(1):11478.
62. Barthel SR, Wiese GK, Cho J, Opperman MJ, Hays DL, Siddiqui J, et al. Alpha 1,3 fucosyltransferases are master regulators of prostate cancer cell trafficking. *Proc Natl Acad Sci USA*. 2009;106(46):19491–6.
63. Datta D, Aftabuddin M, Gupta DK, Raha S, Sen P. Hum prostate cancer Hallm Map. *Sci Rep*. 2016;6:30691.
64. Ward Y, Lake R, Yin JJ, Heger CD, Raffeld M, Goldsmith PK, et al. LPA receptor heterodimerizes with CD97 to amplify LPA-initiated RHO-dependent signaling and invasion in prostate cancer cells. *Cancer Res*. 2011;71(23):7301–11.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

