Oncogene (2008) 27, 323–331

www.nature.com/onc

ORIGINAL ARTICLE

Signal pathway profiling of epithelial and stromal compartments of colonic carcinoma reveals epithelial-mesenchymal transition

KM Sheehan[1,2,7], C Gulmann[1,2,7], GS Eichler[3], JN Weinstein[3], HL Barrett[2], EW Kay[2], RM Conroy[4], LA Liotta[1,5] and EF Petricoin III[5,6]

[1]NCI-FDA Clinical Proteomics Program, Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [2]Department of Pathology, Royal College of Surgeons in Ireland and Beaumont Hospital, Dublin, Ireland; [3]Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [4]Department of Epidemiology, Royal College of Surgeons in Ireland, Dublin, Ireland; [5]Center for Applied Proteomics and Molecular Medicine, George Mason University, Manassas, VA, USA and [6]NCI-FDA Clinical Proteomics Program, Food and Drug Administration, Bethesda, MD, USA

Molecular crosstalk, including reciprocal stimulation, is theorized to take place between epithelial cancer cells and surrounding non-neoplastic stromal cells. This is the rationale for stromal therapy, which could eliminate support of a cancer by its genetically stable stroma. Epithelial-stromal crosstalk is so far poorly documented *in vivo*, and cell cultures and animal experiments may not provide accurate models. The current study details stromal-epithelial signalling pathways in 35 human colon cancers, and compares them with matched normal tissues using quantitative proteomic microarrays. Lysates prepared from separately microdissected epithelium and stroma were analysed using antibodies against 61 cell signalling proteins, most of which recognize activated phosphoisoforms. Analyses using unsupervised and supervised statistical methods suggest that cell signalling pathway profiles in stroma and epithelium appear more similar to each other in tumours than in normal colon. This supports the concept that coordinated crosstalk occurs between epithelium and stroma in cancer and suggests epithelial-mesenchymal transition. Furthermore, the data herein suggest that it is driven by cell proliferation pathways and that, specifically, several key molecules within the mitogen-activated protein kinase pathway may play an important role. Given recent findings of epithelial-mesenchymal transition in therapy-resistant tumour epithelium, these findings could have therapeutic implications for colon cancer.

Oncogene (2008) **27,** 323–331; doi:10.1038/sj.onc.1210647; published online 9 July 2007

Keywords: proteomics; protein microarray; colon cancer; stroma; microdissection

Correspondence: Dr C Gulmann, Department of Pathology, Beaumont Hospital, Dublin 9, Ireland.
E-mail: christian_gulmann@hotmail.com
[7]These authors contributed equally to this study and are to be considered joint first authors.
Received 15 March 2007; revised 3 May 2007; accepted 6 June 2007; published online 9 July 2007

Introduction

Cancer cells can alter adjacent stroma to support their progression (Liotta and Kohn, 2001; Matrisian *et al*., 2001). Tumour desmoplasia, for example, is a characteristic feature of carcinomas, and cancer cells produce multiple growth factors to induce angiogenesis (Bergers and Benjamin, 2003), inflammation (Coussens and Werb, 2002) and fibrosis (DeWever and Mareel, 2003). There is also evidence of reciprocal signalling that tumour stroma also provides oncogenic signals to the epithelial component (Tlsty and Hein, 2001).

There is increasing awareness that interactions between tumour cells and their microenvironment play a significant role in the progression of many cancers (Thiery, 2002; DeWever and Mareel, 2003). Tumour epithelium is supported by specialized stroma in which inflammatory cells, fibroblasts, blood vessels and extracellular matrix proteins establish complex communication networks to collaborate towards invasion and metastasis (Tlsty and Hein, 2001; Bhowmick *et al*., 2004). This relationship is mediated by numerous signals including growth factors, chemokines and cytokines (Littlepage and Werb, 2005) that transmit the neoplastic messages through interconnected cell signalling cascades both inside and outside the affected tumour epithelium. Thus, identification of the aberrant crosstalk between tumour epithelium and stroma may provide a novel therapeutic targets for inhibition of tumour progression (Joyce, 2005).

The interaction between the tumour and the cellular stroma can be bidirectional and stroma may inhibit the tumour by releasing inhibitory molecules, or, alternatively, may provide the necessary pro-stimulatory response to support the tumour's progression (Brigati *et al*., 2002). Assessment of the balance between immunosuppressive and cytotoxic activity in the tumour is complicated by the mixed population microecology of stromal cells within tumours, some of which may also exist in the adjacent normal tissues. Therefore, to investigate fully the alterations and molecular crosstalk between tumour epithelium and its microenvironment, it

is useful to make direct comparisons with uninvolved tissues from which the tumour has arisen.

Several apparent relationships among cellular components of tumours have emerged from *in vitro* experiments using co-culture models with dual-cell populations. Those models allow for exchange of soluble mediators through the culture medium (Gallagher *et al*., 2005). Transplantation models that combine *in vitro* and *in vivo* components have been used to study the kinetics of tumour–stroma interactions during different stages of malignancy (Mueller and Fusenig, 2002). Those models provide valuable information, but they do not indicate what is different from the normal interaction between epithelial and stromal cells. Furthermore, they cannot accurately reflect the characteristics of reciprocal signalling within clinical tumours, which show patient-specific variations in the extent of ischaemia, inflammation and immune response.

High-throughput genomic technologies such as oligonucleotide or cDNA arrays have shown signatures characteristic of colon cancers, as opposed to normal colonic tissue (Brown and Botstein, 1999; Notterman *et al*., 2001). Gene expression profiling, however, cannot assess the activation status of signalling pathways, which depend on post-translational events including phosphorylation, cleavage and proteasomal degradation of involved proteins. Characterization of kinase-based cell signalling events, provide a unique opportunity for new drug target identification since these kinases represent a majority of targeted inhibitors today. Recent research has revealed that gene expression has only partial concordance with protein expression (mostly with more highly abundant structural proteins), and is unable to accurately predict the state of protein signalling pathways driven by post-translational modifications (Nishizuka *et al*., 2003).

Most gene mutations found in colon cancer affect signal transduction pathways (Zhang *et al*., 1997). Those pathways in turn are dependant on activation, which typically occurs through protein phosphorylation. Since a large number of potential drugs currently in the pipeline for cancer target kinases, the 'phosphoproteome' contains essential information, not just about the functional biology, but also about potential drug targets.

