Complement membrane attack and tumourigenesis: A systems biology approach.

Laurence D. Towner*, Richard A. Wheat, Timothy R. Hughes and B. Paul Morgan,

From the Complement Biology Group, Division of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, UK. *Current address: Thomson Reuters IP & Science, London, UK.

Running Title: Complement membrane attack and tumourigenesis

To whom correspondence should be addressed: Prof B. Paul Morgan, Systems Immunity URI, Division of Infection and Immunity, School of Medicine, Cardiff University, Henry Wellcome Building, Heath Park, Cardiff CF14 4XN, UK. Telephone: (+44) 2920687096. E-mail: MorganBP@cardiff.ac.uk

Keywords: cancer biology, complement, gene expression, membrane attack complex, systems biology, tumor promoter.

ABSTRACT

Tumour development driven by inflammation is now an established phenomenon but the role that complement plays remains uncertain. Recent evidence has suggested that various components of the complement (C) cascade may influence tumour development in disparate ways; however, little attention has been paid to that of the membrane attack complex (MAC). This is despite abundant evidence documenting the effects of this complex on cell behaviour, including cell activation, protection from/induction of apoptosis, release of inflammatory cytokines, growth factors and ECM components and regulators and the triggering of the NLRP3 inflammasome. Here we present a novel approach to this issue by using global gene expression studies in conjunction with a systems biology analysis. Using network analysis of MAC responsive expression changes we demonstrate a cluster of co-regulated genes known to have impact in the extracellular space and on the supporting stroma and with well-characterized tumour promoting roles. Network analysis highlighted the central role for EGFR activation in mediating the observed responses to MAC exposure. Overall, the study sheds light on the mechanisms by which sublytic MAC causes tumour cell responses and exposes a gene expression signature that implicates MAC as a driver of tumour progression. These findings have implications for understanding of the roles of C and the MAC in tumour development and progression which in turn will inform future therapeutic strategies in cancer.

Inflammation is now well established as a crucial contributor in the development and progression of tumours; indeed, it has been included among the second generation hallmarks of cancer (1). A key player in inflammatory responses is the complement (C) system, an innate immune effector with important roles in defence against infection. C provides recognition, early warning signals and the initial fast response upon exposure to foreign organisms and has evolved to amplify the response to the initial signal (2). C comprises three activation pathways which converge on a common terminal pathway at the stage of C5 cleavage; release of a 74 amino acid peptide C5a, which has potent anaphylatoxic and chemotactic activities (3), leaves the large fragment, C5b, to form the nidus of a membrane-associated complex. Sequential recruitment of C6, C7, C8 and multiple C9 molecules creates a membrane-spanning pore-like cylindrical protein structure known as the membrane attack complex (MAC). The MAC can cause osmotic lysis of certain susceptible bacteria and of metabolically inert cells (4); however, lysis of self-cells is restricted by a combination of regulatory proteins, ion pumps and MAC removal processes (5). Non-lytic MAC triggers numerous activation events in
cells that likely contribute to the pro-inflammatory activity of C (6).

C has been strongly implicated as an effector in tumour clearance over the past 20 years, largely because of the success of monoclonal antibody (mAb)-based immunotherapies, many of which are designed to harness C as an effector to cause killing of tumour cells (7). In this context, the mAb triggers overwhelming C activation and tumour cell destruction; however, the role of C in tumour clearance in the absence of an activating mAb is much less clear. Indeed, it has been suggested that C activation has a tumour promoting role in many malignancies (8). C activation is known to occur on tumour cells both in vitro and in vivo in many malignancies, including breast (9), papillary thyroid (10,11), colorectal (12) and ovarian (13) cancers. The best evidence implicating a C activation product as a promoter of tumour development was provided by the demonstration that locally generated C5a recruits myeloid-derived suppressor cells (MDSCs) into the tumour where they suppress the anti-tumour activity of CD8+ T-cells (14,15). Others have implicated C5a as a factor influencing the balance between tumour promotion and tumour clearance (16), while both C3a and C5a have been shown to cause proliferation in tumour cells, for example in neuroblastoma (17,18). Studies in knockout mice lacking C3 or C4 confirm important roles for C in tumourogenesis, tumour growth being restricted in both C3 and C4 knockouts (14). Despite the large and growing body of evidence supporting a tumour promoting role of C activation (18-21), the role of MAC in tumour biology has been neglected.

