Quantitative proteomic profiling of primary cancer-associated fibroblasts in oesophageal adenocarcinoma

Antigoni Manousopoulou1♯, Annette Hayden2♯, Massimiliano Mellone2, Diana J. Garay-Baquero3, Cory H. White3♯, Fergus Noble2, Monette Lopez4, Gareth J. Thomas2,

Timothy J. Underwood2* and Spiros D. Garbis1,2*

1Institute for Life Sciences, University of Southampton, Southampton, UK; 2Cancer Sciences Unit, Faculty of Medicine, University of Southampton, Southampton, UK;
3Clinical and Experimental Sciences Unit, Faculty of Medicine, University of Southampton, Southampton, UK; 4University Hospital Southampton NHS Foundation Trust, Southampton, UK

♯These authors contributed equally to the study

&Present address: Merck Exploratory Science Center, Cambridge MA USA

*SDG and TJU jointly led the study and are co-corresponding authors

Corresponding authors:
Spiros D. Garbis & Timothy J. Underwood
Cancer Sciences Unit, Faculty of Medicine
University of Southampton
SO16 6YD, Southampton, UK
Email: S.D.Garbis@soton.ac.uk; T.J.Underwood@soton.ac.uk

Running title: CAF proteomic profiling in OAC
**Abstract** (200 words)

Background: Cancer–associated fibroblasts (CAFs) form the major stromal component of the tumour microenvironment (TME). The present study aimed to examine the proteomic profiles of CAFs vs. normal fibroblasts (NOFs) from patients with oesophageal adenocarcinoma to gain insight into their pro-oncogenic phenotype.

Methods: CAFs/NOFs from four patients were sub-cultured and analysed using quantitative proteomics. Differentially expressed proteins (DEPs) were subjected to bioinformatics and compared with published proteomics and transcriptomics datasets.

Results: Principal component analysis of all profiled proteins showed that CAFs had high heterogeneity and clustered separately from NOFs. Bioinformatics interrogation of the DEPs demonstrated inhibition of Adhesion of Epithelial Cells, Adhesion of connective tissue cells and Cell death of Fibroblast Cell Lines in CAFs vs. NOFs (p < 0.0001). KEGG pathway analysis showed a significant enrichment of the insulin-signalling pathway (p = 0.03). Gene ontology terms related with Myofibroblast phenotype, Metabolism, Cell adhesion/migration, Hypoxia/oxidative stress, Angiogenesis, Immune/inflammatory response were enriched in CAFs vs. NOFs. Nestin, a stem-cell marker up-regulated in CAFs vs. NOFs, was confirmed to be expressed in the TME with immunohistochemistry.

Conclusions: The identified pathways and participating proteins may provide novel insight on the tumour-promoting properties of CAFs and unravel novel adjuvant therapeutic targets in the TME.
Introduction

Oesophageal cancer represents a significant global health burden with 395,000 deaths in 2010, an increase of nearly 15% from 1990 (Lozano et al, 2012). Oesophageal adenocarcinoma (OAC) is the predominant histological subtype in Western countries and age-standardised incidence rates are rising by 40% every 5 years (Lepage et al, 2008). The United Kingdom has the highest incidence of OAC in the world, and outcomes are poor because 60–70% of patients present with late-stage disease too advanced for treatment with curative intent (Arnold et al, 2014).

Using whole genome sequencing the OCCAMS consortium has identified new mutational signatures of OAC disease types that might be suitable for targeted treatments (Dulak et al, 2013; Ross-Innes et al, 2015; Weaver et al, 2014). However, findings from the OCCAMS cohorts require pre-clinical validation prior to implementation in trials, and studies are needed to understand the extent to which the genomic distinction is maintained downstream, at the level of the transcriptome and proteome (Secrier et al, 2016). Moreover, although mutationally corrupted cancer cells are recognized as the driving force of tumour development and progression, a key knowledge gap hindering the prediction of which patients will benefit from treatment is that the contribution of the tumour microenvironment (TME) is not considered (Hanahan and Coussens, 2012).