To elucidate the complex tumour–stroma interactions within a human cancer type, we used a novel protein microarray platform to analyse activation (that is, the phosphorylation status) of key signalling pathways in colorectal cancers and patient-matched, uninvolved colonic mucosa taken at the same time at surgery. To facilitate comparative analyses of different tissue compartments, epithelium and stroma of the tumour and uninvolved tissues were laser capture microdissected to isolate the candidate protein expression profiles, providing a unique 'field effect' study. We then analysed the status of ongoing phosphorylation-driven cellular signalling of four separate patient-matched compartments: tumour epithelium (TE), stroma within the tumour (TS), normal-appearing epithelium from unaffected adjacent areas of tumours (NE) and stroma next to unaffected epithelium (NS).

Results

Patient demographics and clinicopathological parameters
Patient-matched invasive carcinoma and normal colonic mucosa were obtained from 35 patients with a diagnosis of colorectal cancer. There were 16 females and 19 males. The mean patient age was 69 years (range 52–89 years). Seven cases were Dukes' stage A, fourteen Dukes' B, twelve Dukes' C and two Dukes' D. Comprehensive additional clinicopathological information was available for all cases (Table 1).

Proteomic profiling of separate tissue compartments
Reduced molecular distinction between epithelial and stromal cells in tumour compared to normal tissue. Analysis of multiple different kinase substrates revealed a striking degree of patient heterogeneity in the activity of multiple signalling cascades within each patient, between different patients and between epithelium and stroma. Unsupervised hierarchical clustering of the multiplex of signalling endpoints within a 'normal epithelium–normal stroma' comparison and 'tumour epithelium–tumour stroma' comparison shows two distinct heat maps (Figures 1a and b). Epithelium and stroma derived from normal-appearing colonic mucosa were well separated based on their phosphoproteomic signatures (Figure 1a). Epithelium and stroma derived from tumour are less well distinguished by this unsupervised technique (Figure 1b). Although suboptimal branch reordering could account for some of that difference, the dendrogram suggests that the stroma and epithelium within the tumour are less distinct from each other than are the normal stroma and normal-appearing epithelium obtained from matched colonic mucosa.

The reduction of tumour epithelial-stromal distinction by unsupervised hierarchical clustering was further examined by supervised statistical analyses. Global correlation matrices were generated to quantify the degree of coordinated expression between the four tissue types. The correlation of all protein levels in NE and NS was compared with the correlation of all proteins in TE and TS. Using that method, at an overall global level, the TE–TS portrait ($r = 0.64$) was slightly more than the global NE–NS portrait ($r = 0.6$), reaching borderline statistical significance (P-value 0.067) using a Mann–Whitney–Wilcoxon rank-sum test.

Random Forest bioinformatic analysis was also used as an alternative supervised technique to evaluate the ongoing signalling endpoints and probes the relative relationships between tissue types. The algorithm uses an ensemble of decision trees trained on a bootstrapped set of samples and a subset of the features (endpoints) with the capability of assessing each endpoint's potential utility in discerning the different classes of samples (TE, TS, NE and NS). Table 2 shows confusion matrices from each of the four class comparisons. One hundred

Table 1 Clinicopathological data for patients

Case	*Age*	*Sex*	*Tumour site*	*Dukes' stage*[a]	*Nodal status*	*Lymphocytic response*	*Invasive pattern*	*Tumour differentiation*
1	60	F	Caecum	C	Positive	—	—	Moderate
2	53	M	Sigmoid	B	Negative	—	—	Moderate
3	80	F	Rectum	A	Negative	—	—	Moderate
4	72	M	AC	A	Negative	Poor	Pushing	Moderate
5	66	F	Rectum	A	Negative	Poor	Pushing	Moderate
6	62	M	RS	B	Negative	Poor	Infiltrative	Moderate
7	83	F	Caecum	B	Negative	Poor	Infiltrative	Moderate
8	83	M	AC	A	Negative	—	—	Well
9	82	F	Rectum	B	Negative	None	Infiltrative	Moderate
10	60	M	DC	B	Negative	—	—	Moderate
11	63	M	Caecum	C	Positive	—	Infiltrative	Poor
12	71	F	DC	C	Positive	Poor	Infiltrative	Moderate
13	89	F	Caecum	B	Negative	—	—	Moderate
14	66	F	AC	B	Negative	Poor	Infiltrative	Poor
15	58	F	Caecum	C	Positive	—	Infiltrative	Moderate
16	84	F	AC	C	Positive	Poor	Infiltrative	Moderate
17	66	F	Caecum	B	Negative	—	—	Moderate
18	73	F	Sigmoid	C	Positive	—	Infiltrative	Moderate
19	75	M	DC	B	Negative	—	—	Moderate
20	79	F	Sigmoid	B	Negative	Poor	Infiltrative	Moderate
21	67	F	TC	D	Negative	Poor	Infiltrative	Moderate
22	66	M	DC	B	Negative	—	Infiltrative	Moderate
23	71	M	Sigmoid	B	Negative	Poor	Infiltrative	Moderate
24	71	M	Sigmoid	A	Negative	—	Infiltrative	Moderate
25	79	M	AC	B	Negative	—	—	Moderate
26	67	M	AC	A	Negative	Poor	Pushing	Moderate
27	63	M	DC	B	Negative	—	—	Moderate
28	75	M	TC	C	Positive	Poor	Infiltrative	Poor
29	53	F	Caecum	C	Positive	—	—	Moderate
30	78	F	Caecum	C	Positive	—	—	Poor
31	64	M	TC	D	Negative	—	—	Moderate
32	65	M	RS	C	Positive	—	—	Moderate
33	52	M	RS	A	Negative	Poor	Infiltrative	Moderate
34	61	M	RS	C	Positive	—	—	Moderate
35	87	M	AC	C	Positive	—	—	Moderate

Abbreviations: AC, ascending colon; DC, descending colon; F, female; M, male; TC, transverse colon. [a]Astler-Collins modified Dukes' stage.

per cent (33/33) of the 'normal stroma' group were classified correctly as normal stroma. Similarly, 31 of the 'normal epithelium' group were classified as such, and 2 were incorrectly classified as normal stroma. Thus, the mean class error rate for separating normal epithelium from stroma is 3%. Thirty of the tumour epithelia were correctly classified as tumour epithelium and three incorrectly classified using the same approach. Similarly, 27 of the tumour stroma were classified as such, and 6 were incorrectly grouped as tumour epithelium, giving an average class error rate for separation of malignant epithelium from stroma of 13.64%. Thus, overall, the class error rate for separation of epithelium from stroma in the tumour setting was higher than in their matched normal counterparts (McNemer's P-value of <0.039), suggesting that there is a statistically significant difference between the classification performance.

