Identification of proteins associated with development of metastasis from cutaneous squamous cell carcinomas (cSCCs) via proteomic analysis of primary cSCCs*

A. Shapanis,1,2 C. Lai,1,2 S. Smith,3 G. Coltart,1,2 M. Sommerlad,4 J. Schofield,5 E. Parkinson,5 P. Skipp5 and E. Healy1,2

1Dermatopharmacology, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK
2Dermatology, University Hospital Southampton NHS Foundation Trust, Southampton, UK
3Department of Pathology, University of Cambridge, Cambridge, UK
4Histopathology, University Hospital Southampton NHS Foundation Trust, Southampton, UK
5Centre for Proteomic Research, Institute for Life Sciences, University of Southampton, Southampton, UK

Linked Comment: Tobin. Br J Dermatol 2021; 184:593–594.

Summary

Background Cutaneous squamous cell carcinoma (cSCC) is one of the most common cancers capable of metastasizing. Proteomic analysis of cSCCs can provide insight into the biological processes responsible for metastasis, as well as future therapeutic targets and prognostic biomarkers.

Objectives To identify proteins associated with development of metastasis in cSCC.

Methods A proteomic-based approach was employed on 105 completely excised, primary cSCCs, comprising 52 that had metastasized (P-M) and 53 that had not metastasized at 5 years post-surgery (P-NM). Formalin-fixed, paraffin-embedded cSCCs were microdissected and subjected to proteomic profiling after one-dimensional (1D), and separately two-dimensional (2D), liquid chromatography fractionation.

Results A discovery set of 24 P-Ms and 24 P-NMs showed 144 significantly differentially expressed proteins, including 33 proteins identified via both 1D and 2D separation, between P-Ms and P-NMs. Several differentially expressed proteins were also associated with survival in SCCs of other organs. The findings were verified by multiple reaction monitoring on six peptides from two proteins, annexin A5 (ANXA5) and dolichyl-diphosphooligosaccharide–protein glycosyltransferase noncatalytic subunit (DDOST), in the discovery group and validated on a separate cohort (n = 57). Increased expression of ANXA5 and DDOST was associated with reduced time to metastasis in cSCC and decreased survival in cervical and oropharyngeal cancer. A prediction model using ANXA5 and DDOST had an area under the curve of 0.93 (confidence interval 0.83–1.00), an accuracy of 91.2% and higher sensitivity and specificity than cSCC staging systems currently in clinical use.

Conclusions This study highlights that increased expression of two proteins, ANXA5 and DDOST, is significantly associated with poorer clinical outcomes in cSCC.

What is already known about this topic?

- Keratinocyte cancer is the most common cancer in the UK, and the capacity for cutaneous squamous cell carcinomas (cSCCs) to metastasize presents a clinical problem.
- Although there are known clinical risk factors for cSCC metastasis, current staging systems are inaccurate at predicting the development of metastasis in patients with cSCC.
It has been shown that mass spectrometry-based proteomic analysis can quantify and uncover potential key proteins in cancer development and metastasis.

What does this study add?

- This study has identified a number of proteins that are differentially expressed between primary cSCCs that metastasize and primary cSCCs that do not metastasize.
- Expression of the genes encoding several of these proteins influences the outcome in SCCs of other organs (lung, oropharynx, cervix and oesophagus).
- Higher abundances of two key proteins, annexin A5 (ANXA5) and dolichyl-diphosphooligosaccharide–protein glycosyltransferase noncatalytic subunit (DDOST), are associated with the development of, and reduced time to, cSCC metastasis.

What is the translational message?

- This is the first study to undertake proteomic profiling using mass spectrometry to investigate proteins that are differentially expressed between human primary cSCCs that metastasize and those that do not metastasize.
- The results of this proteomic analysis of cSCCs will be useful for identifying potential therapeutic targets in this cancer.
- A prediction model incorporating ANXA5 and DDOST showed higher sensitivity and specificity than cSCC clinical staging systems for estimating the likelihood of cSCC metastases.

The number of keratinocyte cancers in the UK is > 211,120 annually, with cutaneous squamous cell carcinoma (cSCC) accounting for > 44,672, constituting one of the most common types of cancer capable of metastasizing.\(^1,2\) The risk of metastasis for cSCC depends on clinical and histological parameters, including the site, depth of invasion, diameter, differentiation of the tumour, presence of lymphovascular or perineural invasion, and host immunosuppression.\(^3\) Following surgical excision, cSCC metastasizes in 16% of cases with tumour depth > 6 mm,\(^4\) and in 30% of tumours > 2 cm in diameter.\(^5\) Whereas the 3-year disease-specific survival rate for patients with cSCC is 85%,\(^6\) for patients with distant metastasis the median survival is < 2 years.\(^7\)

Staging systems assist identification of patients at greater risk of metastases after excision of primary cSCC.\(^8,9\) However, current staging systems distinguish ‘poorly to moderately’ between patients who do and do not develop cSCC metastases,\(^8\) and one-third of patients are classified incorrectly using these staging systems.\(^10\) There is a need to undertake research into factors that contribute to more aggressive tumours,\(^11\) to understand the mechanisms responsible for development of metastases in cSCC and to identify more accurately those patients at risk of metastases.

Proteomic analysis can aid in understanding the aetiology of cancer progression and provide information of prognostic relevance.\(^12\) In this study we used a mass spectrometry-based proteomic approach on cSCCs to identify proteins involved in development of metastases. The results highlight a number of differentially expressed proteins that associate with occurrence of metastases from cSCC, and reduced survival in lung, cervical, oropharyngeal and oesophageal SCC.