Our group’s work has focused on the relationship between tumour cells and cancer-associated fibroblasts (CAFs), which form the major cellular component of the TME (De Wever et al, 2008). The in vivo “education” or “reprogramming” of fibroblasts by tumour cells is an established mechanism by which cancer cells exploit the plastic nature of reactive cell populations to generate a tumour-supportive microenvironment (Erez et al, 2010). The accumulation of CAFs in tumours correlates with poor prognosis across cancer types, including OAC, where we have shown that the presence of CAFs is more predictive of poor outcome than T, N or M stage (Marsh et al, 2011; Underwood et al, 2015). CAFs
are most commonly characterized by the acquisition of an “activated”, alpha-smooth muscle actin (α-SMA) positive, myofibroblast phenotype (Marsh et al, 2011), which regulates a number of tumour promoting processes (Underwood et al, 2015; Hanley et al, 2016). Additionally, CAFs may be implicated in the development of drug resistance during chemotherapy treatment of cancer patients (Wang et al, 2017; Kalluri et al, 2006). Along these lines, anti-cancer drugs have been found to become ineffective against cancer cells co-cultured with various types of stromal cells (Straussman et al, 2012).

Shotgun proteomics, supported by recent technological advances in mass spectrometry, is gradually becoming an indispensable analytical tool in cancer research since the unbiased protein expression profiling of tumours or their microenvironment can provide novel biological insight but also help identify novel diagnostic, prognostic and therapeutic targets that can eventually influence clinical practice (Larkin et al, 2016; Galanos et al, 2016; Hanley et al, 2016; Bouchal et al, 2015; Zeidan et al, 2015). There are only a limited number of studies that have examined the global proteomic portrait of primary CAFs derived from human cancer patients (Groessl et al, 2014; Fu et al, 2014, Torres et al, 2015).

We have previously reported the shotgun proteomic analysis of primary, patient-matched, CAF/NOF pairs (n=4) from patients with OAC (Hanley et al, 2016). The analysis resulted in the profiling of 3,579 unique proteins, of which 172 were up- and 368 down-regulated in CAFs vs. NOFs. The focus of the study by Hanley et al. was to examine the relative expression levels of extracellular matrix proteins in CAFs vs. NOFs.

The aim of the present study was to apply a more in-depth proteomics methodology in combination with comprehensive bioinformatics analysis to an additional cohort of primary patient-matched CAF/NOF pairs (n=4) derived from patients with OAC in order to gain insight into the pro-oncogenic features of the myofibroblast phenotype. An additional aim was to identify novel therapeutic targets relevant to the TME. An overview of the study workflow is presented in Figure 1.
Materials and Methods

Primary cell culture

Experimental protocols received ethical approval by the Southampton and South West Hampshire Research Ethics Committee (09/H0504/66). All participants signed an informed consent form. Fibroblasts were derived from four patients with oesophageal adenocarcinoma and sub-cultured as previously described (Underwood et al., 2015). Normal fibroblasts (NOFs) were taken from the proximal resection margin (at least 10 cm distant from the cancer) of each patient. Cell culture passage number was consistently under four.

Quantitative proteomics sample processing

Cell pellets were snap frozen at -80 °C. These were dissolved in 0.5 M triethylammonium bicarbonate, 0.05% sodium dodecyl sulphate and subjected to pulsed probe sonication (Misonix, Farmingdale, NY, USA). Lysates were centrifuged (16,000 g, 10 min, 4°C) and supernatants were measured for protein content using infrared spectroscopy (Merck Millipore, Darmstadt, Germany). Lysates were then reduced, alkylated and subjected to trypsin proteolysis. Peptides were labelled using the eight-plex iTRAQ reagent kit (113=NOF patient 1, 114=NOF patient 2, 115= NOF patient 3, 116= NOF patient 4, 117= CAF patient 1, 118= CAF patient 2, 119= CAF patient 3, 121= CAF patient 4) and analysed using multi-dimensional liquid chromatography and tandem mass spectrometry.

Two-dimensional LC-MS proteomic analysis

To enhance peptide separation efficiency and subsequent mass spectrometry analysis, the initial offline peptide fractionation was based on alkaline C4 Reverse Phase
(RP) chromatography (Kromasil 150 x 2.1 mm, 3.5 μm particle, 100Å pore size, Merck KGaA, Darmstadt, Germany) using gradient mobile phase conditions as previously reported by the authors (Manousopoulou et al, 2017). All other method details were as reported by the authors (Manousopoulou et al, 2016; Manousopoulou et al, 2017).