We next used principal components analysis to determine the relative cumulative distribution of the four tissue types in a multidimensional space. As shown in Figure 2, normal epithelium and normal stroma represent relatively distinct and compact data clouds. Tumour epithelium demonstrated a relatively distinguishable proteomic signature. In contrast, tumour stroma showed a more heterogeneous distribution of signalling portraits among the patients, making it difficult to differentiate this data cloud from the other tissue types.

Figure 3 shows a scatter plot of TE–TS correlations vs NE–NS correlations. A total of 23 out of 35 samples lie below the $X = Y$ line, representing samples with greater similarity in TE–TS profile than in NE–NS profile. The binomial probability of such an event is $P \leqslant 0.044$.

Proteins involved in cell proliferation appear to be an important component of the epithelial-stromal likeness observed within the colorectal tumours. To assess the altered stroma/epithelium coordination at an individual protein level, each antibody's correlation between the two tumour sample types (TE and TS) was compared with a null distribution of correlations in normal sample types (NE and NS) generated by 5000 bootstraps. Proteins identified as being more highly correlated (and reaching statistical significance $P<0.0008$) in the tumour tissue compartments compared with normal included cleaved Caspase-3, paxillin as well as the phosphorylated forms of p38, Ras-GRF and interferon κB (IκB). Statistical significance was also reached for the prostaglandin receptors EP2 and EP4 and the phosphorylated forms of c-Abl, Met, focal adhesion kinase and signal transducer and activator of transcription-1

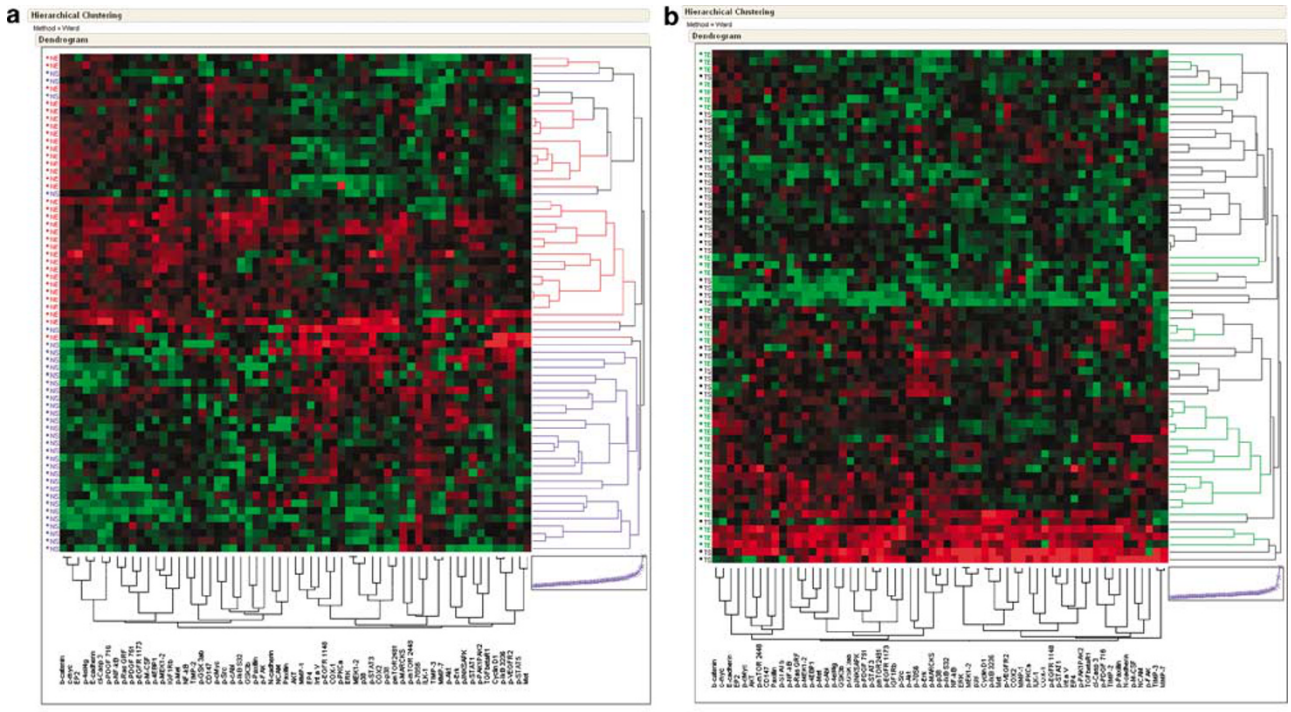

Figure 1 Hierarchical clustering of normal and tumour epithelial and stromal components. (**a**) Normal epithelium vs normal stroma; (**b**) Tumour epithelium vs tumour stroma. The two tissue phenotypes are on the vertical axis, the 61 endpoints are outlined on the horizontal axis and their phosphorylation sites are listed in Table 3. Higher relative levels of signal are represented in red; lower levels are in green.

Table 2 Confusion matrices from each of the class comparisons between the four tissue types

NE vs TE				*NS vs TS*			
	TE	*NE*	*Class error*		*TS*	*NS*	*Class error*
TE	28	5	0.15	TS	27	6	0.18
NE	7	26	0.21	NS	8	25	0.24
NS vs NE				*TS vs TE*			
	NS	*NE*	*Class error*	*A*	*TS*	*TE*	*Class error*
NS	33	0	0	TS	27	6	0.18
NE	2	31	0.06	TE	3	30	0.09

Abbreviations: NE, 'normal' (that is uninvolved) epithelium; NS, 'normal' (that is uninvolved) stroma; TE, tumour epithelium; TS, tumour stroma. Two patients were excluded in this analysis as a result of missing data, leaving 33 patients – this test cannot handle 'missing' data. There are two quadrants in the table, each corresponding to one of the pairwise comparisons. Average class error rates: NE vs TE 0.18; NS vs TS 0.21; NS vs NE 0.03; TS vs TE 0.136.

