A signature predicting poor prognosis in gastric and ovarian cancer represents a coordinated macrophage and stromal-response.

Authors:
Rita A. Busuttil\textsuperscript{1,2,3,*}, Joshy George\textsuperscript{1,*}, Richard W. Tothill\textsuperscript{4}, Kylie Ioculano\textsuperscript{1}, Adam Kowalczyk\textsuperscript{5}, Catherine Mitchell\textsuperscript{6}, Stephen Lade\textsuperscript{6}, Patrick Tan\textsuperscript{7,8,9,10}, Izhak Haviv\textsuperscript{1,11,12} and Alex Boussioutas\textsuperscript{1,2,3}.

Affiliations:
\textsuperscript{1}Cancer Genetics and Genomics Laboratory, Peter MacCallum Cancer Centre, East Melbourne, Australia
\textsuperscript{2}Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, Australia
\textsuperscript{3}Department of Medicine, Royal Melbourne Hospital, The University of Melbourne, Parkville, Australia
\textsuperscript{4}Molecular Genomics Core Facility, Peter MacCallum Cancer Centre, East Melbourne, Australia
\textsuperscript{5}National ICT Australia (NICTA), The University of Melbourne, Parkville, Australia
\textsuperscript{6}Department of Pathology Peter MacCallum Cancer Centre, East Melbourne, Australia
\textsuperscript{7}Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, Singapore
\textsuperscript{8}Cellular and Molecular Research, National Cancer Centre, Singapore
9 Cancer Science Institute of Singapore, National University of Singapore, Singapore
10 Genome Institute of Singapore, Singapore
11 Department of Pathology, The University of Melbourne, Parkville, Australia
12 Faculty of Medicine in the Galilee, Bar Ilan University, Zfat, Israel
* These authors contributed equally to this study

**Running title:** Stromal-response signature in cancer

**Keywords:** stroma, cancer, gene signature, macrophages, tumor

**Financial support:** This work was supported by NHMRC grant 288714.

**Cell line use:** No cell lines have been used in this research

**Conflict of interest:** The authors have no conflicts to disclose

**Microarray data:** Data has been submitted to GEO (Series GSE51105)

**Corresponding author:**
Associate Professor Alex Boussioutas
Peter MacCallum Cancer Centre.
Locked Bag 1, A’Beckett Street,
Melbourne, Victoria 8006. Australia
Telephone: +61 3 96561287, Fax: +61 3 96561414
Email: alex.boussioutas@petermac.org

**Word count:** 5241

**Total number of figs and tables:** 6
TRANSLATIONAL RELEVANCE:

Cancer is a disease involving complex interactions of the tumor cell along with a multitude of other, distinct cell types which comprise the tumor microenvironment. It is this cross-talk that promotes cancer invasion and metastasis. In this study we use a gastric dataset to identify a tumor specific gene signature termed, the stromal response cluster (SRC). This set of co-expressed genes can also be detected in multiple other cancer types and have prognostic significance in the data sets studied. We show that genes in this signature are predominantly expressed in cancer associated stroma. In addition a strong correlation was observed between the SRC and macrophages of the M2 subclass which are known to promote cancer progression. It is conceivable that identifying patients with high expression of the SRC using biopsy samples taken at the time of diagnosis could help identify candidates for adjunct therapies which act by preventing the polarization of M2 macrophages.
ABSTRACT

Purpose: Gene-expression profiling has revolutionized the way we think about cancer and confers the ability to observe the synchronous expression of thousands of genes. The use of putative genome-level expression profiles has allowed biologists to observe the complex interactions of genes that constitute recognized biological pathways. We used gastric and ovarian data sets to identify gene-expression signatures and determine any functional significance.

Experimental design:
Microarray data of 94-tumor and 45-benign samples derived from gastric cancer patients was interrogated using HOPACH analysis identifying clusters of co-expressed genes. Clusters were further characterized with respect to biological significance, gene ontology and ability to discriminate between normal and tumor tissue. Tumor tissues were separated into epithelial and stromal compartments and immunohistochemical analysis performed to further elucidate specific cell lineages expressing genes contained in the signature.

Results: We identified a “stromal-response” expression signature, highly enriched for inflammatory, extracellular matrix, cytokine and growth-factor proteins. The majority of genes in the signature are expressed in the tumor associated stroma but were absent in associated premalignant conditions. In gastric cancer, this module almost perfectly differentiates tumor from non-malignant gastric tissue and hence can be regarded as a highly tumor-specific gene-expression signature.
Conclusions We show these genes are consistently co-expressed across a range of independent gastric datasets as well as other cancer types suggesting a conserved functional role in cancer. In addition we show that this signature can be a surrogate marker for M2 macrophage activity and has significant prognostic implications in gastric and ovarian high grade serous cancer.