Database searching

Unprocessed raw files were submitted to Proteome Discoverer 1.4 for target decoy searching against the UniProtKB homo sapiens database comprised of 20,159 entries (release date January 2015), allowing for up to two missed cleavages, a precursor mass tolerance of 10ppm, a minimum peptide length of six and a maximum of two variable (one equal) modifications of; iTRAQ 8-plex (Y), oxidation (M), deamidation (N, Q), or phosphorylation (S, T, Y). Methylthio (C) and iTRAQ (K, Y and N-terminus) were set as fixed modifications. FDR at the peptide level was set at <0.05. Percent co-isolation excluding peptides from quantitation was set at 50. Reporter ion ratios from unique peptides only were taken into consideration for the quantitation of the respective protein. Raw iTRAQ intensity values of unique peptides were median-normalized and log2 transformed. A Student’s T-Test using the normalised raw iTRAQ intensity was performed to identify differentially expressed unique peptides between CAFs and NOFs. Significance was set at p ≤ 0.05. A protein was considered to be differentially expressed in CAFs vs. NOFs when it had at least one differentially expressed unique peptide and a mean iTRAQ log2ratio of ≥ ± 0.2. In adherence to the Paris Publication Guidelines for the analysis and documentation of peptide and protein identifications (http://www.mcponline.org/site/misc/ParisReport_Final.xhtml), only proteins identified with at least two unique peptides were further subjected to bioinformatics. All mass spectrometry data have been deposited to the ProteomeXchange Consortium via PRIDE with the dataset identifier PXD005444.
Bioinformatics analysis

Principal component analysis using the log\textsubscript{2} ratio of each sample over the mean of all samples was performed using the online tool ClustVis (http://biit.cs.ut.ee/clustvis/). DAVID (https://david.ncifcrf.gov/) was applied to differentially expressed proteins in order to identify over-represented gene ontology terms and KEGG pathways. Fisher exact corrected p-values $\leq$ 0.05 were considered significant. Subcellular localization of top up- and down-regulated proteins in CAF vs. NOF was manually assessed using ExPASy (www.expasy.org). The diseases and functions module of Ingenuity Pathway Analysis (IPA) (Qiagen, Hilden, Germany) was used to predict upstream biological processes activated or inhibited based on a combination of up- and down-regulated proteins observed. Biological processes with a Fisher’s exact p-value < 0.05 and a false discovery rate score (z-score) of $\geq$ 2 or $\leq$ -2 were considered significantly activated or inhibited respectively (Krämer et al, 2014; Al-Daghri et al, 2014).

Comparison of DEPs with published proteomics and transcriptomics datasets

DEPs were compared with our previously published proteomics dataset of primary CAFs/NOFs from patients with OAC (n=4) (Hanley et al, 2016). To define DEPs in this previous dataset, the exact same criteria as described above for the present study were used. Common DEPs in the two proteomics experiments were compared with a publically available transcriptomics dataset of laser-capture micro-dissected oesophageal stroma (n=44; 17 with intestinal metaplasia, 16 with dysplasia and 11 with adenocarcinoma) (NCBI/NIH; GEO; dataset ID: GSE19632).

In silico evaluation of the prognostic value of DEPs in OAC

Proteins identified to be differentially expressed in CAFs vs. NOFs in both proteomics experiments were in silico evaluated for their prognostic value in OAC using PrognoScan (http://www.abren.net/PrognoScan/), a database of published cancer microarray experiments linking gene expression to patient prognosis (Mizuno et al, 2009).
**Immunohistochemical validation of key findings**

Immunohistochemical staining was performed in sections derived from a cohort of 183 OAC patients as previously described (Underwood *et al.*, 2015). Briefly, sections of thickness 5μm were taken from the recipient paraffin block for IHC staining. Primary antibody dilution for polyclonal rabbit anti-human Nestin was 1:100 (DAKO no. M3515). Slides were de-paraffinized with xylene and rehydrated with alcohol. Incubation in 3% H₂O₂ (in deionized water) for 10 minutes was used to suppress endogenous peroxidase activity. Slides were incubated in 1mM ethylenediaminetetraacetic acid (EDTA) for 15 minutes at 98°C and pH 8.0, allowing antigen retrieval. Tissue was sequentially incubated in avidin, biotin, primary and biotinylated secondary antibody (at appropriate dilutions), streptavidin biotin-peroxidase complexes and DAB (3-3′-diaminobenzidine). Cores were counter-stained with Mayers Haematoxylin, dehydrated and mounted with DPX. The automated immunostainer DAKO® Autostainer Link 48 (Cambridge, UK) was used in a CPA-accredited cellular pathology department with the use of antibodies optimised to national diagnostic standards (NEQAS).