Most tumours express, and indeed often over-express membrane bound C regulators CD55, CD59 and CD46 (22). As a consequence, although C is activated in the tumour micro-environment, activation will be restricted and thus terminal pathway activation and MAC deposition may be insufficient to kill the tumour cell. Nevertheless, MAC deposition on tumour cells at a sublytic level may have a profound impact on the target, for example by causing an immediate increase in intracellular Ca2+ (23) and downstream activation of signalling cascades (24). Effects of sublytic MAC on cells in vitro include: release of inflammatory mediators such as ROS/RNS, leukotrienes, arachidonic acid metabolites and prostaglandins (5); the release of cytokines such as IL-1, TNF α, IL-8, IL-6 and MCP-1 (25); increased expression of adhesion molecules such as E-selectin, ICAM-1, VCAM-1 and ELAM-1 (26); release of growth factors such as bFGF, PDGF, EGF, PIGF and RANTES (27,28); secretion of extracellular matrix (ECM) components such as collagen IV and fibronectin (29) and regulators such as MMP2 and MMP9 (30); increased cell proliferation (31); accelerated or inhibited apoptosis (32-34) and activation of the NLRP3 inflammasome (35). Given this catalogue of effects on cell function, it is likely that sublytic MAC will significantly influence tumour cell fate in vivo. Here we take a novel approach to addressing how MAC influences tumour cell fate by adopting an unbiased systems analysis of the effects of sublytic MAC on the patterns of gene expression in a tumour cell, we identify key pathways implicated and discuss the impact that these might have on tumour survival.

RESULTS

Sublytic C attack and MAC inhibition on tumour cell lines- The C5-binding protein OmCI has been extensively characterised and shown to specifically block formation of MAC in human and rodent plasma (36,37). The dose of OmCI required to completely block MAC formation was titrated by assessing inhibition of haemolysis of ShEA exposed to pNHS, an assay where target haemolysis is absolutely dependent on MAC formation (Fig1A). At 10 μg/ml, OmCI caused complete inhibition of pNHS induced haemolysis. This dose was used in all subsequent experiments. The sensitivity of each of the selected tumour cell lines, CT26 and B16, to pNHS-induced CDC was determined in a calcein release assay immediately prior to each experiment. Both CT26 and B16 cells were efficiently killed by pNHS without need for sensitisation, calcein release correlating with dose of pNHS (Fig1B). The dose of pNHS causing <10% specific calcein release at 1 hour was chosen as the maximum sublytic dose for the subsequent experiments. The 1 hour timepoint was chosen based on our previous work showing that MAC killing of nucleated cell targets is an acute event and does not increase further with prolonged incubation (6). In order to identify MAC-specific effects, cells were also exposed to the same
sublytic dose of pNHS but preincubated with the inhibitory dose of OmCI (10 μg/mL) to block MAC formation.

Preliminary qPCR experiments were carried out on RNA harvested from CT26 cells used in the above experiments in order to validate targets for subsequent microarray expression analyses and determine optimum time points for RNA collection. Initially, RNA was harvested at 1 hour and expression of osteopontin (OPN), a candidate gene chosen based on evidence from the literature (38), was measured by qPCR. Relative expression of OPN increased more in response to treatment with pNHS compared to OmCI-treated pNHS after 1 hour exposure (Fig1C). To further refine the time points of exposure to sublytic MAC prior to RNA harvest, the experiment was repeated for 6 or 12 hour time points, calculating the expression change in relation to untreated cells. OPN expression increased significantly at 6 (p<0.01) and 12 (p<0.05) hours in MAC-exposed cells compared to the OmCI control (Fig. 1D). For expression analysis, we chose to use 1 hour as an initial time point to capture an immediate response and 12 hours to capture sustained changes.