INTRODUCTION

Cancer is characterized by the invasion of malignant cells into surrounding supportive tissue and distant sites[1]. It is well recognized that the consequent tissue deformation leads to an inflammatory host response with elevated levels of infiltrating inflammatory and immune cells [2]. While this host reaction could merely represent a generic response of the host to injury, the genes that control heterotypic cell interactions in wound healing and cancer stroma, are subject to germ line sequence variations [3-8] suggesting that the tumor microenvironment may vary among patients exhibiting a variety of phenotypic responses. One such response was reported recently suggesting an immunosuppressive role of FAPα (Fibroblast Activation Protein alpha) expressing cells in the tumor stroma that was tumor permissive [9]. The relative contribution of stromal cells, such as Tumor Associated Macrophage (TAMs) and Carcinoma-Associated Fibroblasts (CAFs) to the cancer phenotype in different patients is of great interest [10]. More importantly, the identification of the specific stromal cell lineage responsible for cancer promotion would enable therapeutic trials aimed at patient cohorts
that would best respond to an agent targeted against specific stromal components. FAPα expressing cells may be one example as well as TAMs that are thought to promote cancer and interact with both the cancer cells [11, 12] and other stromal cells [13].

Genomic profiling using gene expression arrays was designed as a discovery tool, aimed at increased insight into the cause and therapy of disease. The approach proved useful to dissect molecular subtypes of malignancies including Diffuse Large B Cell Lymphoma (DLBCL) [14] and solid tumors such as breast cancer [15] and gastric cancer [16]. Gene expression analysis of a whole tumor is the sum of the mRNA contributions from different cell lineages, driven by cell-autonomous genomic changes as well as non-cell autonomous cell-matrix and cell-cell interactions [17]. We previously described a gene expression signature of co-expressed genes exhibiting increased expression in malignant tissue [16, 18, 19]. More recently, we found a highly related signature in a group of ovarian cancer patients, which was associated with the poorest survival in the Australian Ovarian Cancer Study (AOCS) cohort [20].

In this study we describe the discovery and exploration of a conserved signature of genes derived from gastric cancer. We extended the analysis to include independent gastric data sets [21, 22] as well as a broader group of malignancies including ovarian cancer [20], colorectal cancer [23], lung cancer [24] and breast cancer [25]. This novel finding of a conserved signature in diverse epithelial malignancies, the ontology of the genes in this signature (tissue remodeling) and the specific up-regulation of this signature in gastric cancer over the respective premalignant tissues, all suggest this
signature reflects a cancer-specific inflammatory response. We therefore termed the signature as “stromal-response” to emphasize that it is a response to the presence of malignant epithelium. Here we examine the factors that potentially control the expression of this gene expression signature and shows, for the first time, data that suggests that this signature arises from the stromal compartment of the cancer and is strongly associated with the presence of M2 polarized macrophages or TAMs.

MATERIALS AND METHODS

Patients and Samples

Patients diagnosed with operable Gastric Cancer (GC) were included in the study as part of a large cohort study where clinical, demographic and surgical pathology information was collected (PMCC cohort; Supplementary Table S1). Histology of specimens was assessed by an independent pathologist (S.L.). Patients were recruited from six metropolitan hospitals in Melbourne with written informed consent. All procedures were ethically approved by individual hospital Institutional Review Boards that was overseen by Peter MacCallum Cancer Centre IRB. Cases of Ovarian Cancer were obtained through the Australian Ovarian Cancer Study (AOCS) as previously described [20].

Microarray experiments

Fresh frozen tissue was collected at the time of surgery as previously described [16]. Whole tissue sections from 94 tumors and 45 benign tissues
were profiled. All tissues required pathology evaluation and the minimum acceptable percentage of tumor cells was 60%. Total RNA was isolated by acid phenol extraction (Trizol, Invitrogen) and column chromatography (RNeasy, Qiagen). RNA quantity and quality was assessed by spectrophotometer (Nanodrop Technologies) and Bioanalyser (Agilent Technologies). Microarrays were hybridized using Affymetrix U133+2 chips (Affymetrix) according to the protocol described in the manual. Microarrays were scanned using the Genechip Scanner (Affymetrix). Data has been submitted to GEO (Series GSE51105).

Additional/Validation Microarray datasets
Details of additional and validation data sets are provided in the Supplementary Materials and Methods.

Laser capture microscopy (LCM)
Laser capture microscopy was performed for 4 cases of gastric cancer and 5 cases of ovarian cancer. 10µm sections were placed into membrane coated slides (MDS Analytical Technologies) and stained with cresyl violet (Ambion). Tumor and stromal compartments were identified by a pathologist and individually isolated using laser capture microdissection (LCM; Arcturus Veritas, MDS Analytical Technologies). RNA isolation, quantitation and Affymetrix arrays were run as described for microarray experiments.

In situ hybridization and Immunohistochemistry
GC and premalignant tissues were ethanol fixed, paraffin embedded and used to create tissue microarrays (TMA’s) which were subsequently used for ISH and IHC. Details of relevant protocols are described in the Supplementary Materials and Methods.