**Results**

**Proteomic profiling of primary oesophageal fibroblasts**

We compared the global proteomic profiles of matched pairs of primary CAFs and NOFs taken from oesophageal resections of four OAC patients in order to identify proteins and pathways that may be responsible for the pro-oncogenic CAF phenotype and the poor patient prognosis associated with the accumulation of CAFs in OAC. Proteomic analysis resulted in the profiling of 7,718 unique protein groups (peptide FDR p-value<0.05) ([Supplementary Table 1](#)), a substantial improvement of more than double the number of profiled unique proteins compared to our previously published proteomics dataset. Principal component analysis of all profiled proteins demonstrated the more homogeneous
NOF phenotype clustered separately from that of the more heterogeneous CAFs (Figure 2A).

The differentially expressed proteome (DEP) comprised 699 up- and 987 down-regulated proteins in CAFs compared to NOFs (Supplementary Table 2). A volcano plot representation of the mean iTRAQ log2 ratio of proteins in CAF vs. NOF plotted against the minus log10(p-value) is presented in Figure 2B. Alpha-SMA expression was found to be variable but with a mean log2 ratio of 0.2 ± 0.9 (p-value<0.0001 at the peptide level) across all CAFs vs. NOFs examined (Figure 2C).

Comparison of DEPs with published proteomics and transcriptomics datasets

Of the DEPs, 136 proteins were also identified with the same trend of modulation in our previously published proteomic analysis of an independent primary CAF/NOF cohort from patients with OAC (Hanley et al, 2016) and the expression trend of five up- and 11 down-regulated proteins was confirmed in the publically available microarray dataset of OAC micro-dissected stromal cells. These proteins are presented in heatmap format in Figure 2D. Proteins identified in both proteomic experiments and confirmed with the same trend of modulation in the microarray dataset are highlighted in grey (Figure 2D). Among the proteins identified in both proteomics and confirmed at the transcriptomics dataset to be up-regulated in CAFs vs. NOFs were α-smooth muscle actin, lamin A (LMNA) and actin-1 (ACTN1).

Bioinformatics Analysis

The diseases and functions module of IPA predicted, based on the downstream up- and down-regulated proteins, that Adhesion of Epithelial Cells (z-score = - 2.4 | p=6.3E-06), Adhesion of connective tissue cells (z-score = - 2.3 | p=1.8E-05) and Cell death of Fibroblast Cell Lines (z-score = - 2.2 | p=1.7E-09) were significantly inhibited in CAFs vs. NOFs (Figure 3A). KEGG pathway analysis using DAVID showed a significant enrichment of the insulin-signalling pathway (Fisher exact p-value= 0.03 for the common
proteins between the two proteomics experiments and 0.05 for the DEPs analysed in the present study) (Figure 3B).

DAVID gene ontology analysis, accounting for both up- and down-regulated proteins constituting the DEP, demonstrated that processes related with Myofibroblast phenotype, Metabolism, Cell adhesion/migration, Hypoxia/oxidative stress, Angiogenesis, and Immune/inflammatory response were over-represented (Figure 4A). The top 10 up- and down-regulated proteins mapping to each GO term group are presented in heatmap format in Figure 4B. The sub-cellular localization of these proteins is also presented in the heatmap. Top up-regulated proteins that are either secreted or localized in the membrane are highlighted in the heatmap as potential therapeutic targets in CAF (gene names of the respective proteins are: CD9, MIF, HMGB2, HMGB1, CSPG4, CACNB3, APC, BCAM, CD97, LPP, LCT, TJP2, PLCD3, SLC9A3R1, CAV1, RAPGEF2, MAP3K7 and CD44) (Figure 4B).