Materials and methods

Tissue samples

Formalin-fixed paraffin-embedded (FFPE) human primary cSCCs were acquired from Histopathology, University Hospital Southampton NHS Foundation Trust (UHS-NHSFT) under ethics committee approval (South Central Hampshire B National Research Ethics Service Committee; LREC number 07/H0504/187). Samples were categorized as primary cSCCs that metastasized (P-M) or primary cSCCs that had not metastasized at 5 years post-surgery (P-NM), with the latter based on no evidence of metastasis during 5 years of follow-up and/or patient review for another reason after 5 years in Dermatology, UHS-NHSFT.

Sample preparation for mass spectrometry

FFPE tissue sections were mounted onto glass slides, and tumour and surrounding immune infiltrate were microdissected and transferred into protein extraction buffer (Appendix S1; see Supporting Information). Samples were heated to 105 °C for 30 min, cooled, then heated to 80 °C for 2 h before reduction using dithiothreitol and alkylation with iodoacetamide. Samples were digested with sequencing-grade trypsin overnight and the resulting peptides were desalted using C18 reverse-phase clean-up plates.
Immunostaining

Standard immunostaining protocols were used. Briefly, slides were deparaffinized and rehydrated, and endogenous peroxidase blocked, before incubation overnight at 4 °C with primary antibodies for CD1a (1: 50, M3571, Dako, Glostrup, Denmark), L-plastin (1: 200, ab109129, Abcam, Cambridge, UK), annexin A5 (ANXA5) (EPR3979, Abcam), dolichyl-diphosphooligosaccharide–protein glycosyltransferase noncatalytic subunit (DDOST) (C340633, LSBio, Seattle, WA, USA) and CD8 (1: 50, 998254C, Invitrogen, Carlsbad, CA, USA). Subsequent incubation with biotinylated secondary antibody (antimouse, 1: 400, JIR 315-066-045; antirabbit, 1: 400, Vector Laboratories Ltd, Peterborough, UK) and 3,3’-diaminobenzidine as the chromogen. Slides were imaged using an Olympus VS110 virtual microscopy system (Olympus, Center Valley, PA, USA).

Discovery liquid chromatography mass spectrometry

Samples were fractionated using a nanoACQUITY UPLC System (Waters, Milford, MA, USA) and electrosprayed into a Waters SYNAPT-G2-Si mass spectrometer operating in MS² mode with ion mobility activated (Appendix S1; see Supporting Information). Estimates of absolute quantification using the Top3 approach were obtained using one-dimensional (1D) and two-dimensional (2D) liquid chromatography (LC) separation strategies. Data from the 1D and 2D LC procedures were analysed separately. Three blank runs were conducted between samples to ensure avoidance of carry-over into subsequent samples.

Multiple reaction monitoring

A spectral library from the discovery proteomic data was generated using Skyline software to identify unique peptides for proteins of interest. Heavy stable-isotope-labelled (SIL) peptides were synthesized by Cambridge Research Biochemicals (Billingham, UK). Calibration curves were created using 1 µg cSCC ‘proteomic-ready’ sample as background. The High Definition MRM (multiple reaction monitoring) acquisition mode was used for targeted acquisition. Transitions for each peptide were identified using Skyline and imported into MassLynx (Waters) for targeted acquisition. Samples were analysed containing 100 fmol of each heavy SIL peptide. Raw data were imported into Skyline for interpretation and calculation of native peptide quantity.

Gene expression in other squamous cell carcinomas

Expression levels of relevant genes were analysed in publicly available RNA-sequencing data from The Cancer Genome Atlas (TCGA) research network (http://cancergenome.nih.gov). Computational analysis and statistical testing of next-generation sequencing data were conducted using R statistical programming language. Filtered and log₂-normalized RNA expression data, alongside available clinical data, were downloaded from the GDAC firehose database (run: std-data_2015_06_01; https://gdac.broadinstitute.org). Plotting of TCGA data was performed using the ggplot2 package in R. Survival analysis was performed using the R packages survminer and survival. Kaplan–Meier survival curves were constructed using TCGA clinical data. Statistical testing of differences between survival curves used the G-rho family of tests, as implemented in the survdiff function of the survival package.

Data analysis

A 1% false discovery rate was applied to searching for peptide identification. Each protein was inferred from identification of at least one unique peptide. Only proteins detected in ≥ 50% of samples were subsequently analysed. Data were normalized to the median protein concentration for each sample and P-values were obtained by Mann–Whitney U-test. Topological data analysis, using Ayasdi (Palo Alto, CA, USA), was performed on complete, normalized proteomic data with a hammer metric and two neighbourhood lenses. For Kaplan–Meier survival analysis, P-values were obtained by log-rank test. Machine learning was performed using the statistical programming language, R, with the packages caret and caretensembles.

Results

Discovery proteomics

This study investigated proteomic differences between P-M and P-NM cSCCs to identify proteins associated with metastasis in cSCC. As expected, more patients in the P-M than in the P-NM group had poorly differentiated tumours or perineural invasion or were immunosuppressed (Table 1). A discovery group of 24 P-M and 24 P-NM samples was subjected to proteomic profiling using 1D, and independently 2D, separation to identify and quantify differences in protein abundance between P-Ms and P-NMs. Microdissected cSCC samples included tumour keratinocytes and stromal regions containing the immune cell infiltrate (Figure 1a). Volcano plots demonstrated higher numbers of upregulated than downregulated proteins in P-M compared with P-NM cSCCs (Figure 1b, c). Overall, 4018 unique proteins were identified in the cSCCs (Figure 1d), of which 144 were significantly differentially expressed between P-Ms and P-NMs (P < 0.05; Tables S1 and S2; see Supporting Information). This included 33 proteins identified via both 1D and 2D proteomics (Figure 1e and Table 2). Topological data analysis of the 48 proteomes from the discovery set of 24 P-M and 24 P-NM cSCCs, performed without including input information on metastases or any other clinical data, demonstrated separation of samples in both 1D and 2D analyses according to development of metastases (Figure 1f, g). This provides support for distinct proteomic profiles of P-M and P-NM cSCCs.
Table 1 Clinicopathological characteristics of the study participants