Statistics

Differential expressed genes between gastric cancer and normal gastric epithelium were identified using empirical Bayes method available in the R-package limma. An FDR of 5% was used as the cut-off to select differentially expressed genes. Hierarchical Ordered Partitioning and Collapsing Hybrid (HOPACH) analysis was performed on the gastric dataset using the R statistical package, which builds a hierarchical tree of maximally homogeneous clusters by recursively partitioning a dataset while ordering and collapsing clusters at each level [26]. Eisen’s Cluster was used for hierarchical clustering using average linkage [27]. The statistical significance of co-expression of a gene set was computed using a Monte Carlo procedure. The average correlation of the pair-wise correlation of the set of genes was noted. The average correlation of a randomly selected set of the same number of genes within the same dataset was computed N times. The proportion of times the randomly selected genes had higher of equal correlation to that of the original set is used as the statistical significance of the correlation. GSEA (Gene Set Enrichment Analysis) [28] was used for assessment of gene enrichment across Australian and Hong Kong datasets as well as the tumor/normal datasets referred to above. Kaplan Meier analysis was performed for the progression free survival and log rank statistics used to
analyze the data. The mean expression level of the cluster of genes was used to stratify samples into high and low status for that particular cluster. The top and bottom 33% of the samples were classified as high and low respectively.

RESULTS

The stromal-response signature is a set of co-expressed genes that discriminates gastric cancer from benign gastric tissues

We and others have used gene expression profiling to classify cases according to groups of co-expressed gene signatures [14, 16, 19]. Often, such signatures are associated with histopathological and clinical phenotype, such as drug response, metastatic spread or overall survival. In order to employ this information to instruct patient management, further mechanistic insight as to the co-regulation of the signature genes is required. We have expanded on our previously published gene expression profiling of gastric cancers with the addition of an extra 29 tumour samples [16]. All samples in the cohort were reprofiled this time using Affymetrix U133+2 gene chips. The HOPACH [29] method was used to identify “clusters” of genes that are strongly co-expressed within a given dataset; irrespective of the pathology of the sample. All genes with a cluster membership value <0.3 were disregarded. This left nine major clusters, each including at least 34 genes, which showed significant co-expression with an FDR smaller than $10^{-8}$ (Table 1; see Supplementary Table S2 for gene lists).

Five of these HOPACH clusters (Clusters 27, 21, 12, 0 and 26) included a significant proportion of genes (over 60% of genes in the cluster) that were
also found to be statistically over-represented as discriminatory between tumor and normal based on empirical Bayes tests (Table 1; Supplementary Table S3).

Figure 1A shows hierarchical clustering of genes using Pearson correlation as a similarity matrix and average linkage method in our Affymetrix data. The relative positions of the clusters from Table 1 are also indicated. Cluster 27 not only exhibited the most significant co-expression, but also contained the highest percentage of differentially expressed genes. Figure 1B shows Eisen clustering of all 144 probes (representing 100 non-redundant genes – Supplementary Table S4) in cluster 27 which clearly differentiates normal and tumor samples.

In order to confirm that this signature is only expressed in tumor samples, a separate HOPACH analysis was also performed using all normal and premalignant samples in the cohort. None of the clusters generated in this analysis significantly overlapped with cluster 27 (data not shown) suggesting a tumor specific signature which is not seen in normal or inflamed premalignant tissue.

**Biological significance of the HOPACH clusters**

To gain more biological insight to the mechanism of the co-expression observed for each of the clusters, we performed GeneGo® MetaCore analysis on each of the top nine cancer-specific tightly co-expressed clusters. The top network from each of those clusters in GeneGo analysis was used to give a
representative ontogenic title to each cluster (Table 1). Whilst cluster 14 represented recognized stomach function, without an obvious assignment to which cell expresses those genes (tumor or stroma) some clusters were composed of cell autonomous ontologies, such as cell cycle (clusters 21 and 28). By contrast, some clusters represented reactive stroma, either of immune cell origin (clusters 26 and 30), or with fibrotic characteristics (cluster 27).

Cluster 27- termed the Stromal-Response Cluster (SRC)- was predominantly an extracellular matrix signature as shown in Figure 2A. Other major processes represented by cluster 27 genes include cell adhesion and development pathways (Figure 2B) which are driven essentially by WNT-TCF, VEGF-anoxia, and TGFβ signaling.