Global gene expression analysis of sublytic C-For microarray analysis cells were exposed to sublytic C attack and compared to a MAC-inhibited control as established in the optimisation experiments described above: CT26 cells seeded at 1.6x10⁶/mm² were exposed to 5% pNHS with or without addition of OmCI, then incubated for 1 or 12 hours (4 replicates of each condition in wells of a 12-well plate) prior to harvest of RNA. RNA was also harvested from untreated control cells (4 replicates) to establish a baseline. A total of 16 samples were used for microarray analysis; 3 from each of the 4 serum conditions and all 4 untreated controls. Quality control of Microarray data was performed using principal components analysis (PCA), a method used to compress a high content dataset, enabling its description with a limited number of contributors to variation. PCA allows the effect of experimental parameters on the data to be observed and identifies data quality issues. Initial PCA showed a batch effect which was eliminated using the Partek batch remover tool (www.partek.com) to reveal the most important components (Fig. 2A). Data were plotted to explore the contributions of the top three components; PCA#1 and 3 were best correlated with time point and serum exposure, and presence of OmCI respectively. Batch removal was not retained for downstream analysis because the robustness of the ANOVA model used rendered it unnecessary.

To better visualise the impact of experimental conditions, scatter plots of median baseline adjusted data were prepared to compare pNHS to pNHS + OmCI at 1 and 12 hours (Fig 2B(i) & (ii)). This graphical representation revealed that, for most parameters, expression changes were more apparent 1h after sub-lytic attack than at 12h, indicating that in this system most of the changes were transient in nature. To measure significance of differential gene expression in cells exposed to pNHS with or without MAC blockade, a 2-way ANOVA model was applied using method of moments (39). Gene lists were prepared using the ANOVA generated fold change (FC) and p-values to identify the most significantly differentially expressed (up or down) genes at each time point (Table 1).These show those genes which altered expression significantly (p<0.05), filtered to include genes changing by greater than 2-fold at 1 hour or 1.5-fold at 12 hours; different filters were chosen to reduce the disparity in number of differentially expressed genes at these timepoints. Five genes were upregulated and 1 downregulated at 1 hour post-exposure, while two genes were upregulated 12 hours post-exposure with none downregulated. This difference in the number of genes differentially regulated between the timepoints supports the trends shown in the scatter plots, substantially more points falling outside the set confidence intervals at 1 hour compared to 12 hours (Fig 2B).

Identification of secreted effectors induced by sublytic C- To provide functional insight into the data a new gene list was prepared using less stringent thresholds for inclusion by applying the following filters to the ANOVA statistics: FC > 1; unadjusted p-value cut-off <0.05 for both 1 hour and 12 hour comparisons. This latter filter selected for MAC-induced upregulation events that were apparent at both 1 and 12 hour time points, representing sustained changes. To understand the interactions between these genes the list was interrogated using MetaCore network building tools to automatically map genes to a
representative component termed the network object (NO). The ‘Shortest Path’ algorithm was selected and canonical pathways included for network building; to interpret the resulting network, nodes were arranged to identify the starting NOs and their overlapping connections. NOs not connected to the main network were removed and those groups displaying little connectivity to the larger network were pruned. The network was then organised by cellular location from top to bottom (Fig. 3). The analysis revealed four key highly interrelated NOs, co-location from top to bottom (Fig. 3). The analysis showed that the starting 11 NOs are well connected and therefore hidden. The network contained 11 of the total 12 starting NOs; FAM110C, ITPRIP and HBB-BH1 were unconnected and therefore hidden. The network generated using the gene lists identified above as secreted effectors during network analysis were selected (Table 2), together with four genes identified as showing the most significant differential expression (up or down at either time point) between MAC-exposed and MAC-inhibited pNHS-exposed CT26 cells when stringent thresholds were applied (Table 1). FAM110C (encoding Family With Sequence Similarity 110, Member C) and RGS16 (encoding Regulator of G-protein Signalling 16) were both increased at 1 hour post-attack; IRF1 (encoding Interferon Regulatory Factor 1) was decreased at 1 hour; HBB-BH1 (encoding Hemoglobin Z, beta-like embryonic chain) was increased at 12 hours. For each of these eight genes, qPCR was performed twice, first on cDNA prepared from RNA extracted for the microarray experiment and second on RNA from a fresh replication experiment. The qPCR expression data was replicated in these experiments and largely confirmed the expression patterns found in microarray for these same genes; data are presented together for comparison (Fig 4 and 5). In a few instances, qPCR and microarray data did not completely replicate: CXCL1 showed highest upregulation at 1 hour by microarray but at 12 hours by qPCR; RGS16 peak upregulation at 1 hour in microarray was confirmed by qPCR using the same original RNA but in RNA from the second experiment, further upregulation was seen at 12 hours post exposure. Despite these minor differences, the data strongly correlated, confirming the capacity of the microarray to accurately detect expression changes.