($P<0.025$), but none of those relationships survived Bonferroni multiple comparisons correction.

A few proteins were less correlated in the tumour epithelial and stromal compartments in comparison with normal. Those reaching statistical significance ($P<0.0004$) were matrix metalloproteinase (MMP)-1, nuclear factor κB (NF-κB) and the phosphorylated forms of m-CSF and NF-κB. Proteins reaching significance ($P<0.010$) but failing to withstand multiple comparisons correction included integrin α_V, MMP-7, cyclin D1 and the phosphorylated 4EBP1, vascular endothelial growth factor receptor 2 and c-Jun N-terminal kinase/stress-activated protein kinase.

A further analysis, based on clustering of variables around latent components, was used to categorize proteins into groups based on the similarity of their profiles. Analyses were then performed to assess statistical interactions between the proteins in the subclusters for tissue 'type' (normal/tumour) and 'location' (epithelium/stroma). Two sets of proteins were identified. One cluster consisting of 32 proteins showed a significant interaction between tissue type and location in that the epithelial and stromal expression levels were closer in tumour than in normal samples ($P=0.038$) but did not survive multiple comparisons testing. In particular, the significant proteins identified in the previous analysis were also within the cluster. Therefore, using two different and independent analytical approaches, we identified proteins that best segregate tissue types and locations.

Lymphocyte infiltration of colorectal cancers does not contribute to signalling portraits

The tissue analysed in this study was microdissected by laser capture microdissection (LCM), which can isolate and procure distinct cell populations within a highly

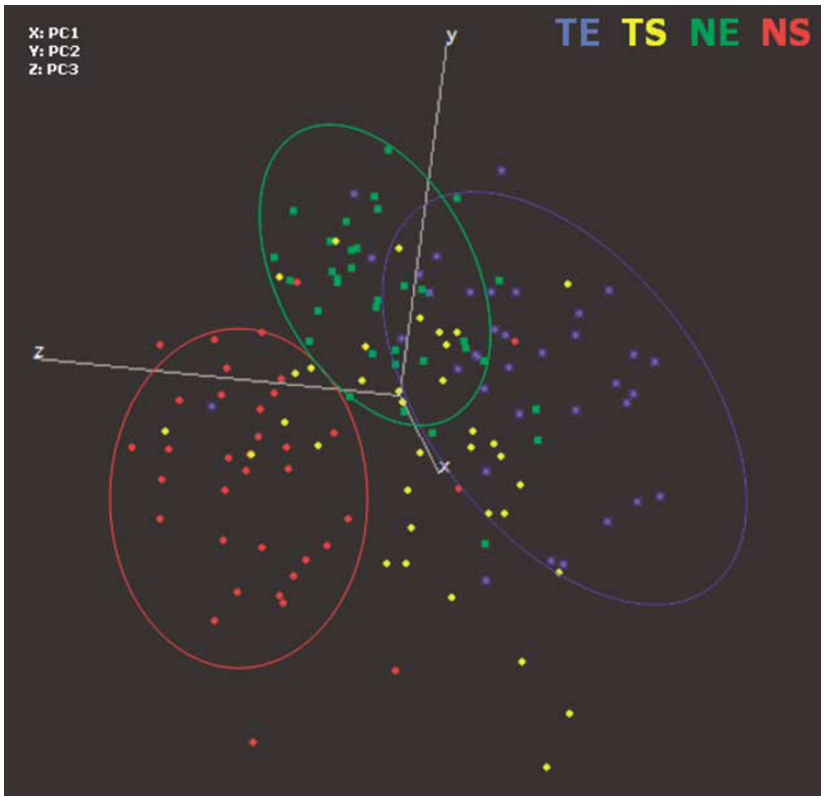

Figure 2 Principal component analysis. This analysis demonstrates all protein endpoints in a multidimensional space using the first three principal components. Each spot represents the global proteomic signatures of one of the four tissue types.

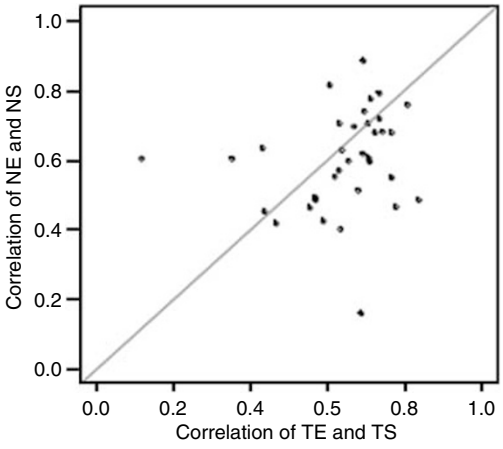

Figure 3 Scatter plot of TE–TS correlations vs NE–NS correlations. A total of 23 out of 35 samples lie below the X = Y line, representing samples with greater similarity in TE–TS profile than in NE–NS profile ($P \leqslant 0.044$).

heterogeneous tissue specimen such as epithelium and stroma. Lymphocytes were the dominant inflammatory cell type identified in these sections and immuno-LCM (Fend *et al*., 1999) was not feasible. Therefore, to rule out the possibility of inflammatory cell signalling as a confounding factor in the protein expression portraits, whole sections of the normal and cancerous colonic tissues were immunostained with a lymphocyte marker (CD45) as described in the methodology. No protein was found to significantly correlate with the degree of lymphocyte invasion in either the epithelial or stromal compartments of the uninvolved or cancerous tissues and thus the results are not unfairly biased by the presence of lymphocytes.

Analysis of inter-pathway crosstalk identifies correlation between phosphorylated AKT in the epithelium and TIMP-2 in the stroma

Thus far, we have described analyses focused on phosphorylation and protein expression changes in the same protein in different tissue compartments of uninvolved and malignant samples (for example extracellular signal-regulated kinase (ERK) in epithelium vs ERK in stroma). An alternative approach to determine relative likeness of tissue compartments is to assess the coordinated signalling/phosphorylation between proteins in different pathways (example ERK in epithelium vs SMAD in stroma, and so on). Such an analysis was performed by computing matrices of all antibodies against all antibodies in both the uninvolved colon and tumour epithelium and stromal compartments.