The stromal-response signature genes arise from stromal components specifically juxtaposed to cancer

The input RNA hybridized onto the Affymetrix arrays described in the previous analysis was derived from whole tumor pieces which ultimately included both tumor and stromal compartments. SRC could represent reactive stroma, but could also represent a mesenchymal state of the cancer cells, such as EMT. For example, the serum response signature of cell lines was originally expected to represent reactive fibroblasts but was later found to be expressed by the epithelial cells [30, 31]. Using species specific expression profiling, we and others found that some genes in stromal reaction signatures are in fact expressed by epithelial cancer cells [32, 33].
In order to better determine the specific cellular compartment of the tumor expressing the SRC, we performed laser capture microdissection (LCM) of 4 cases of gastric cancer and 5 cases of high grade serous ovarian cancer (OVCA). Epithelial and stromal components were separated by LCM and Affymetrix expression analysis on RNA derived from the two areas was performed independently.

Overall, 370 genes were overexpressed in the epithelial compartment of the gastric samples and 654 in the stromal compartment. The greatest ontological enrichment in the epithelial compartment was by genes identified in the Regulation of gene metabolism: Bile acid cluster (cluster 14) and the Cell cycle: The metaphase checkpoint cluster (cluster 21). The greatest enrichment for stromal expression was exhibited by the cell cycle regulation of G1/S (cluster 28) and the SRC (cluster 27) clusters (Table S5).

A heat map was generated showing expression of the SRC genes (cluster 27). Preferential clustering was observed in the stromal fraction for both gastric cancer and ovarian cancer (Figure 2C). We then further analyzed the distribution of the 100 unique stromal-response genes with respect to both the epithelial and stromal compartments of GC and OVCA. We observed that most (GC n=38; OVCA n=54) of the SRC genes are predominantly expressed in the GC and OVCA stroma and only 2 or 4 were highly expressed in the epithelium of GC and OVCA respectively (Figure 2D&E). Interestingly, there are a large number of genes (GC n=60 and OVCA n=42) that are not specifically up-regulated in stroma or epithelial compartments (Figure 2D&E),
suggesting that they are equally expressed in both compartments and is suggestive of a dynamic interaction at the epithelial-stromal interface (Supplementary Table S6 lists the compartmental location for each gene). Genes expressed by the epithelial cells which are tightly co-expressed with SRC, serve as evidence that variation in expression of SRC is the product of heterotypic cancer-stroma cross talk[34].

The stromal-response signature is co-expressed in multiple carcinoma types.

To determine whether each of the clusters is universal across multiple cancer types we interrogated a variety of publically available datasets as outlined in Table 1. A permutation approach (Monte Carlo Method) was used to compute the likelihood that these clusters of co-expressed genes were also over-represented in other solid cancers. The datasets we used were: colorectal carcinoma [23], lung carcinoma [24], breast carcinoma [25] and two independent gastric carcinoma cohorts [21, 22]. We found that many tumor datasets that are publicly available had few, if any, non-malignant tissue profiled as part of the dataset limiting the number of datasets available for the analysis to determine if the cluster was tumor specific.

Across all cohorts, several clusters including cluster 27, (the SRC) were tightly co-expressed. This general concordance between the cluster ontology and the robustness of each cluster expression in different organ further supports our contention that stromal-response is a robust, universal and consistent aspect of carcinoma gene expression.
To identify the specific stromal cell types expressing genes from the SRC, we performed *in situ* hybridization for the SPARC gene on a TMA comprising GC and its premalignant counterparts. SPARC was chosen as it is among the top 50 most significantly differentiated genes between tumor and normal and acted as a surrogate marker for the metagene. Representative images showing SPARC ISH on normal and tumor samples obtained from the same patient are shown in Figure 3A. As expected, staining was predominantly observed in stromal areas and whilst most stromal cell types expressed these genes to varying extents, fibroblast cells were preferentially stained. SPARC ISH was indicative of the stromal-response signature and was highly expressed in stroma juxtaposed to tumor epithelium. SPARC expression was not high in non-malignant stroma despite the presence of significant amounts of stroma in our tissue sections and despite considerable inflammation in non-malignant tissues. This suggests a tumor specific function for the stromal-response genes.

**The stromal-response signature is correlated with macrophage infiltrates.**

Although the stromal-response signature was exceptionally specific and consistent in all gastric cancer cases, the overall degree of expression was nevertheless variable across multiple cases. Across the GC and benign gastric tissues, the expression of SPARC in the microarray and the SPARC ISH signals were highly correlated. We therefore aimed to exploit this
variation to determine which stromal cell type most tightly correlates with the overall expression level of the stromal-response signature, which was defined by SPARC ISH as a surrogate of the signature.