To explore whether the observed expression changes were cell-type specific, RNA extracted from MAC-exposed and control B16 melanoma cells was analysed by qPCR for expression of the four network identified hits, AREG, MMP3, MMP13 and CXCL1. Expression of MMP3 RNA was negligible in this cell type. Expression of RNA for both AREG and CXCL1 was markedly increased in MAC-exposed cells at both 1 and 12 hours, replicating the results obtained in CT26 cells (Figure 5). MMP13 RNA expression was low in B16 cells and not significantly different between MAC exposed and control cells.

Inter-connectivity of secreted effectors and regulatory genes- In an effort to identify pathways and mechanisms by which MAC effect ed changes in expression of the identified genes, a model was developed that analysed the combination of those genes identified as significantly differentially expressed under stringent statistics and those identified by network analysis as downstream secreted factors. Figure 6 shows the ‘Shortest Path’ network generated using the gene lists shown in Tables 1 and 2 as input. The main hub of the network contained 11 of the total 12 starting NOs; FAM110C, ITPRIP and HBB-BH1 were unconnected and therefore hidden. The network shows that the starting 11 NOs are well connected with a central triangle containing EGR1, EGR2 and IRF1, suggesting they are key drivers of the gene expression response to MAC. NOs added by the algorithm included AP-1 transcription factor subunit c-JUN, several NFkB subunits, the glucocorticoid receptor-alpha, c-Myc and the epidermal growth factor receptor (EGFR). Ap-1 and NFkB are the only two transcription factor NOs connected to all 4 secreted NOs validated by qPCR.

Transcriptional regulation network - In order to interrogate the data further and explore transcription regulation patterns from a greater
number of data points, a gene list was generated by applying an FDR adjusted p-value cut-off of <0.05 to the ANOVA statistics for the pNHS vs. pNHS+OmCI comparison. This was applied separately for 1 hour and 12 hour time points and the two lists combined to identify the most significant MAC induced expression changes at either time point and regardless of direction. This new list was combined with that created to generate figure 3 and the entire genelist uploaded to MetaCore (https://portal.genego.com/). The list was used as the starting list for the Analyze Network (transcription factors) (AN(tf)) algorithm applied using the default settings.

The AN(tf) algorithm identifies transcription factors for which there are enriched numbers of targets in the starting list, then uses the list to find the shortest path back to a receptor for which there are ligands in the starting list, thereby creating networks for each transcription factor, ranked by significance, based on enrichment of starting NOs via calculated g-scores, z-scores and p-values. z-score indicates the saturation of starting NOs, the g-score is a modified z-score describing the number of Canonical Pathways used to build the network; p-Value assesses the probability of the number of starting NOs falling on the generated network by chance accounting for the total number of NOs in the network and in the entire database (40).

The network with the highest g-score, z-score and smallest p-value was selected: g & z-scores = 187.12, p-value = 7x10^{-211}. To assist in interpretation, the network was organised by aligning the most connected NOs in the centre and placing the remaining NOs by protein class and in context around these main hubs (Fig. 7). With c-Myc and CREB1 as the main controlling transcription factors, EGFR and TrkB are introduced as non-seed nodes to the network as putative receptor starting points with EGFR the most interacting of the two. Other important TF hubs include c-jun, p53, ESR1, and Oct3/4. All four secreted effectors identified in microarray are present and the NO with most direct connections with these is c-Jun. Other signalling molecules with high connectivity include AKT and ERK1/2. The network contains 6 of the 8 validated genes. Overall the AN(tf) network includes 67 of the 118 candidate objects and provides evidence for a central role of EGFR activation by sublytic MAC.