Two matrices of all-pairs protein expression correlations were generated, NE vs NS and TE vs TS. Following a Fisher's *z*-transformation as described in 'Materials and methods', the protein pairs with large differences in correlation *z*-score between the tumour

and the normal were noted. One protein pair, phosphorylated AKT in epithelium and tissue inhibitors of metalloproteinase (TIMP)-2 in stroma was found to become more positively correlated in the tumour epithelium and stroma compared with uninvolved tissues (z-score 3.78; false discovery rate (FDR) < 0.2).

Discussion

Using a series of unsupervised and supervised bioinformatic techniques, the major finding of this study is that coordinated activation of cell signalling proteins and pathways in stroma and epithelium are more alike in the tumour compartment compared to those cell types in normal mucosa taken from the same patient. This result is not entirely unexpected. An important concept to consider in the interpretation of these results is epithelial mesenchymal transformation (EMT) of the tumour cells. A carcinoma cell that has undergone mesenchymal or stromal transformation takes on some characteristics of stromal fibroblasts. Recently it has been proposed that EMT plays a role in the promotion of the invasion phenotype (Grunert *et al*., 2003; Cristofoli, 2006; Kopfstein and Christofori, 2006; Orimo and Weinberg, 2006; Yang *et al*., 2006). While this process is thought to be required physiologically during embryogenesis, its persistence in tumour cells may be causally related to invasion and metastasis. EMT is associated with a loss of cell–cell contacts and a morphologic shape shift from a cuboidal state to an elongated configuration associated with migration. A number of molecular changes accompany the EMT (Orimo and Weinberg, 2006; Yang *et al*., 2006). Alteration in the transcription of a series of genes has been implicated in the in EMT including Twist, Slug, Snail2, E12/E47 and S1P1 (Peinado *et al*., 2004; Cristofoli, 2006; Orimo and Weinberg, 2006; Yang *et al*., 2006). In addition, a number of cytokines, motility factors, cytoskeletal elements and degradative enzymes are also altered during EMT in a manner thought to promote invasion (McCawley *et al*., 1998; Derycke *et al*., 2005; Thomson *et al*., 2005a; Cristofoli, 2006). These recent insights support the concept that invasion and metastasis does not require a wholly new set of genes or programmes. Instead the invading cancer cell is merely using the programme that is already pre-existing for the normal EMT. Under this interpretation, the correlation herein observed in the signalling pathways for the stromal compartment compared to the carcinoma cells may be influenced by carcinoma cells undergoing EMT, which 'cross over' into the stroma.

It was found that activation of proteins involved in regulation of cell proliferation appears to be more closely matched in tumour stroma and epithelium and many are associated with the mitogen-activated protein kinase (MAPK) pathway including phospho Ras-GRF, phospho p38 and phospho p-IκB. These results were confirmed by western blotting (data not shown). This observation suggests the MAPK pathway as an interesting therapeutic target whose modulation could effect both stromal and epithelial compartments. Specific selective inhibitors of the MAPK family are already in pre-clinical and clinical trial testing, and may thus play a role in combinatorial strategies for the treatment of colon cancer (Sun and Sinicrope, 2005). Also, the observed coordinated correlation of phosphorylation of AKT in the epithelium with TIMP-2 in the stroma of tumour tissue is intriguing. TIMPs inhibit metalloproteinase activity, and hence the prevention of tumour progression. However, recent studies demonstrate that they are multifunctional proteins, and may have both an inhibitory and stimulatory effect on tumourigenesis (Jiang *et al*., 2002). Immunohistochemical studies have demonstrated increased localization of TIMPs in stromal cells during colonic tumourigenesis (Tomita and Iwata, 1996). Moreover, TIMP-1-induced cell survival is mediated via the phosphatidyl-inositol 3 kinase/AKT cascade (Lambert *et al*., 2003). Since targeted inhibitors against AKT are being widely developed given this proteins important role in tumourigenesis, further studies on the observed correlation between phosphorylated AKT and TIMP-2 and possible interaction in colorectal cancer are justified.

The phenomenon of tumour–stroma crosstalk, and the mimicry of cellular signalling by tumour stromal and epithelium, may have important clinical implications. It has been recently reported that EMT, and proteins more typically associated with the stromal compartment, such as vimentin and E-cadherin, appear predictive of response to certain targeted therapy and associated outcome (Thomson *et al*., 2005b). In the light of these findings, an important attribute of this study is that analysing the stromal and epithelial compartments using phosphoproteomic portraits of kinase substrates generates data on endpoints that are important drug targets themselves. Knowledge of which signalling pathways are being utilized by both the tumour cells and their surrounding cellular milieu permits the potential for building functional interaction maps, at a patient-specific level, and then providing a basis for direct therapeutic modulation of aberrant kinase signalling. Mathematical models have demonstrated that targeting multiple signalling interconnections will achieve better efficacy than inhibition of interdependent 'nodes' simultaneously (Araujo *et al*., 2005). Moreover, many of the novel therapeutic agents in development have the capacity to target extracellular matrix proteins, endothelial cells, pericytes and stromal fibroblasts (Joyce, 2005). If stroma actively supports cancer survival and progression, therapy aimed at stromal targets may be useful alone or, more likely, in combination with agents directed at targets in the cancer cells. Stromal components are not subject to the same degree of genetic instability and potential for escaping therapy (Sawyers, 2004).

Materials and methods

Patient samples

From 2003 to 2004, fresh paired tissue samples of colorectal cancer and corresponding normal mucosa were collected from

patients who underwent colectomy due to colorectal cancer in Beaumont Hospital, Dublin, Ireland. A pathologist dissected cancerous tissue from the tumour mass as well as uninvolved mucosa from the adjacent uninvolved colon. The paired tissue samples were rapidly processed to ensure preservation of molecular endpoints placed in separate vials and snap-frozen in liquid nitrogen. The time from removal of a colectomy specimen to snap-freezing of samples was <10 min. Samples were stored at -80°C. All specimens were obtained with informed patient consent, and the study was approved by local ethics committees at Beaumont Hospital, Ireland, as well as the National Institutes of Health, Bethesda, MD, USA. All tumours were conventional colonic adenocarcinomas; the patients had not received neoadjuvant therapy before colectomy. Histological sections from all frozen specimens were evaluated by a pathologist to determine suitability before inclusion into the study. Clinical and pathological parameters and the tissue study set description are presented in Table 1.