In silico evaluation of the prognostic value of Nestin in OAC

Using the in silico PrognoScan meta-analysis microarray database for the common DEPs in both proteomics experiments, increased levels of nestin was found to be associated with poor OAC patient prognosis [COX p-value=0.003; HR (95% CI) = 78.0 (4.3 to 1409.8)] (Figure 5A). Immunohistochemical staining of nestin was performed in a well-described cohort of 183 oesophageal tumours where the presence of α-SMA positive CAFs correlated strongly with poor overall survival (Underwood et al, 2015). The patient clinico-pathological characteristics of this cohort have been reported before (Underwood et al, 2015). Nestin showed a conserved expression pattern in the tumour microenvironment of OAC, being confined to CAF, blood vessels and smooth muscle cells. Example staining is shown in Figure 5B.

Discussion
The seminal work of Paget over a century ago proposed that cancer cells constitute the “seeds” that colonize a favourable stromal microenvironment as the receptive “soil” (Han et al, 2015; Baulida et al, 2017). A key “soil” constituent is the normal fibroblast that acquires a cancerous phenotype by the “seed” cancer cell to facilitate its proliferation, invasion, and metastasis (Martin et al, 2011). However, the proteomic characterization of such cancer fibroblasts remains limited.

To address this need, our study made use of a comprehensive quantitative proteomics approach (Figure 1) and reports the most extensive proteome coverage to date of primary CAF/NOF pairs from patients with OAC, a considerable improvement in proteome coverage compared to our published CAF/NOF dataset. Principal Component Analysis (PCA) against the reporter ion ratios of the 7,718 unique protein analysed across all samples resulted in distinct molecular clusters for CAFs relative to NOFs (Figure 2A). In keeping with previous findings (Ishii et al, 2016; Underwood et al, 2015), PCA analysis showed marked heterogeneity in proteome expression for CAF relative to the more homogeneous proteome expression for NOF. Significant differential expression was observed for 699 up- and 987 down-regulated proteins across all CAFs relative to all NOFs, as highlighted in the volcano plot of Figure 2B (log2ratio ≥ 0.2, p ≤ 0.05, t-test). Alpha-SMA (ACTA2) was analysed to be marginally up-regulated in CAFs vs. NOFs (as illustrated in the volcano plot of Figure 2B) (log2ratio = 0.2 ±0.9; p-value<0.0001 at the peptide level) (Figure 2C). By contrast, our quantitative proteome revealed a large spectrum of novel proteins exhibiting a higher and more consistent level of differential expression that may constitute more robust candidate markers of the CAF phenotype (Figure 2B and 2C, Supplementary Table 2). Consistent protein differential expression of CAF canonical markers was observed between the current quantitative proteome, a proteomics dataset reported by the authors (Hanley et al, 2016) and a publically available transcriptomics microarray dataset (Figure 2D). Notable surrogate markers of the CAF phenotype include the up-regulated proteins lamin A (LMNA) and actin-1 (ACTN1). LMNA has been implicated in the modulation of TGF-β1 on collagen production and
mesenchymal differentiation (Van Berlo et al, 2005), and ACTN1 up-regulation has been described in stromal fibroblasts derived from oral cancers (Chatzistamou et al, 2011).

The diseases and functions module of IPA predicted the inhibition of Adhesion of Epithelial Cells (z-score = - 2.4 | p=6.3E-06) and Adhesion of connective tissue cells (z-score = - 2.3 | p=1.8E-05) (Figure 3A). The inhibition of these processes suggests the involvement of CAFs on increasing the tumour’s metastatic potential. These findings confirm and extend the current knowledge of the CAF phenotype also affecting cell adhesion/cell migration processes (Underwood et al, 2015; Hanley et al, 2016).