| Primary cutaneous squamous cell carcinoma | Metastasized, n = 58 | Not metastasized, 5 years postsurgery, n = 65 |
|------------------------------------------|----------------------|---------------------------------------------|
| Male/female                              | 44 (76)/14 (24)      | 47 (72)/18 (28)                              |
| Age (years), median (range)              | 82 (51–98)           | 75 (47–94)                                   |
| Site                                      |                      |                                             |
| Head and neck                            | 43 (74)              | 44 (68)                                     |
| Trunk                                     | 3 (5)                | 5 (8)                                       |
| Upper limb                                | 8 (14)               | 9 (14)                                      |
| Lower limb                                | 4 (7)                | 7 (11)                                      |
| Well differentiated                       | 1 (2)                | 22 (34)                                     |
| Moderately differentiated                 | 20 (34)              | 33 (51)                                     |
| Poorly differentiated                     | 37 (64)              | 10 (15)                                     |
| Perineural invasion                       | 12 (21)              | 3 (5)                                       |
| Immunosuppressed                          | 8 (14)               | 5 (8)                                       |
| Clark level                               |                      |                                             |
| I                                         | 0                    | 0                                           |
| II                                        | 0                    | 2 (3)                                       |
| III                                       | 2 (3)                | 10 (15)                                     |
| IV                                        | 21 (36)              | 39 (60)                                     |
| V                                         | 34 (59)              | 9 (14)                                      |
| Tumour diameter (mm), mean ± SD           | 28.6 ± 29.0          | 13.2 ± 8.26                                 |
| Tumour depth (mm), mean ± SD              | 7.46 ± 5.78          | 3.91 ± 2.49                                 |

The data are presented as n (%) unless stated otherwise. Some samples were used for both proteomic and immunohistochemical analysis, whereas other samples were used for proteomic or immunohistochemical analysis according to the amount of tissue available. Two samples from each group were removed during multiple reaction monitoring analysis due to the limited amount of tissue available.

Pathway analysis

Weighted gene coexpression network analysis (WGCNA) of the proteomics data was conducted and, following construction of a signed topological overlap matrix (TOM) of corresponding dissimilarity, hierarchical clustering was used on the dissimilarity TOM to produce modules of genes (Figure S1a, b; see Supporting Information). Modules were examined for correlation with clinical and immunohistochemical characteristics (Figure S1c) in addition to analysis for pathway enrichment (Figure S1d).

Immunohistochemical characterization showed significantly fewer CD8+ T cells and CD1a+ Langerhans cells in P-M than in P-NM samples (Figure 2a, b; and Table S3; see Supporting Information). Lower numbers of CD8+, and separately CD1a+, cells were significantly associated with reduced time to metastasis (P = 0.0041 and P = 0.0057, respectively; Figure 2c, d). In WGCNA, one module (denoted by the colour ‘blue’) correlated inversely with intratumoral CD1a+ Langerhans cell numbers (P = 0.04) but positively with Forkhead box P3+ regulatory T cell (Treg) numbers (P = 0.005) and with development of metastasis (P = 0.04). Conversely, another module of proteins (represented by the colour ‘brown’) correlated positively with the number of intratumoral CD1a+ Langerhans cells (P = 0.03) but inversely with CD3+ T-cell numbers (P = 0.03). A different module (‘turquoise’) demonstrated strong correlation with greater Clark level of invasion, an inverse correlation with peritumoral CD1a+ cell numbers, and increased pathway enrichment, including neutrophil degranulation (P = 3.3 × 10−13). Another module (‘yellow’), which was heavily enriched in the keratinization pathway (P = 2.1 × 10−17), correlated with CD3+ and CD8+ cell numbers, but inversely with tumour differentiation and CD20+ B-cell numbers.

To identify cell-signalling pathways associated with development of cSCC metastasis, Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) analysis with Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway mapping of significantly differentially expressed proteins in the 1D and 2D data was conducted. STRING analysis demonstrated highly connected structures with clusters (Figure S2a, b; see Supporting Information). KEGG pathway enrichment highlighted ribosomal proteins, protein processing in the endoplasmic reticulum, focal adhesion, extracellular matrix–receptor interactions, phosphoinositide 3-kinase (PI3K)–Akt signalling, and antigen processing and presentation as key differences between P-Ms and P-NMs (Figure S2c; see Supporting Information).

Comparison with The Cancer Genome Atlas

To determine whether proteins involved in development of cSCC metastases influence development of metastases in other SCC types, the 33 significantly differentially expressed proteins in the 1D and 2D proteomic data were compared against gene expression in cervical, oropharyngeal, oesophageal and lung SCC using TCGA. Expression levels of genes encoding for several proteins differentially expressed between P-M and P-NM cSCCs were identified as having significant effects on survival in SCCs arising at these other sites, with reduced survival associated, separately, with high expression of POSTN, DDOST, HNRNPK, COL6A3, ANXA5 and LCP1, and with low expression of CALML5 (Figure 3a–n; and Table S3; see Supporting Information). Furthermore, as immune dysfunction is important for cSCC development, and as L-plastin (LCP1) can stimulate the T-cell receptor and activate T cells,19 immunohistochemistry for LCP1 was conducted on the discovery group of cSCCs. This demonstrated more LCP1+ cells in P-Ms than in P-NMs (Figure 3o, p; and Table S3; see Supporting Information).

Multiple reaction monitoring

MRM was used to validate the discovery proteomics. MRM is a highly sensitive and specific mass spectrometry method that involves filtering the mass spectrometer on specific peptides of interest and quantifying these against known concentrations of isotopically labelled peptides spiked into the samples, enabling greater sensitivity and more accurate quantification of protein
concentrations. Firstly, machine learning (using a generalized linear model) was conducted on significantly differentially expressed proteins between P-M and P-NM cSCCs, in which a model predicting cSCC metastases was produced for every combination of two proteins on a training set and tested on a holdout cohort (two-thirds and one-third split, respectively). From >300 models, the combination of ANXA5 and DDOST gave one of the best area under the curve results, and because expression of both of these genes had been identified via TCGA as being important in reducing survival in SCCs of other organs, ANXA5 and DDOST were selected for targeted verification and validation using MRM.