Laser capture microdissection
LCM was performed to evaluate epithelium and stroma separately for cell signalling analysis. Two histopathologists (KS and CG) performed the microdissection. In the light of the results, it should be noted that avoidance of stromal-epithelial cross-contamination during LCM was easier in tumour given the large epithelial islands and broad fibrous bands. Using LCM (Pixcell II, Arcturus Bioscience, Mountain View, CA, USA) (Emmert-Buck *et al*., 1996), approximately 6000 laser shots (estimated $>20\,000$ cells) were obtained from both epithelium and stroma in each sample, from separate cryostat sections. Therefore, four samples were generated from each patient: TE, TS, NE and NS. No attempt was made to target specific regions of stroma or epithelium in either tumour or uninvolved mucosa, and multiple separate areas of tissue were dissected so that signalling analysis could be performed on a cell population-wide scale within each patient sample. Tissue processing and preparation of tissue lysates have been described previously (Wulfkuhle *et al*., 2003; Sheehan *et al*., 2005).

Reverse phase protein microarray construction
Detailed descriptions of the technology and associated methodology have been published previously (Paweletz *et al*., 2001; Liotta *et al*., 2003; Wulfkuhle *et al*., 2003; Sheehan *et al*., 2005). Briefly, lysates were loaded into 384-well plates in serial dilutions (neat, 1:2, 1:4, 1:8 and 1:16) with negative control wells containing lysis buffer only. Each dilution series was printed in duplicate onto nitrocellulose-coated glass slides (Schleicher and Schuell Bioscience, Keene, NH, USA) using a ring-and-pin robotic arrayer (GMS 417, Affymetrix, CA, USA). Each array could accommodate 30 cases, necessitating five separate arrays for the entire study. To bridge endpoints across separate arrays and compare endpoints across the entire study set, reference lysates were also printed onto every array. Lysates from A431 cell lines ($\pm$EGF stimulation; BD Pharmingen, San Diego, CA, USA) as well as a reference lysate prepared from several tissue specimens from multiple cases included in the study. For a detailed discussion of the technique, please see recent review (Gulmann *et al*., 2006).

Total protein determination
Before antibody staining, the total protein concentration was determined using Sypro Ruby Protein Blot Stain (Molecular Probes, Eugene, OR, USA) on one representative slide from each array run. This technique has been described previously (Sheehan *et al*., 2005). The slides were analysed using a Fluorchem Imaging System (Alpha Innotech, San Leandro, CA, USA).

Immunostaining
Microarrays were immunostained as previously described (Wulfkuhle *et al*., 2003) on an automated slide stainer (Dako, Carpinteria, CA, USA) using a biotinyl-linked catalysed signal amplification system (Dako). All antibodies were validated by western blot as described previously (Paweletz *et al*., 2001) and antibody concentrations were optimized using test arrays similar to those included in the study. A list of antibodies used in the study is shown in Table 3.

Image analysis
Stained slides were scanned using Adobe Photoshop 6.0. (Adobe, San Jose, CA, USA) on a UMAX PowerLook III scanner (UMAX, Dallas, TX, USA) as 16-bit TIFF images at 600 d.p.i. The TIFF images were analysed using MicroVigene software (VigeneTech, Boston, MA, USA) (this is discussed in a recent review, Gulmann *et al*., 2006). Briefly, the slopes for the dilution curve of each case were plotted. The antibody intensity was normalized to total protein for each sample (Sypro stain as above) and the replicates for each case were averaged. The Y-intercept intensity values for each antibody were imported into Microsoft Excel (Microsoft, Redmond, WA, USA). Finally, each case was standardized to the bridging case or (when bridging cases were poorly analysable) to the median value of the corresponding array.

Assessment of lymphocyte infiltration in stroma and epithelium
Because many signalling pathway and specific protein phosphorylations are ubiquitous across different cell types (for example phosphorylation of AKT), it is important to ensure that the signalling endpoint intensities determined emanate from the appropriate cell type. Thus, lymphocyte density within epithelium and stroma was analysed to ensure the proteomic tissue portraits were not being unfairly biased by lymphocytes. Sections from two random blocks from each case, one containing tumour and the other uninvolved mucosa, were stained immunohistochemically with antibody for CD45 (Dako) using a standard protocol with heat-induced antigen retrieval and counterstained with haematoxylin. Under a $\times 40$ objective lens, using a grid covering half the visual field, 10 random fields that included epithelium as well as stroma were assessed for lymphocyte density. A lymphocyte was defined as any nucleated cell with a size and shape consistent with a lymphocyte and with continuous membrane positivity for CD45. Intraepithelial lymphocyte density was assessed by, in each field, counting the number of intraepithelial lymphocytes per 50 epithelial cell nuclei in the portion of glandular tissue closest to or at the lower right corner of the grid. In 10 fields, therefore, the total lymphocyte count was per 500 epithelial cells. Stromal lymphocyte density was assessed by dividing the number of grid intercepts with stromal lymphocytes by the number of stromal intercepts in 10 fields.

Data analysis
Principal components analysis and two-way hierarchical clustering analyses (Weinstein *et al*., 1997; Figures 1a and b) were performed using JMP 5.0 software (SAS Institute, Cary, NC, USA).

Clustering of variables Analysis of homogeneity between variables (proteins) rather than patients was performed using Stata/SE Release 9.2. Clustering of variables around latent components works in a stepwise manner, grouping variables