Of relevance, given the endoergic character of increased cellular proliferation and pro-metastatic phenotypes observed, the insulin-signalling pathway was significantly enriched in the DEP of the present study as well as the common proteins with our previously published proteomics dataset (Figure 3B). Increased expression of the insulin-like growth factor 1 (IGF-I) and its receptor (IGF-IR) has been found to be associated with tumour progression and poor prognosis in different cancer types including gastrointestinal tumours (Giovannucci et al, 2001; Woodson et al, 2004). The tumour promoting properties of IGF-IR are dependent on the activation of the down-stream insulin receptor substrates (IRS) (Chan et al, 2008; Ramocki et al, 2008). IGF-I also plays a key role in the autocrine and paracrine induction of CAF “activation” (Kalluri et al, 2006). A recent study showed that NT157, an inhibitor of the IGF-IR-IRS signalling pathway, resulted in inhibition of CAF “activation”, as well as reduced expression of pro-oncogenic chemokines, cytokines and growth factors, including several interleukins (IL-6, IL-11, IL-23) and TGFβ (Sanchez-Lopez et al, 2016). The de-regulation of the insulin signalling pathway in CAFs could also be linked to the “Reverse Warburg effect”, a model describing the metabolic coupling between stromal and cancer cells (Pavlides et al, 2009). One interesting protein mapping in the insulin-signalling pathway was hexokinase-1 (HK1). HK1 was consistently upregulated in both proteomic experiments and further confirmed at the microarray dataset (Figures 2D and 3B). HK1 catalyzes the first obligatory and rate-limiting step of glucose metabolism, more specifically the phosphorylation of glucose to G6P (Smith, 2000).
Furthermore, HK1 has been suggested to regulate cell death, a process associated with abnormal proliferation and tumorigenesis (Pastorino and Hoek, 2003). HK1 has also been found to be upregulated in different cancer types, including kidney and breast carcinomas (Hooft et al, 2005; Millon et al, 2011). Furthermore, a recent study showed that HK1 over-expression is associated with poor patient prognosis in colorectal cancer (He et al, 2016). HK1 expression in CAFs and its implication with tumour aggressiveness warrants further investigation.

DAVID GO analysis identified terms related with *Myofibroblast phenotype*, *Metabolism*, *Cell adhesion/migration*, *Hypoxia/oxidative stress* (including *DNA damage response*), *Angiogenesis*, and *Immune/inflammatory response* processes to be over-represented in the DEPs (Figures 4A). The gene names of the top-10 differentially expressed proteins observed for each these processes, including those classified as secreted or membrane associated, constitute novel observations and may reveal candidate therapeutic targets (Figure 4B).

Hypoxia, oxidative stress and DNA damage response were significantly enriched GO terms. Oxidation-reduction is an established process in CAFs (Balliet et al, 2011; Martin et al, 2011). CAFs have been shown to overproduce reactive oxygen species (ROS), leading to oxidative stress, inflammation and significant cellular damage, which could in turn affect DNA damage response (Martin et al, 2011; Trinchieri, 2012). The over-production of ROS by CAFs can induce oxidative stress in normal fibroblasts that further triggers CAF activation, thus leading to a positive feedback loop between ROS production and CAF activation (Jeziersdka-Drutel et al, 2013; Chan et al, 2017). Furthermore, the *Cell death of Fibroblast Cell Line* process was found to be inhibited (z-score = -2.2 | p=1.7E-09), showing that CAFs may evade apoptosis possibly as a result of their enhanced DNA damage response.

Immune and inflammatory responses were also significantly over-represented terms in CAFs vs. NOFs (Figure 4A). Previous studies have reported on the immunomodulatory effects of CAF (Takahashi et al, 2017; Wen et al, 2017; Yeh et al,
Specific pathways and their participatory proteins responsible for the interplay between CAF and the host immune response may be of relevance to a number of current clinical trials using immune checkpoint inhibitors in unselected patients with OAC. The success rate of these therapies may not depend entirely on the immune system, but also implicate CAF-induced alterations of the tumour microenvironment in preventing immune cell entry. This may warrant the combined use of immunotherapy and CAF permeability modifiers (Hanley et al., 2018). At the same token, CAFs have been reported to promote angiogenesis through different mechanisms, including ECM remodelling, recruitment of epithelial progenitor cells, and increased leucocyte infiltration through chemokine secretion, that in turn produce angiogenic factors (Tao et al., 2017).

An up-regulated protein identified in both proteomics experiments was nestin. Nestin was further investigated as it was found to correlate with decreased overall survival in patients with oesophageal cancer based on the in silico microarray meta-analysis tool PrognoScan (Figure 5A), suggesting that nestin may play an important role in OAC biology. Nestin is an intermediate filament protein originally detected in neuronal stem cells during development (Lendahl et al., 1990). Nestin has been detected in various types of solid tumours, including mesenchymal tumours and cancers (e.g. breast, lung, ovarian and gastrointestinal) (Ishiwata et al., 2011). Nestin has been suggested as a stem-cell marker indicating an undifferentiated and thus more invasive phenotype of transformed cells (Neradil et al., 2015). Immunohistochemical staining showed that nestin protein expression was confined to the tumour microenvironment of OAC (Figure 5B). A recent study showed that nestin suppression reduced the metastatic potential of endometrial cancer cells by inhibiting the TGFβ signalling cascade (Bokhari et al., 2016), the main pathway promoting aberrant CAF “activation” (Hawinkels et al., 2014).