Three unique peptides per protein were identified using Skyline software and synthesized as SIL peptides (Figure S3; see Supporting Information). MRM of the discovery cSCC group (22 P-Ms and 22 P-NMs) verified that there was more DDOST and ANXA5 in P-M than P-NM cSCCs (DDOST \( P < 0.0036 \), ANXA5 \( P = 0.0046 \); Figure 4a–d; and Table S3; see Supporting Information). MRM for DDOST and ANXA5 was then conducted in a different (i.e. validation) group of cSCCs, comprising 28 P-Ms and 29 P-NMs. Again, DDOST and ANXA5 levels were significantly higher in P-M than P-NM cSCCs (DDOST \( P < 0.001 \), ANXA5 \( P < 0.001 \); Figure 4e–h; and Table S3; see Supporting Information).
Table 2: Significantly differentially expressed proteins identified in both one-dimensional and two-dimensional proteomics and a summary of their respective Uniprot descriptions

| Gene ID | Uniprot ID | Description* | One-dimensional | Two-dimensional |
|---------|------------|---------------|-----------------|-----------------|
|         |            |               | Fold change<sup>b</sup> | P-value | Fold change<sup>b</sup> | P-value |
| ANXA5   | P08758     | Anticoagulant protein | 0.44 | 0.01 | 0.46 | 0.001 |
| CALML5  | Q9NZT1     | Binds to calcium; may be involved in terminal differentiation of keratinocytes | −0.90 | 0.013 | −0.81 | 0.021 |
| CCT8    | P50990     | Assists folding of proteins after ATP hydrolysis | 0.64 | 0.019 | 0.57 | 0.022 |
| COL6A3  | P12111     | Cell-binding protein | 0.66 | 0.006 | 0.50 | 0.017 |
| DDOST   | P39656     | Essential subunit of OST. Catalyses transfer of oligosaccharide to asparagine residues | 0.69 | 0.015 | 0.54 | 0.029 |
| FGB     | P02675     | Major function in haemostasis. Guides cell migration during re-epithelialization | 0.66 | 0.025 | 0.69 | 0.02 |
| GANAB   | Q14697     | Subunit of glucosidase 2 | 0.75 | 0.01 | 0.48 | 0.019 |
| GDI1    | P50395     | Regulates GDP/GTP | 0.42 | 0.024 | 0.44 | 0.019 |
| HCC1    | P62805     | Inhibits GDP dissociation from Rab proteins to allow binding of GTP | 0.31 | 0.014 | 0.27 | 0.044 |
| HNRNPA2B1 | P22626    | Helps package other nascent hnRNPs | 0.67 | 0.008 | 0.72 | 0.002 |
| HNRNPK  | P61978     | Major pre-mRNA-binding protein. Important for response of p53 to DNA damage | 0.45 | 0.008 | 0.54 | 0.008 |
| HSP9OA1 | P07900     | Molecular chaperone | 0.42 | 0.047 | 0.42 | 0.042 |
| HSP9OA1B1 | P08238    | Molecular chaperone | 0.70 | 0.005 | 0.50 | 0.008 |
| KRT2    | P35908     | Keratinocyte activation, proliferation and keratinization. Role in epidermal barrier and terminal cornification | −1.47 | 0.016 | −1.25 | 0.022 |
| KRT6B   | P04259     | Expressed in filiform papillae of tongue, epithelial lining of oral mucosa and oesophagus and outer sheath of hair follicles | −0.43 | 0.021 | −0.52 | 0.014 |
| LCP1    | P13796     | Costimulates activation of T cells with CD3, CD2 and CD28 | 0.63 | 0.024 | 0.69 | 0.026 |
| LUM     | P51884     | Extracellular protein involved in collagen fibril organization and epithelial cell migration and tissue repair | 0.80 | 0.007 | 0.53 | 0.042 |
| MSN     | P26038     | Involved in cytoskeletal structuring. Helps regulate the proliferation, migration and adhesion of lymphoid cells | 0.37 | 0.019 | 0.49 | 0.023 |
| MYL6    | P60665     | Involved in muscle contraction and ATP-dependent actin-based motility | 0.57 | 0.008 | 0.44 | 0.014 |
| NCL     | P19334     | Major constituent of nucleolus in growing cells | 0.57 | 0.015 | 0.37 | 0.044 |
| P4H8    | P07237     | Catalyses formation, breakage and rearrangement of disulfide bonds. Can promote T helper cell 2 migration | 0.73 | < 0.001 | 0.63 | 0.017 |
| PHB2    | Q99623     | Recruits histone deacetylases to mediate transcriptional repression by hormone receptors | 0.63 | 0.02 | 0.67 | 0.017 |
| POSTN   | Q15063     | Secreted EMP associated with epithelial-to-mesenchymal transition. Binds to integrins, activating Akt-PKB and FAK signalling pathways | 1.03 | 0.004 | 0.98 | 0.001 |
| PPLA    | P62937     | Catalyses folding of proteins. Helps induce inflammatory response in the presence of ROS | 0.52 | 0.017 | 0.54 | 0.008 |
| PRDX5   | P30044     | Reduces hydrogen peroxide to water. Helps protect against oxidative stress | 0.66 | 0.048 | 0.68 | 0.023 |
| RPS13   | P62277     | Ribosomal protein. Catalyses protein synthesis | 0.59 | 0.007 | 0.48 | 0.004 |
| RPS20   | P60866     | Ribosomal protein. Catalyses protein synthesis | 0.31 | 0.039 | 0.32 | 0.027 |
| RPS7    | P62081     | Ribosomal protein. Catalyses protein synthesis. Required for rRNA maturation | 0.58 | 0.001 | 0.36 | 0.034 |
| SFQ     | P23246     | Required for pre-mRNA splicing | 0.59 | 0.006 | 0.64 | 0.008 |
| TGFB1   | Q15582     | Involved in cell adhesion and possibly cell−collagen adhesion. Binds several integrins | 0.93 | 0.001 | 1.12 | < 0.001 |