Serial sections of the TMA were prepared. One section was stained by ISH for SPARC as described above. The remaining sections were stained with antibodies for specific markers of stromal cell lineages; anti-smooth muscle actin (anti-SMA), anti-CD31, anti-CD45, and anti-CD68 as markers of connective tissue, endothelial, leukocyte and macrophage cells, respectively. We then correlated between degree of ISH staining for SPARC and abundance of specific cell lineage(s) in the tumor, based on the markers listed above. Figure 3B shows an example of SPARC ISH on a representative tumor sample with IHC stains performed on serial sections for cell markers described. The association of expression of SPARC ISH and cell markers is summarized in Table 2 for all cases, including the non-malignant ones. We find that the expression of the SPARC correlated best with CD31 and CD68 positive cells when assessed using Fisher's exact test ($p=3.2\times10^{-6}$ and $p=3.7\times10^{-6}$, respectively). Indeed, when we assess the tumors specifically then CD68 positive cells demonstrate the most significant association (Table 2 right panel; $p=0.007$) with the expression of the SPARC ISH. Whilst we do not propose that the stromal response genes themselves are being produced by the macrophages these results suggest that macrophage (CD68-positive cell) presence in the tumor microenvironment is important for the expression of the SRC signature locally, in the setting of invasive cancer. Macrophage activation in vitro leads to elongated fibrous cell morphology of these cells,
hindering the identification of the cell type. Alternatively, it is possible that
cross talk between fibroblasts, the main cells expressing the SRC signature,
and macrophages, is a critical modulator of the overall stromal-response gene
signature level of expression.

Macrophage polarization in the cases that exhibit high stromal-response
signature

Macrophages, which originate from monocytes, can be activated or polarized
into two phenotypically distinct subtypes, each with distinct cytokine secreting
profiles and functions. M1 macrophages are said to be classically activated by
the Th1 cytokine interferon gamma (IFNγ), in the absence or presence of
lipopolysaccharide (LPS). M1 macrophages are pro-inflammatory and also
promote apoptosis and ECM destruction. M2 macrophages, in contrast, are
alternatively activated and are known to promote proliferation, invasion and
ECM construction.

Based on the results described in Table 2 we investigated whether M1 or M2
macrophages were more highly enriched in samples with high representation
of the SRC genes and analyzed our data with respect to previously published
M1 and M2 gene signatures [35].

Firstly, samples were ranked according to their overall stromal-response
signature, and classified them as either SRC high or low (see materials and
methods). The expression values for each M1 signature gene was then
determined for all samples in the cohort and the fold-change expression of
each gene between SRC high and low was determined. A similar analysis was performed for M2 signature genes. The box plot in Figure 3C shows that M2 genes are more positively correlated with SRC high samples (p=0.032). The absence of M2 genes in the SRC signature can be attributed to the small cellular contribution of macrophages to the overall tumor microenvironment and expected dilution of expression of these genes in the context of the tumor and other stromal components.

In more advanced stages of tumor development, it is generally well accepted that TAMs have an M2 phenotype based on their cytokine secretion profile with high levels of CD163 expression being a key marker. Using the remaining serial section from the TMA we then performed IHC for CD163 and correlated its expression with SPARC ISH as before (Figure 3D). Table 2 (lower panels) shows a significant correlation (p=0.002; Fishers exact test) when both normal and tumor samples were considered. There was no significant correlation of SPARC ISH/CD163 staining in tumors most likely due to a large number of missing data which can be attributed to TMA exhaustion. These results suggest that CD163+ M2 macrophages may be driving the SRC signature.

**High level-expression of the stromal-response signature predicts poor outcome.**

Comprehensive clinical data available for the PMCC and Ooi et al gastric, as well as the AOCS ovarian cohort allowed us to compare the association of each of the cluster expression levels with patient outcome (progression free survival), using Kaplan Meier analyses (Table S7, Figure 4). We analyzed the
core group of genes by using a metagene expression signature and
determined cases that express high and low levels of the signature based
upon the median expression of each metagene. In all datasets, elevated
expression of genes in SRC and cluster 28 was correlated with worse
outcome (Figure 4A&B). Consistently, across all cohorts, expression of
clusters 21, 0 and 26 was not significantly predictive of outcome (data not
shown). High expression of Cluster 30 was associated with poor outcome only
in ovarian cancer (Figure 4C; p=0.007). We then used the PMCC dataset to
show that these survival differences were not associated with the AJCC stage
of the patients, which in itself is a known prognostic factor for GC (Table S8).

These observations suggest that the metagene which may be driven by M2
polarization has a prognostic effect. Similarly other cancer intrinsic features,
such as cell cycle, or other stromal-responses do not appear to be reflective
of overall patient survival.

Discussion
Gene signatures derived from gene expression data have previously been
used by ourselves and others to diagnose GC [36], predict patient prognosis
[19, 37, 38] and distinguish cancer stage and grade [39]. Indeed, through our
collaborators we have previously published on a stromal signature derived
from a Singapore cohort and validated on a group of Australian samples from
our laboratory [19]. A number of stromal signatures have been described by
others in different cancer contexts [30, 31, 40], which are composed of clearly
have distinct gene lists. Indeed, when those gene expression signatures are
derived from whole tumour mass, it is impossible to exclude the possibility that the genes are expressed by the cancer, and not the stromal cells. We have previously addressed this concern by profiling xenografted human cancer cells grown in a mouse host [33]. Bioinformatic analysis allowed a species specific profiling of both the host and cancer cell within a heterogenous tumour sample, but, such experiments do not confirm that the observation is physiologically relevant in the human patient body. The study described here used a discovery based approach to identify clusters of co-expressed genes specifically co-segregated with gastric cancers in a more comprehensive data set. The most significant cluster in terms of co-expression and number of differentially expressed genes was SRC. We termed this the “stromal-response” cluster as it is rich in genes which appeared to be derived from the stromal elements of the tumor as determined by annotation, ISH and LCM experiments, and was specifically associated in a juxatumoral microenvironment and is not found in benign conditions. This is a key advantage of our study, which better distinguishes what genes are expressed from what cell types.