Table 3 Antibodies used in reverse phase protein microarrays

Antibody[a]	*Source*	*Species*	*Dilution*
COX-1	Cay	M	1:100
COX-2	Up	R	1:50
EP2 Receptor	Cay	R	1:100
EP4 Receptor	Cay	R	1:100
Cyclin D1	CS	M	1:100
β-Catenin	CS	R	1:200
c-Myc	CS	R	1:50
p-c-Myc (Thr58/Ser62)	CS	R	1:100
p-c-Abl (Thr735)	CS	R	1:100
p-ras-GRF (Ser916)	CS	R	1:50
p-PDGF Receptor β (Tyr751)	CS	R	1:50
p-PDGF Receptor β (Tyr716)	Up	R	1:250
p-EGF Receptor (Tyr1148)	BS	R	1:100
p-EGF Receptor (Thr1173)	Up	M	1:50
p-VEGF Receptor 2 (Tyr951)	CS	M	1:50
p-PKCα (Ser657)	Up	R	1:2000
NF-κB	CS	R	1:100
p-NF-κB (Ser536)	CS	R	1:50
p-IκB-α (Ser32)	CS	R	1:25
p-IκB-α (Ser32/36)	CS	M	1:50
p-SAPK/JNK (Thr183/Tyr185)	CS	R	1:100
p38	CS	R	1:100
p-p38 (Thr180/Tyr182)	CS	R	1:50
ERK	CS	R	1:200
p-ERK 1/2 (Thr202/Tyr204)	CS	R	1:1000
MEK1/2	CS	R	1:750
p-MEK1/2 (Ser217/221)	CS	R	1:100
AKT	CS	R	1:200
p-AKT (Ser473)	CS	R	1:50
GSK-3β	CS	R	1:50
p-GSK-3α/β (Ser21/9)	CS	R	1:50
p-mTOR (Ser2448)	CS	R	1:50
p-mTOR (Ser 2481)	CS	R	1:50
p-MARCKS (Ser152/156)	CS	R	1:20
p-STAT1 (Ser727)	Up	R	1:200
p-STAT3 (Ser727)	BS	R	1:100
p-STAT5 (Tyr694)	CS	M	1:50
p-70 S6 (Tht421/Ser424)	CS	R	1:500
p-4E-BP1 (Thr37/46)	CS	R	1:50
p-eIF4G (Ser1108)	CS	R	1:100
p-PAK1 (Ser199/204)/PAK2 (Ser192/197)	CS	R	1:25
c-Caspase 3 (Asp175)	CS	R	1:50
CD147	Zy	R	1:100
E-Cadherin	CS	R	1:100
ILK-1	CS	R	1:200
p-M-CSF Receptor (Tyr723)	CS	R	1:50
p-Src (Tyr416)	CS	R	1:100
TGF-β Receptor I	CS	R	1:50
Met	CS	M	1:50
p-Met (Tyr1234/1235)	CS	R	1:30
Integrin α_V	BD	M	1:50
N-Cadherin	BD	M	1:50
NCAM	Ch	M	1:50
MMP-1	Ch	R	1:500
MMP-7	Ch	R	1:300
TIMP-2	Ch	R	1:100
TIMP-3	Ch	R	1:500
Paxillin	CS	R	1:50
p-Paxillin (Tyr118)	CS	R	1:30
IGF-I Receptor β	CS	R	1:100
p-FAK (Typ397)	BD	M	1:25

Abbreviations: BD, BD Biosciences (San Jose, CA, USA); BS, Biosource Int. (Camarillo, CA, USA); Cay, Cayman Chemical (Ann Arbor, MI, USA); Ch, Chemicon International Inc. (Temecula, CA, USA); CS, Cell Signalling Technology (Beverly, MA, USA); M, mouse; R, rabbit; Up, Upstate (Charlottesville, VA, USA); Zy, Zymed Laboratories (San Francisco, CA, USA). [a]p, phosphorylated at residue shown; c, cleaved at residue shown.

into clusters so as to minimize the variation within clusters and maximize the variation between clusters. Clusters were identified based on variation in the T criterion (used to identify the point in the clustering process in which homogeneity within clusters and heterogeneity between clusters reached a maximum). Analysis of variance and linear regression was used to examine the interaction between tissue type and location.

Sample classification Supervised sample classification was performed using Random Forests (Breiman, 2001) in R (www.r-project.org). This algorithm conveniently has only two tunable parameters (mtry and ntrees) and is capable of assessing its own performance error distinguishing training from test examples. The standard settings of mtry = 7 and ntrees = 500 were found to perform best in the analysis.

Correlations Global correlation matrices were generated to quantify the degree of coordinated protein expression within each specimen and between the four sample types (NS, NE, TS and TE). To determine the significance of the difference in correlations, paired Mann–Whitney tests were performed across the 35 specimens. To assess the altered stroma/epithelium coordination concurrent with tumourigenesis, each antibody's correlation between the two tumour sample types (TE and TS) was compared with a null distribution of 5000 bootstrapped correlations of the normal sample types (NE and NS). *P*-values were estimated by the number of bootstrapped correlations that were more extreme than the correlations observed with the tumour samples. Significance level was defined as $P<0.05$. Multiple comparisons correction was performed using the Bonferroni technique.

To elucidate changes in protein interactions among different proteins that occur in the epithelium and stroma from normal to tumour tissues, an analysis of the changes in protein correlations was performed between all protein pairs. The endpoints were first filtered to remove those within the bottom 60% of variability within each NE, NS, TE or TS. For every pair of endpoints that survived filtering, Pearson's correlation coefficients were computed between the stromal and epithelial compartments in both the normal and tumour samples. This procedure resulted in two matrices of all-pairs protein expression correlations, NE vs NS and TE vs TS. The correlation coefficients were then transformed to a *z*-score using a Fisher's *z*-transformation to make them comparable between the normal samples and the tumour samples. The protein pairs with large differences in correlation *z*-scores between the tumour matrix and the normal matrix were noted. An FDR, the likelihood of discovering false correlation differences of equivalent or greater magnitude by chance, was computed by permuting the patient labels and performing the same analysis 10 000 times.

For the analysis of lymphocyte infiltration and antibody staining, correlation coefficients were calculated between the antibody and the degree of lymphocyte invasion.

References

Araujo RP, Doran C, Liotta LA, Petricoin EF. (2005). Network targeted combination therapy: a new concept in cancer treatment. *Drug Discov Today* **1**: 425–433.

Bergers G, Benjamin LE. (2003). Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* **3**: 401–410.

Bhowmick NA, Neilson EG, Moses HL. (2004). Stromal fibroblasts in cancer initiation and progression. *Nature* **432**: 332–337.

Breiman L. (2001). Random Forests. *Machine Learning* **45**: 5–32.

Brigati C, Noonan DM, Albini A, Benelli R. (2002). Tumor and inflammatory infiltrates: friends or foes? *Clin Exp Metastasis* **19**: 247–258.

Brown PO, Botstein D. (1999). Exploring the new world of the genome with DNA microarrays. *Nat Genet* **21**: 33–37.

Coussens LM, Werb Z. (2002). Inflammation and cancer. *Nature* **420**: 860–867.