The main study limitation is that only four matched pairs of fibroblasts were used to generate the proteomic expression profiles. This is partly compensated, however, by the evaluation of the analysed proteins using our previously published proteomics dataset (n=4) and an independent microarray dataset (n=44).
In conclusion, this study reports the proteomic profiling of primary CAFs from patients with OAC, a cancer with a vast unmet clinical need. The biological pathways and networks observed for the primary CAFs examined were found to emulate all the intrinsic hallmarks of cancer, as expected given the strong functional cross-talk between fibroblasts and cancer cells. Furthermore, the participating proteins to these biological processes may constitute novel adjuvant therapeutic targets for OAC in the TME as part of precision medicine protocols.

References


Lendahl U, Zimmerman LB, McKay RD. CNS stem cells express a new class of intermediate filament protein. *Cell* 1990; **60**: 585–95


Declaration: The authors have no conflict of interest to declare

Funding: Wessex Cancer Trust, CRUK – Southampton Internal Pilot Grant, EU-FP7 Marie Curie (CANOMICS), Annual Adventures in Research – University of Southampton, EU-Excellence II – Systems Biology Framework FRA-SYS (Grant 4072).

Authors’ contributions: A.M. and A.H. designed study, performed experiments, interpreted results and wrote manuscript; M.M. interpreted results and edited manuscript; D.J.G.B., provided analytic tools; C.H.W. performed biostatistical analysis; F.N. and M.L. sample procurement; G.J.T. interpreted results and edited manuscript, T.J.U. and S.D.G. raised funding, designed study, interpreted results and wrote manuscript.

Acknowledgements: We are indebted to Mr. Roger Allsopp, Mr. Derek Coates and Hope for Guernsey for establishing the clinical mass spectrometry infrastructure at the University of Southampton. The authors are grateful to the support of King Saud University, Deanship of Scientific Research Chair, Prince Mutaib Bin Abdullah Chair for Biomarkers of Osteoporosis, College of Science, as well as the Visiting Professor Program of King Saud University, Riyadh, Saudi Arabia.
Figure Legends

Figure 1. Study workflow

Figure 2. (A) Principal component analysis using the reporter ion log2ratios of all analysed proteins showed that CAF have a distinct proteomic profile and higher heterogeneity compared to NOF. (B) Volcano plot highlighting the differentially expressed proteins in CAF vs. NOF (red=up-regulated proteins; green=down-regulated proteins). (C) Alpha smooth muscle actin (ACTA2) was found to be significantly up-regulated in CAF vs. NOF (Mean log2ratio (SD) = 0.2 (0.9); p-value < 0.0001 at the peptide level) (D) In total, 136 DEPs were also analysed with the same trend of modulation in a previously published proteomics dataset of primary CAF/NOF cells from patients with OAC. Of these five up-and 11 down-regulated proteins were confirmed in the microarray dataset (highlighted in grey).

Figure 3. (A) The diseases and functions module of IPA predicted the significant inhibition of Adhesion of Epithelial Cells (z-score = - 2.4 | p=6.3E-06), Adhesion of connective tissue cells (z-score = - 2.3 | p=1.8E-05) and Cell death of Fibroblast Cell Lines (z-score = - 2.2 | p=1.7E-09) in CAFs vs. NOFs. (B) KEGG pathway analysis using DAVID showed a significant enrichment of the insulin-signalling pathway (Fisher exact p-value= 0.03 for the common proteins between the two proteomics experiments and 0.05 for the DEPs analysed in the present study).

Figure 4. (A) DAVID gene ontology analysis showed that gene ontology terms related with Myofibroblast phenotype, Metabolism, Cell adhesion/migration, Hypoxia/oxidative stress, Angiogenesis, Immune/inflammatory response were significantly over-represented in the DEP. (B) Heatmap of top 10 up- and top 10 down-regulated proteins mapping to each gene ontology terms category. The subcellular location of each protein is also presented.
and up-regulated proteins that are either secreted or membrane are highlighted as potential therapeutic targets.