The SRC comprises genes that are predominantly involved in the normal process of wound healing. These complex coordinated processes involve extensive intercellular cross-talk between stromal cells (macrophages, endothelial cells, lymphocytes, fibroblasts), and the epithelial cells. Normal wound healing proceeds in three steps, initially mounting an innate immune rejection of foreign invaders, often associated with injury and inflammation [41-43]. This is followed by proliferation and tissue remodeling. These stages
involve angiogenesis [17, 44], ECM deposition, and stem cell recruitment [17, 45, 46]. Genes involved in the processes of matricellular remodeling, cell movement and angiogenesis are highly represented in the stromal-response cluster supporting the well established paradigm that chronic wounds and/or inflammation are predisposing factors for tumorigenesis [42, 47, 48].

The multicellular nature of wound healing impairs the identification of a single critical stromal cell that would coordinate this reaction in cancer. In prostate cancer it is recognized that carcinoma associated fibroblasts (CAFs) may influence tumor behavior by encouraging growth of the primary tumor [45]. In models of breast cancer it was found that the stromal components derived from mesenchymal stem cells play an important role in divesting metastatic properties to the primary cancer [49]. Macrophages are known to play a role in cancer development and progression. In the early stages of tumor initiation M1 macrophages are believed to be activated in response to inflammation and illicit Th1/Th17 responses, a pathway which was found to be associated with cluster 26. Tumor associated macrophages (TAMs) are a primary source of EGF in breast cancers [43, 50, 51] and hence putatively growth promoting. Cancer related inflammation, when innate immune cells are mobilized to assist in tumor formation, has recently been described as a tumor enabling consequence which acts in support of the multiple hallmarks of cancer by providing the necessary factors and enzymes to aid invasion, metastasis, angiogenesis and EMT [52].
Our data showed a strong correlation between SRC expression and the presence of macrophages, more specifically those of the M2 subclass when all samples were considered. This suggests that M2 macrophages may act as a master regulator of the stromal response. There are examples in human cancer where the immune infiltrate in a carcinoma is a beneficial factor and engenders good prognosis. For instance, lymphocytic infiltrates of breast, ovarian and colorectal cancer [41, 53] are all positive prognostic variables. On the other hand, the complex interaction between CAFs and M2 macrophages and tumor cells has been attributed to an increase in cancer cell motility and ultimately in poor prognosis in prostate cancer [54]. Our findings show a ubiquitously expressed signature of genes derived primarily from tumor stroma that appears to have negative prognostic implications for both gastric and ovarian cancer.

It is conceivable that using the stromal-response signature as a marker for M2 activity in patient biopsies provides tangible criteria for the selection of patients to treat with anti-inflammatory drugs. M2 macrophages are known to secrete IL10, PDGF and TGFβ cytokines. Each of those cytokines is a potential target for novel cancer drugs, making those drugs potential tools in the improvement of patient outcome, by eliminating the observed stromal-response signature in suitable patients. Thus, the association between the recognized M2 macrophage polarity and this signature sheds light on the physiological and clinical significance of M2 macrophages. Our data suggests M2 polarized macrophages may influence stromal-response to the presence
of cancer cells in patients, and consequently may accelerate cancer cell survival and attenuate cancer cell immune surveillance.

Acknowledgements

The authors would like to thank Professor David Bowtell for useful discussions and advice.

Grant support

This work was supported by NHMRC grant 288714. AK acknowledges support by National ICT Australia (NICTA). NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program.

References


Figure legends

**Figure 1 (A)** Heat map derived from a Peter Mac gastric dataset of 139 samples run on Affymetrix U133+2 arrays. Hierarchical clustering shows segregation of samples based on HOPACH clusters of genes. Relative positions of the top 9 clusters representing genes with a cluster membership value >0.3 are highlighted. Red and green areas indicate up or down regulated genes respectively. **(B)** Data from the same 139 cases showing only Cluster 27 were reclustered using Eisen Clustering and show clear segregation of tumors and normals.