(continued)
Table 2 (continued)

| Gene ID | Uniprot ID | Description* | One-dimensional | Two-dimensional |
|---------|------------|---------------|-----------------|-----------------|
|         |            |               | Fold change<sup>b</sup> | P-value          |
| TKT     | P29401     | Connects glycolysis to pentose phosphate pathway. Important for NADPH production in tissues undergoing biosynthesis | 0.47 | 0.026 |
|         |            |               | 0.52 | 0.004 |
| TNC     | P24821     | EMP that guides migrating neurones and axons during development. Thought to stimulate angiogenesis in cancers | 0.93 | 0.011 |
|         |            |               | 0.85 | 0.009 |
| TUBB    | P07437     | Major constituent of microtubules | 0.43 | 0.01 |
|         |            |               | 0.47 | 0.011 |

ATP, adenosine 5'-triphosphate; EMP, extracellular matrix protein; FAK, focal adhesion kinase; GDP, guanosine 5'-diphosphate; GTP, guanosine 5'-triphosphate; hnRNP, heterogeneous nuclear ribonucleoprotein; NADPH, reduced nicotinamide adenine dinucleotide phosphate; OST, N-oligosaccharyltransferase; PKB, protein kinase B; ROS, reactive oxygen species. *Sourced from Uniprot (www.uniprot.org). <sup>b</sup>Log2 fold change.

Proteomic analysis and ANXA5 and DDOST in SCC, A. Shapanis et al.

Figure 2 Lower CD8<sup>+</sup> and CD1a<sup>+</sup> immune cell frequencies in the primary tumour associate with development of cutaneous squamous cell carcinoma (cSCC) metastases. (a, b) Immunohistochemical staining for CD8 and CD1a revealed significantly more CD8<sup>+</sup> cells and CD1a<sup>+</sup> cells in primary tumours without metastasis (P-NMs) than in primary tumours with metastasis (P-Ms). (c, d) Decreased frequencies of CD8<sup>+</sup> cells and CD1a<sup>+</sup> cells in the tumour or tumoral immune infiltrate are significantly associated with reduced time to metastasis.

Survival analyses were conducted to investigate the relationship between ANXA5 and DDOST expression and clinical outcome. High expression of ANXA5 and DDOST was associated with reduced time to cSCC metastasis (P < 0.001, Figure 5a). P-M cSCCs were associated with a reduced time to death compared with P-NM cSCCs (P < 0.001, Figure 5b), and high expression of ANXA5 and DDOST was also associated with reduced 5-year overall survival (P = 0.024, Figure 5c). Moreover, TCGA analysis demonstrated that high coexpression of ANXA5 and DDOST significantly reduces survival in cervical and oropharyngeal SCC (P = 0.046 and P = 0.0072, respectively; Figure 5d, e).

A stacked ensemble prediction model with the ANXA5 and DDOST MRM data was created using R software and base-level algorithms comprising k nearest neighbours, naive Bayes, glmnet, AdaBoost, xgbDART and the stochastic gradient boosting GBM. The predictions of these individual algorithms were then subjected to a top-layer algorithm, xgbTree, to form final predictions for each sample. Data were split into two-thirds (n = 67) for training and one-third (n = 34) for testing, and models were trained using 10-fold cross-validation repeated three times. The resulting prediction-model receiver operating characteristic curve gave an area under the curve of 0.93 (Figure 5f).

This ANXA5–DDOST prediction model was compared on the same cSCC samples with cSCC clinical staging systems, including American Joint Committee on Cancer 7th and 8th editions, Brigham and Women’s Hospital,9 British Association of Dermatologists, European Dermatology Forum7 and Union for International Cancer Control,24 and with results of the validation study of some of these staging systems by Roscher et al. on their patient cohort.8 This comparison showed that the ANXA5–DDOST prediction model has higher sensitivity and specificity than each of these staging systems.

Discussion

This proteomics-based study identified multiple proteins associated with development of cSCC metastases and ascertained that high levels of expression of several respective genes...
Figure 3  Genes encoding proteins that were significantly differentially expressed between primary tumours with metastasis (P-Ms) and primary tumours without metastasis (P-NMs) were identified as markers of survival in other types of squamous cell carcinoma (SCC). (a, c, e, g, i, k, m) Examples of proteins that were significantly differentially expressed between P-M and P-NM cutaneous SCCs (cSCCs). The median and interquartile range are shown; the Mann–Whitney U-test was used for significance. (b, d, f, h, j, l, n) The Cancer Genome Atlas data demonstrate that expression of genes encoding for relevant proteins have significant effects on survival in cervical, oropharyngeal, oesophageal and lung SCC. High and low expression was defined as above and below the median, respectively. (o) Representative immunohistochemistry stains of L-plastin in P-M and P-NM cSCCs. (p) Immunohistochemical quantification of L-plastin in cSCCs corroborated the proteomic results, identifying significantly more L-plastin-positive cells in P-M than in P-NM tumour groups.
encoding for these proteins associate with reduced survival in SCCs of the cervix, oropharynx, oesophagus and lung. Although mass spectrometry for proteomic analysis of cSCCs has been employed previously, to our knowledge, the current study is the first to investigate differential expression of proteins in primary cSCC with respect to metastasis and clinical outcome. Our topological data analysis was largely able to separate cSCCs according to development of metastases, providing strong support for involvement of the detected proteins in the metastatic process, although it is not possible to

Figure 4 Multiple reaction monitoring (MRM) mass spectroscopy confirms higher ANXA5 and DDOST expression in primary cutaneous squamous cell carcinomas (cSCCs) with metastasis (P-Ms) than primary cSCCs without metastasis (P-NMs). MRM of (a–d) discovery-group cSCCs and (e–h) validation-group cSCCs demonstrated that ANXA5 and DDOST protein levels are increased in P-Ms compared with P-NMs. Data are shown for individual peptides in (a, e) DDOST and (c, g) ANXA5, and mean ± SD for (b, f) DDOST and (d, h) ANXA5.
conclude from this study what proportion of these are drivers or passengers in this process.