Cristofoli G. (2006). New signals from the invasive front. *Nature* **441**: 444–450.

Derycke L, Van Marck V, Depypere H, Bracke M. (2005). Molecular targets of growth, differentiation, tissue integrity, and ectopic cell death in cancer cell. *Cancer Biother Radiopharm* **20**: 579–587.

De Wever O, Mareel M. (2003). Role of tissue stroma in cancer cell invasion. *J Pathol* **200**: 429–447.

Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR *et al*. (1996). Laser capture microdissection. *Science* **274**: 998–1001.

Fend F, Emmert-Buck MR, Chuaqui R, Cole K, Lee J, Liotta LA *et al*. (1999). Immuno-LCM: laser capture microdissection of immunostained frozen sections for mRNA analysis. *Am J Pathol* **154**: 61–66.

Gallagher PG, Bao Y, Prorock A, Zigrino P, Nischt R, Politi V *et al*. (2005). Gene expression profiling reveals cross-talk between melanoma and fibroblasts: implications for host–tumor interactions in metastasis. *Cancer Res* **65**: 4134–4146.

Gulmann C, Sheehan KM, Kay EW, Liotta LA, Petricoin III EF. (2006). Array based proteomics: mapping of protein circuitries for diagnostics, prognostics and therapy guidance. *J Pathol* **208**: 595–606.

Grunert S, Jechlinger M, Beug H. (2003). Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. *Nat Rev Mol Cell* **4**: 657–665.

Jiang Y, Goldberg ID, Shi YE. (2002). Complex roles of tissue inhibitors of metalloproteinases in cancer. *Oncogene* **21**: 2245–2252.

Joyce JA. (2005). Therapeutic targeting of the tumor microenvironment. *Cancer Cell* **7**: 513–520.

Kopfstein L, Christofori G. (2006). Metastasis: cell-autonomous mechanisms versus contributions by the tumor microenvironment. *Cell Mol Life Sci* **63**: 449–468.

Lambert E, Boudot C, Kadri Z, Soula-Rothhut M, Sowa ML, Mayeux P *et al*. (2003). Tissue inhibitor of metalloproteinases-1 signalling pathway leading to erythroid cell survival. *Biochem J* **372**: 767–774.

Liotta L, Espina V, Mehta A, Calvert V, Rosenblatt K, Geho D *et al*. (2003). Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* **3**: 317–325.

Liotta LA, Kohn EC. (2001). The microenvironment of the tumor-host interface. *Nature* **411**: 375–379.

Littlepage LE, Werb Z. (2005). Coevolution of cancer and stromal cellular responses. *Cancer Cell* **7**: 499–500.

Matrisian LM, Cunha GR, Mohla S. (2001). Epithelial-stromal interactions and tumor progression: meeting summary and future directions. *Cancer Res* **61**: 3844–3846.

McCawley LJ, O'Brien P, Hudson LG. (1998). Epidermal grow factor (EGF), and scatter factor/hepatocyte (SF/HGF)-mediated keratinocyte migration is a coincident with induction of matrix metalloproteinase (MMP)-9. *J Cell Physiol* **176**: 255–265.

Mueller MM, Fusenig NE. (2002). Tumor-stroma interactions directing phenotype and progression of epithelial skin tumor cells. *Differentiation* **70**: 486–497.

Nishizuka S, Charboneau L, Young L, Major S, Reinhold WC, Waltham M *et al*. (2003). Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *Proc Natl Acad Sci USA* **100**: 14229–14234.

Notterman DA, Alon U, Sierk AJ, Levine AJ. (2001). Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. *Cancer Res* **61**: 3124–3130.

Orimo A, Weinberg RA. (2006). Stromal fibroblasts in cancer: a novel tumor-promoting cell type. *Cell Cycle* **5**: 1597–1601. [E-pub 2006 Aug 1].

Paweletz CP, Charboneau L, Bichsel VE, Simone NL, Chen T, Gillespie JW *et al*. (2001). Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* **20**: 1981–1989.

Peinado H, Portillo F, Cano A. (2004). Transcriptional regulation of cadherins during development and carcinogenesis. *Int J Dev Biol* **48**: 365–375.

Sawyers C. (2004). Targeted cancer therapy. *Nature* **432**: 294–297.

Sheehan KM, Calvert VS, Kay EW, Lu Y, Fishman D, Espina V *et al*. (2005). Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Mol Cell Proteomics* **4**: 346–355.

Sun Y, Sinicrope FA. (2005). Selective inhibitors of MEK1/ERK44/42 and p38 mitogen-activated protein kinases potentiate apoptosis induction by sulindac sulfide in human colon carcinoma cells. *Mol Cancer Ther* **4**: 51–59.

Thiery JP. (2002). Epithelial—mesenchymal transitions in tumor progression. *Nat Rev Cancer* **2**: 442–454.

Thomson EW, Newgreen DF, Tarin D. (2005a). Carcinoma invasion and metastasis: a role for epithelial-mesenchymal transition? *Cancer Res* **65**: 5991–5995.

Thomson S, Buck E, Petti F, Griffin G, Brown E, Ramnarine N *et al*. (2005b). Epithelial to mesenchymal transition is a determinant of sensitivity of non-small-cell lung carcinoma cell lines and xenografts to epidermal growth factor receptor inhibition. *Cancer Res* **65**: 9455–9462.

Tlsty TD, Hein PW. (2001). Know thy neighbor: stromal cells can contribute oncogenic signals. *Curr Opin Genet Dev* **11**: 54–59.

Tomita T, Iwata K. (1996). Matrix metalloproteinases and tissue inhibitors of metalloproteinases in colonic adenomas-adenocarcinomas. *Dis Colon Rectum* **39**: 1255–1264.

Weinstein JN, Myers TG, O'Connor PM, Friend SH, Fornace Jr AJ, Kohn KW *et al*. (1997). An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**: 343–349.

Wulfkuhle JD, Aquino JA, Calvert VS, Fishman DA, Coukos G, Liotta LA *et al*. (2003). Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase protein microarrays. *Proteomics* **3**: 2085–2090.

Yang J, Mani SA, Weinberg RA. (2006). Exploring a new twist on tumor metastasis. *Cancer Res* **66**: 4549–4552.

Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR *et al*. (1997). Gene expression profiles in normal and cancer cells. *Science* **276**: 1268–1272.