**Figure 2 (A)** GeneGo® schematics illustrating the top ranking pathway for the HOPACH generated gene list associated with cluster 27. This pathway is primarily associated with ECM and cell adhesion as has been termed the “stromal-response” cluster. Thermometer icon to the right of the gene name indicates altered expression, with red and blue indicating up and down regulation, respectively. **(B)** The top 10 GeneGo pathways associated with cluster 27 (based on p-value) are enriched for cell adhesion related processes. **(C)** Laser Capture Microdissection (LCM) of 4 gastric and 5 ovarian cancer samples where stroma (Str) and epithelium (Epi) compartments were dissected and profiled on Affymetrix U133+2 arrays independently. Stromal-response genes show higher expression in stroma of gastric and ovarian cancers compared to the epithelial compartments. **(D)** Venn diagram showing the interaction of genes overexpressed in gastric epithelium and stroma with the stromal-response cluster. Here both epithelial and stromal components of gastric cancers are intersected with stromal-response genes and many (n=60) do not fall into epithelial or stromal
compartments specifically suggesting they may be shared between them. (E) Venn diagram of the Stromal-Response Genes (Cluster 27) with ovarian stroma and epithelium enriched genes. In this instance 42 genes are equally expressed by both compartments.

**Figure 3** Correlations between ISH and IHC in GC samples. (A) In situ hybridisation using SPARC probes as a surrogate of the stromal-response cluster was performed on a TMA. These images show SPARC ISH and H&E staining a representative pair of normal and tumor samples derived from the same patient. In all cases on the TMA mRNA appears to be generated from stromal component of the tumor. (B) Correlation of in situ hybridisation and immunohistochemistry staining of SPARC and cell markers of activated fibroblasts (anti-SMA), endothelial cells (anti-CD31), macrophages (anti-CD68) and leucocytes (anti-CD45) was performed. Representative images from one tumor sample is shown. The summary of data analyzed by Fisher’s exact test is found in Table 2. (C) To determine whether the cluster 27 was more associated with M1 or M2 macrophages all samples in the cohort were classified as Cluster 27 high or low based in the formula median +/- median absolute deviation. The expression levels of M1 and M2 associated genes was determined and expressed, for each gene as the log fold change of the ratio of cluster 27 high versus low. (D) Correlation of SPARC ISH and anti-CD163 a key marker of the M2 phenotype.

**Figure 4** Kaplan Meier analysis of gene expression of specific clusters using a metagene approach. Cluster-wide metagene values for each gene cluster was calculated for each sample, within three cancer cohort datasets; PMCC GC (left), AOCS ovarian cancer (centre), and Ooi et al GC (right). Samples
were divided into those below (Low) and above (High) the median value of the metagene. Plots compare the overall survival of patients from the cluster high and low groups, for clusters (A) 27 (B) 28 and (C) 30. Patients with high expression of clusters 27 and 28 do significantly worse with shorter disease free survival in both diseases. Cluster 30 is only prognostic in patients with ovarian cancer and not GC. Statistical analysis was performed using log rank test.
Figure 4

**Cluster 27**

- **PMCC_GC**: p=0.038
- **AOCS_Ovarian**: p=0.0004
- **Ooi_GC**: p=0.0004

**Cluster 28**

- **PMCC_GC**: p=0.023
- **AOCS_Ovarian**: p=0.0002
- **Ooi_GC**: p=0.031

**Cluster 30**

- **PMCC_GC**: p=0.203
- **AOCS_Ovarian**: p=0.007
- **Ooi_GC**: p=0.634
Table 1. Top-ranking HOPACH clusters and differentially expressed genes in PMCC_gastric dataset

<table>
<thead>
<tr>
<th>Cluster</th>
<th>MainGeneGo biological process or network</th>
<th># Genes in cluster</th>
<th># Diff. exp Genes</th>
<th>% Diff. exp Genes</th>
<th>p-value</th>
<th>FDR*</th>
</tr>
</thead>
<tbody>
<tr>
<td>cluster_27</td>
<td>Cell adhesion: ECM remodelling (Stromal response)</td>
<td>100</td>
<td>83</td>
<td>83.00</td>
<td>3.75E-61</td>
<td>5.25E-60</td>
</tr>
<tr>
<td>cluster_21</td>
<td>Cell cycle: The metaphase checkpoint</td>
<td>77</td>
<td>67</td>
<td>87.01</td>
<td>1.42E-51</td>
<td>9.92E-51</td>
</tr>
<tr>
<td>cluster_12</td>
<td>Androsterodione and testosterone biosynthesis</td>
<td>73</td>
<td>58</td>
<td>79.45</td>
<td>2.32E-40</td>
<td>1.08E-39</td>
</tr>
<tr>
<td>cluster_0</td>
<td>Blood coagulation</td>
<td>68</td>
<td>54</td>
<td>79.41</td>
<td>1.35E-37</td>
<td>4.72E-37</td>
</tr>
<tr>
<td>cluster_14</td>
<td>Regulation of metabolism: Bile acid</td>
<td>64</td>
<td>45</td>
<td>70.31</td>
<td>1.22E-27</td>
<td>3.41E-27</td>
</tr>
<tr>
<td>cluster_26</td>
<td>Immune response: IL17 signalling</td>
<td>34</td>
<td>27</td>
<td>79.41</td>
<td>4.00E-19</td>
<td>9.34E-19</td>
</tr>
<tr>
<td>cluster_28</td>
<td>Cell cycle regulation of G1/S</td>
<td>91</td>
<td>40</td>
<td>43.96</td>
<td>1.00E-14</td>
<td>2.00E-14</td>
</tr>
<tr>
<td>cluster_15</td>
<td>Niacin_HDL metabolism</td>
<td>37</td>
<td>23</td>
<td>62.16</td>
<td>6.84E-13</td>
<td>1.20E-12</td>
</tr>
<tr>
<td>cluster_30</td>
<td>Immune response: Alternative complement pathway</td>
<td>71</td>
<td>28</td>
<td>39.44</td>
<td>2.21E-09</td>
<td>3.44E-09</td>
</tr>
</tbody>
</table>