Some differences in protein expression between P-M and P-NM cSCCs may be due to variation in tumour parameters (e.g. cell proliferation, differentiation status) or composition of the immune infiltrate between the two tumour groups. However, bioinformatic analysis highlighted several pathways and processes likely to be causally involved in permitting cSCC metastases. STRING/KEGG identified differences between P-Ms and P-NMs in PI3K–Akt signalling, which influences development of cancer metastasis and can affect cSCC growth. Indeed, PI3K–Akt signalling pathways differ between well-differentiated and moderately or poorly differentiated cSCCs and oncogenic mutations affecting PI3K signalling are frequent in metastatic cSCCs.

Figure 5 High ANXA5 and DDOST expression is associated with reduced time to metastasis in cutaneous, cervical and oropharyngeal squamous cell carcinoma (SCC). (a) ANXA5 and DDOST levels have a significant effect on time to cutaneous SCC (cSCC) metastasis. (b, c) Kaplan–Meier plots showing 5-year overall survival for cSCCs based on (b) primary tumour with metastasis (P-M) and primary tumour without metastasis (P-NM) status and (c) expression level of ANXA5 and DDOST. (d, e) Data from The Cancer Genome Atlas signifying that expression of the genes encoding ANXA5 and DDOST has a significant effect on survival in (d) cervical SCC and (e) oropharyngeal SCC. ‘High’ denotes both ANXA5 and DDOST protein abundance or gene expression above the median. (f) Receiver operating characteristic curve of the model produced from multiple reaction monitoring (MRM) data performs better than current guidelines in clinical use; a stacked ensemble model was created using all peptide MRM data as predictors. AJCC, American Joint Committee on Cancer; AUC, area under the curve; BAD, British Association of Dermatologists; BWH, Brigham and Women’s Hospital; EDF, European Dermatology Forum; UICC, Union for International Cancer Control.
STRING/KEGG also identified extracellular matrix–receptor interaction and enrichment of focal adhesion, which are important for cancer invasion and metastases.\textsuperscript{30,31} in P-M compared with P-NM samples. Additionally, STRING/KEGG identified ‘antigen processing and presentation’ differences between P-M and P-NM, consistent with our observations that lower numbers of CD1a\textsuperscript{+} Langerhans cells and CD8\textsuperscript{+} T cells in cSCCs associate with metastasis, and our previous work demonstrating that cSCC Tregs suppress effector T cells in this tumour.\textsuperscript{32} Furthermore, the current study shows that P-Ms have higher levels of transforming growth factor-\(\beta\)1, which exerts immunosuppressive effects via Tregs\textsuperscript{33} and induction of programmed death (PD)-1 on CD8\textsuperscript{+} T cells.\textsuperscript{34}

More proteins were upregulated than downregulated in the comparison of P-M with P-NM cSCCs, which may relate to limitations with mass spectrometry in detecting reduced protein expression below the sensitivity threshold. There were also substantial variations between samples, confirming our previous observations that cSCCs and their immune infiltrates are highly heterogeneous.\textsuperscript{35} In addition, although many proteins that were differentially expressed between P-Ms and P-NMs were identified using both 1D and 2D separation, the 1D and 2D separation methodologies yielded differences in the overall numbers of unique proteins. Moreover, correction for multiple parameters was not feasible given the large number of variables, including varying levels of infiltration of different immunocyte populations. However, we processed cSCC samples including the tumour and surrounding stroma/immune infiltrate, instead of microdissecting the tumour without the stroma, because there is evidence that immune, as well as tumour, parameters are determinants of clinical outcome in cSCC.\textsuperscript{3,4,32,35} We acknowledge there is likely to have been a loss of resolution with this approach, and that future studies undertaking proteomic profiling of cSCCs following purification of separate tumour regions, and deconvolution of data based on heterogeneous cell populations, would allow identification of additional pathways relevant to development of metastases and clinical outcome.

MRM verified differential expression of ANXA5 and DDOST in the discovery group of P-M and P-NM cSCCs and validated this in a separate cohort of tumours, highlighting the relevance of ANXA5 and DDOST in development of cSCC metastasis. However, as both proteins were expressed in tumour and immune cells (Figure S4; see Supporting Information), it is unclear whether the mechanism underlying this association is due to expression of the proteins in the tumour or immune infiltrate, or both these sites. High ANXA5 expression is associated with metastases from colorectal cancer,\textsuperscript{36} and reduced survival in renal cell carcinoma.\textsuperscript{37} Additionally, the Human Protein Atlas indicates that, using TCGA data, ANXA5 is an unfavourable prognostic marker in renal, liver, urothelial, and head and neck cancers, but a favourable marker in endometrial and stomach cancers.\textsuperscript{38} ANXA5 has also been identified as a potential biomarker in a DNA microarray study of cSCC cell lines and tissue,\textsuperscript{39} and in a proteomic analysis of head and neck SCC.\textsuperscript{40}