Figure 5. (A) Using the in silico PrognoScan meta-analysis microarray database, higher expression levels of nestin was found to correlate with poor patient prognosis [COX p-value=0.003; HR (95% CI) = 78.0 (4.3 to 1409.8)]. (B) Immunohistochemical staining of nestin in OAC showed a conserved expression pattern in the tumour microenvironment, with expression being confined to CAF, blood vessels and smooth muscle cells.
Sub-culture of primary CAF/NOF cells from patients with oesophageal adenocarcinoma (OAC) (n=4)

**Differentially expressed proteins** (DEPs) in CAF vs. NOF

Quantitative in-depth proteomic analysis

Biostatistics

Bioinformatics

Enriched biological processes

Compare with published proteomics dataset of primary CAF/NOF cells from patients with OAC (n=4)

Enriched biological processes

Bioinformatics

Identify common DEPs in CAF vs. NOF

Compare with publically available transcriptomics dataset of micro-dissected oesophageal stroma (n=44)
**Figure 2**

### Down-regulated proteins
- MARS
- ANXAS
- MAMDC2
- CEBPz
- TSPAN3
- FN1
- TNND5C
- PURQ
- CYP1B1
- ORC2
- COL5A3
- UBE2T
- SERPINB2

### Up-regulated proteins
- PBXIP1
- SURF1
- IKBKE
- GSN
- FAM181A
- SLC14A1
- TGM1
- PAX1
- HMG2
- HIST1H2AH

### Mean iTRAQ log2ratio [CAF vs. NOF]
- p=0.05

### α-SMA (ACTA2)

### DEPs in CAF vs. NOF confirmed in both proteomics and one microarray datasets

#### Table: iTRAQ log2ratio of unique peptides

<table>
<thead>
<tr>
<th>Gene</th>
<th>Subcellular location</th>
<th>Experiment A</th>
<th>Experiment B</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTA2</td>
<td>Membrane</td>
<td>-1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>ACTNB</td>
<td>Membrane</td>
<td>-1.0</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Abbreviations**
- EC = Extracellular region
- G = Golgi apparatus
- M = Membrane
- L = Lysoosome
- C = Cytosol
- E = Endosome
- CS = Cytoskeleton
- MI = Mitochondrion
- ER = Endoplasmic reticulum
- N = Nucleus
- P = Peroxisome

#### Proteomics Microarray
- Gene name
- Subcellular location
- Experiment A
- Experiment B

#### Proteomics Microarray
- Gene name
- Subcellular location
- Experiment A
- Experiment B

---

**Legend**
- Up-regulated
- Down-regulated
Fisher exact p-value = 0.05 in DEPs of present study and 0.03 in common DEPs with published proteomics dataset

Key

<table>
<thead>
<tr>
<th></th>
<th>up-regulated in CAF vs. NOF</th>
<th>down-regulated in CAF vs. NOF</th>
<th>Predicted inhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td></td>
<td>Green</td>
<td>Blue</td>
</tr>
<tr>
<td>Yellow</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Proteins analyzed in present study

- Adhesion of epithelial cells
  - z-score = -2.4, p-value = 6.3E-06
- Adhesion of connective tissue cells
  - z-score = -2.3, p-value = 1.8E-05
- Cell death of fibroblast cell lines
  - z-score = -2.2, p-value = 1.7E-09
Gene ontology enrichment analysis of DEPs in CAF vs. NOF

- **GO term groups** (colour key):
  - Myofibroblast phenotype
  - Hypoxia/oxidative stress
  - Metabolism
  - Cell adhesion/migration
  - Angiogenesis
  - Immune/inflammatory response

- **Potential therapeutic targets**

**B**

- **Up-regulated**
  - CAF vs. NOF GO term groups
  - Gene Name
  - Subcellular location
  - Ratio [analysed vs. mapped proteins in GO term]

- **Down-regulated**
  - CAF vs. NOF GO term groups
  - Gene Name
  - Subcellular location
  - Ratio [analysed vs. mapped proteins in GO term]

**Abbreviations**
- EC = Extracellular region
- M = Membrane
- C = Cytosol
- CS = Cytoskeleton
- ER = Endoplasmic reticulum
- N = Nucleus
Figure 5

Nestin

A Kaplan–Meier plot

B

Oesophageal adenocarcinoma

Figure 5