Clusters in italics indicate those with >60% total genes differentially expressed between normal and tumour.
Table 2. Association between SPARC and IHC markers

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>Tumour</th>
<th>Tumour only</th>
<th>SPARC ISH</th>
<th>p value (Fisher’s exact test)</th>
<th>p value (Fisher’s exact test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1+</td>
<td>2+</td>
<td>3+</td>
<td>0 1+ 2+ 3+</td>
<td></td>
</tr>
<tr>
<td>α-CD163</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0</td>
<td>0 0 0 0</td>
<td>p=0.95</td>
</tr>
<tr>
<td>α-CD68</td>
<td>23</td>
<td>8</td>
<td>9 4</td>
<td>1+</td>
<td>1 6 6 9</td>
<td>p=0.66</td>
</tr>
<tr>
<td>α-CD44</td>
<td>11</td>
<td>8</td>
<td>5 2</td>
<td>2+</td>
<td>2 0 5 5</td>
<td></td>
</tr>
<tr>
<td>α-CD45</td>
<td>7</td>
<td>3</td>
<td>2 1</td>
<td>3+</td>
<td>3 1 2 2</td>
<td></td>
</tr>
<tr>
<td>α-SMA</td>
<td>0</td>
<td>1</td>
<td>0 1</td>
<td>0</td>
<td>0 0 1 0</td>
<td>p=3.2E-06</td>
</tr>
<tr>
<td>α-SMA</td>
<td>37</td>
<td>13</td>
<td>4 3</td>
<td>1+</td>
<td>6 10 4 3</td>
<td>p=0.046</td>
</tr>
<tr>
<td>α-SMA</td>
<td>3</td>
<td>7</td>
<td>9 3</td>
<td>2+</td>
<td>1 3 9 3</td>
<td></td>
</tr>
<tr>
<td>α-SMA</td>
<td>0</td>
<td>0</td>
<td>2 2</td>
<td>3+</td>
<td>0 0 2 2</td>
<td></td>
</tr>
<tr>
<td>α-SMA</td>
<td>2</td>
<td>0</td>
<td>0 0</td>
<td>0</td>
<td>2 0 0 0</td>
<td>p=0.005</td>
</tr>
<tr>
<td>α-SMA</td>
<td>35</td>
<td>11</td>
<td>3 0</td>
<td>1+</td>
<td>6 6 3 0</td>
<td>p=0.031</td>
</tr>
<tr>
<td>α-SMA</td>
<td>3</td>
<td>8</td>
<td>8 5</td>
<td>2+</td>
<td>3 8 8 5</td>
<td></td>
</tr>
<tr>
<td>α-SMA</td>
<td>1</td>
<td>1</td>
<td>5 2</td>
<td>3+</td>
<td>1 1 5 2</td>
<td></td>
</tr>
<tr>
<td>α-SMA</td>
<td>0</td>
<td>1</td>
<td>0 0</td>
<td>0</td>
<td>0 0 0 0</td>
<td>p=0.615</td>
</tr>
<tr>
<td>α-SMA</td>
<td>13</td>
<td>5</td>
<td>0 0</td>
<td>1+</td>
<td>0 1 0 0</td>
<td></td>
</tr>
<tr>
<td>α-SMA</td>
<td>3</td>
<td>3</td>
<td>9 3</td>
<td>2+</td>
<td>1 3 5 3</td>
<td></td>
</tr>
<tr>
<td>α-SMA</td>
<td>2</td>
<td>3</td>
<td>4 0</td>
<td>3+</td>
<td>0 2 4 0</td>
<td></td>
</tr>
</tbody>
</table>
Clinical Cancer Research

A signature predicting poor prognosis in gastric and ovarian cancer represents a coordinated macrophage and stromal-response.

Rita A. Busuttil, Joshy George, Richard W. Tothill, et al.

Clin Cancer Res  Published OnlineFirst March 21, 2014.