The mode of action of ANXA5 in relation to development of metastases is not fully understood, but it has been shown to promote migration and invasion of keratinocyte,\textsuperscript{41} oral SCC,\textsuperscript{41} renal cell carcinoma\textsuperscript{37} and hepatocarcinoma\textsuperscript{42} cell lines in ANXA5 knockdown experiments. Potential mechanisms for this include effects of ANXA5 on regulation of genes implicated in cell motility (including S100A4, TIMP3 and RHOC),\textsuperscript{41} activation of PI3K–Akt–mammalian target of rapamycin signalling leading to tumour cell proliferation,\textsuperscript{37} promotion of migration and invasion via upregulation of matrix metalloproteinases 2 and 9,\textsuperscript{37} and effects on integrin signalling and mitogen-activated protein kinase kinase–extracellular signal-regulated kinase pathways.\textsuperscript{42} Conversely, ANXA5 may have a protective role in some cancers because ANXA5 overexpression can inhibit proliferation and metastasis, including in uterine and cervical carcinoma cell lines.\textsuperscript{33} In addition, administration of ANXA5 in a murine model of human papillomavirus 16-associated cancer augmented antitumour immunity by binding to phosphatidylserine externalized by apoptotic tumour cells, which enhanced the immunogenicity of tumour antigens.\textsuperscript{44}

While there is limited published research on DDOST in cancer, the Human Protein Atlas documents DDOST as an unfavourable prognostic marker in renal, liver, and head and neck cancers but as a favourable marker in endometrial cancer.\textsuperscript{45} Gene expression profiling interactive analysis of TCGA, and genome-scale CRISPR-Cas9 knockout screening data have demonstrated DDOST as an essential gene across many cancer cell lines, with DDOST upregulated in colon adenocarcinoma and overlapping with expression of genes required for cell growth and viability (although in that study, higher DDOST expression was associated with increased survival in colon adenocarcinoma).\textsuperscript{46} Furthermore, another study investigating susceptibility variants for oesophageal SCC reported missense variants in DDOST in two cases.\textsuperscript{47} The mechanism whereby DDOST permits metastasis is unclear, but it may involve protein glycosylation and the impact of this via various biological processes relevant to cancer.\textsuperscript{48} For example, DDOST functions as a subunit for an accessory protein required for stabilization of the STT3 protein subunits of oligosaccharyltransferase (OST),\textsuperscript{49,50} which promotes tumour immune evasion via PD ligand 1 (PD-L1).\textsuperscript{51,52} Moreover, STT3, which is induced by epithelial-to-mesenchymal transition, is required for PD-L1 N-glycosylation, which stabilizes and upregulates PD-L1 in breast cancer stem cells.\textsuperscript{53} OST is also required for epidermal growth factor receptor (EGFR) cell-surface localization and signalling in non-small lung cancer cells. Furthermore, in EGFR-driven tumour cells, OST inhibition induces senescence.\textsuperscript{54} Likewise, OST inhibition reduces tumour growth in EGFR-mutant non-small lung cancer\textsuperscript{55} and glioma\textsuperscript{56} xenografts.

The absolute quantification of ANXA5 and DDOST via MRM in primary cSCCs in this study, and confirmation of higher levels of these proteins in P-M tumours in the discovery and validation groups, suggest that they may have potential for use as biomarkers for development of metastasis in cSCC following surgical excision of the tumour. This is supported by our findings that high expression levels of ANXA5 and DDOST...
are associated with shorter time to metastasis and reduced 5-year overall survival in patients with cSCCs, and, similarly, reduced survival in cervical and oropharyngeal SCC. Indeed, the incorporation of our ANXA5 and DDOST MRM data into a prediction model demonstrated higher sensitivity and specificity than commonly used clinical staging systems for cSCC, indicating that ANXA5 and DDOST offer the potential to provide additional useful information on the likelihood of metastatic spread in this cancer.

As MRM was conducted on FFPE cSCC samples in the current study, future evaluation of ANXA5 and DDOST will be possible in larger cohorts of FFPE samples, and in their subsequent study and use in clinical practice as an adjunct to current staging systems that use FFPE samples. Although conjectural, based on evaluation of ANXA5 and DDOST in larger cohorts of patients, the future incorporation of these markers with other relevant clinicopathological risk factors into a prediction model may offer clinical benefits through improved staging and consequently more personalized treatment and/or follow-up of patients with cSCC.

In conclusion, this proteomics study has identified multiple proteins associated with cSCC metastasis, with several of our markers with other relevant clinicopathological risk factors conjectural, based on evaluation of ANXA5 and DDOST in large squamous cell carcinomas. Reduced survival in cervical and oropharyngeal SCC. Indeed, the incorporation of our ANXA5 and DDOST MRM data into a prediction model may offer clinical benefits through improved staging and consequently more personalized treatment and/or follow-up of patients with cSCC.

In conclusion, this proteomics study has identified multiple proteins associated with cSCC metastasis, with several of our markers with other relevant clinicopathological risk factors conjectural, based on evaluation of ANXA5 and DDOST in large squamous cell carcinomas. Reduced survival in cervical and oropharyngeal SCC. Indeed, the incorporation of our ANXA5 and DDOST MRM data into a prediction model may offer clinical benefits through improved staging and consequently more personalized treatment and/or follow-up of patients with cSCC.

References

1. Venables ZC, Nijsten T, Wong KP et al. Epidemiology of basal and cutaneous squamous cell carcinoma in the UK 2013–15: a cohort study. Br J Dermatol 2019; 181:473–81.
2. Nehal KS, Bichakjian CK. Update on keratinocyte carcinomas. N Engl J Med 2018; 379:364–74.
3. Thompson AK, Kelley BF, Prokop LJ et al. Risk factors for cutaneous squamous cell carcinoma recurrence, metastasis, and disease-specific death: a systematic review and meta-analysis. JAMA Dermatol 2016; 152:419–28.
4. Brantsch KD, Meisner C, Schonfisch B et al. Analysis of risk factors determining prognosis of cutaneous squamous-cell carcinoma: a prospective study. Lancet Oncol 2008; 9:711–20.
5. Alam M, Rainer D. Cutaneous squamous-cell carcinoma. N Engl J Med 2001; 344:975–83.
6. Clayton GL, Lee JI, Holsinger FC et al. Mortality risk from squamous cell skin cancer. J Clin Oncol 2005; 23:759–65.
7. Stratigos A, Garbe C, Lebbe C et al. Diagnosis and treatment of invasive squamous cell carcinoma of the skin: European consensus-based interdisciplinary guideline. Eur J Cancer 2015; 51:1989–2007.
8. Roscher I, Falk RS, Vos L et al. Validating 4 staging systems for cutaneous squamous cell carcinoma using population-based data: a nested case–control study. JAMA Dermatol 2018; 154:428–34.
9. Karia PS, Jambusaria-Pahlajani A, Harrington DP et al. Evaluation of American Joint Committee on Cancer, International Union Against Cancer, and Brigham and Women’s Hospital tumor staging for cutaneous squamous cell carcinoma. J Clin Oncol 2014; 32:327–34.
10. Abraham I, Curiel-Lewandrowski C. Staging systems to predict metastatic cutaneous squamous cell carcinoma: unsatisfactory for clinical use, but some less so? JAMA Dermatol 2018; 154:1391–2.
11. Nagarajan P, Asgari MM, Green AC et al. Keratinocyte carcinomas: current concepts and future research priorities. Clin Cancer Res 2019; 25:2379–91.
12. Iglesias-Gato D, Thysell E, Tyanova S et al. The proteome of prostate cancer bone metastasis reveals heterogeneity with prognostic implications. Clin Cancer Res 2018; 24:5433–44.
13. Silva JC, Goreinstein MV, Li GZ et al. Absolute quantification of proteins by LCMS/MS: a virtue of parallel MS acquisition. Mol Cell Proteomics 2006; 5:144–56.
14. MacLean B, Tomazela DM, Shulman N et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 2010; 26:966–8.
15. Weinstein JN, Collisson EA, Mills GB et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013; 45:1113–20.
16. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2013.
17. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag, 2016.
18. Therneau TM, Grambsch, PM. Modeling Survival Data: Extending the Cox Model. New York: Springer, 2000.
19. Wabnitz GH, Kocher T, Lohnes P et al. Costimulation induced phosphorylation of L-plastin facilitates surface transport of the T cell activation molecules CD69 and CD25. Eur J Immunol 2007; 37:649–62.
20. Farasat S, Yu SS, Neel VA et al. A new American Joint Committee on Cancer staging system for cutaneous squamous cell carcinoma: creation and rationale for inclusion of tumor (T) characteristics. J Am Acad Dermatol 2011; 64:1051–9.
21. Lydiatt WM, Patel SG, O’Sullivan B et al. Head and neck cancers – major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 2017; 67:122–37.
22. Moltye RJ, Preston PW, Lawrence CM. Multi-professional guidelines for the management of the patient with primary cutaneous squamous cell carcinoma. Available at: https://www.bad.org.uk/library-media/documents/SCC_2009.pdf (last accessed 12 August 2020).
23. Breuninger H, Brantsch K, Eigentler T et al. Comparison and evaluation of the current staging of cutaneous carcinomas. J Dtsch Dermatol Ges 2012; 10:579–86.
24. Sobin LH, Gospodarowicz MK, Wittekind C. TNM Classification of Malignant Tumours, 7th edn. New York: Wiley-Blackwell, 2009.
25. Azimi A, Yang P, Ali M et al. Data independent acquisition proteomic analysis can discriminate between actinic keratosis, Bowen’s disease, and cutaneous squamous cell carcinoma. J Invest Dermatol 2020; 140:212–22.
26. Pierrobon M, Ramos C, Wong S et al. Enrichment of PI3K-AKT-mTOR pathway activation in hepatic metastases from breast cancer. Clin Cancer Res 2017; 23:4919–28.
27. Zou Y, Ge M, Wang X. Targeting PI3K-AKT-mTOR by LY3023414 inhibits human skin squamous cell carcinoma cell growth in vitro and in vivo. Biochem Biophys Res Commun 2017; 490:385–92.
28. Inman GJ, Wang J, Nagano A et al. The genomic landscape of cutaneous SCC reveals drivers and a novel azathioprine associated mutational signature. Nat Commun 2018; 9:3667.
29. Li YY, Hanna GJ, Laga AC et al. Genomic analysis of metastatic cutaneous squamous cell carcinoma. Clin Cancer Res 2015; 21:1447–56.
Proteomic analysis and ANXA5 and DDOST in SCC, A. Shapanis et al. 721

© 2020 The Authors. British Journal of Dermatology published by John Wiley & Sons Ltd on behalf of British Association of Dermatologists

British Journal of Dermatology (2021) 184, pp709–721

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s website:

Appendix S1 Supplementary materials and methods.

Figure S1 Weighted gene coexpression network analysis reveals clusters of proteins that can be related to clinical and histological characteristics.

Figure S2 STRING analysis with KEGG pathway mapping identified several pathways significantly enriched in both one-dimensional and two-dimensional data.

Figure S3 Multiple reaction monitoring of cutaneous squamous cell carcinoma samples.

Figure S4 Immunostaining of DDOST and ANXA5 shows presence of these proteins in tumour cells and in cells within the surrounding immune infiltrate.

Table S1 List of significantly differentially expressed proteins in one-dimensional proteomics between primary metastatic and nonmetastatic cutaneous squamous cell carcinoma.

Table S2 List of significantly differentially expressed proteins in two-dimensional discovery proteomics data between primary metastatic and nonmetastatic cutaneous squamous cell carcinomas.

Table S3 Medians, interquartile ranges and P-values for comparison of primary metastatic and nonmetastatic cutaneous squamous cell carcinoma groups.