**DISCUSSION**

The role of C as a tumour promoter has attracted a great deal of attention over the last few years because of evidence for significant C activation in diverse tumours. MAC is suspected to be influential given its published activatating and proliferative effects on nucleated cells (5,6,14); however, signalling mechanisms underlying many of these effects remain ill defined. We took a novel approach to understanding the role of the MAC, taking advantage of an available terminal pathway inhibitor, global gene expression technology and systems biology methodology. Sublytic conditions were optimised using pNHS as a C source and OmCI to block terminal pathway activation. CT26 colon carcinoma cells were selected as a model tumour cell and MAC-specific gene expression changes mapped by microarray, qPCR and network analysis. These approaches revealed a gene expression pattern in tumour cells exposed to sublytic MAC which could significantly impact cell survival and proliferation as well as reshape surrounding ECM. The key findings were replicated in an unrelated tumour cell line, B16 melanoma.

Statistical analysis of array data comparing MAC exposed cells with controls confirmed a set of expression changes; genes involved in Ca\(^{2+}\) and G-protein signal transduction (ITPRIP, RGS16), early response transcription factors (EGR1, EGR2), and inflammatory responses (IRF1) were significantly altered. Network analysis to map the interactions of genes upregulated at the 1 hour and 12 hour time points highlighted 4 further expression changes in genes encoding proteins with extracellular localization, AREG, CXCL1, MMP3 and MMP13 genes; these were co-regulated by putative canonical signalling cascades including PKC, PI3k/AKT, JNK, Erk1/2 and p38.

The product of the AREG gene is the amphiregulin protein (AR), an EGF-like ligand capable of triggering erbB2 activation (41). CXCL1 ligand is a potent neutrophil chemoattractant, important in infection and signals via CXCR2, a G-protein coupled receptor (42). MMP3 and MMP13 both code for matrix metalloproteinases (MMPs) that function in extracellular matrix regulation and remodelling; they are important during development, wound healing, proliferation and inflammation (43).
Validation by qPCR confirmed that the expression changes highlighted in gene array and supported by statistics and network building were real and robust. Critically, expression changes for two of these genes, AREG and CXCL1 closely replicated in an unrelated tumour cell line, B16. MMP expression in this line was extremely low so could not be replicated.

Network generation using these and the remaining statistically significant changes found interactions between all identified genes apart from FAM110C and highlighted the central importance of IRF1, EGR1, and EGR2 in mediating the changes. EGFR is noteworthy in that it is placed in between two starting NOs, Amphiregulin and Rgs16, the three connected in an extracellular to nuclear direction, supportive of EGFR activation. Rgs16 is further connected to Eg1, Eg2 and Ir1 via c-Myc. Eg1 protein is known to positively regulate EGR2 gene expression, while Eg2 protein negatively regulates EGR1 (44,45); Eg1 protein is reported to inhibit IRF1 expression (46). Ap-1 and NFkB were both highlighted as possible transcriptional regulators in the network. The Ap1 and NFkB complexes are known to regulate MMP3, MMP13 and CXCL1 gene expression in mouse and human cells (47-52).

Ap-1 (a heterodimer of c-fos and c-jun) and NFkB are known to be responsive to MAC (53) and have been cited as important regulators of the response to sublytic C (6). In particular, c-fos is upregulated rapidly in MAC-exposed cells and is linked to Ca\textsuperscript{2+} flux and MAPK (particularly ERK) involvement (54). MMP3 and MMP13 upregulation has been described in MAC-attacked chondrocytes in human disease and an experimental model of osteoarthritis (55). Each of the canonical signalling cascades identified in the network, PKC, PI3k/AKT, JNK, Erk1/2 and p38 have been reported to be activated in cells exposed to sublytic MAC (6).

The larger network generated using genes significantly changed either at 1 or 12 hours exposure alongside those upregulated at both time points provided greater insight into mechanisms responsible for observed gene expression changes in CT26 cells. The network placed EGFR central in the response, an assignment supported by the presence of this receptor in all three generated networks (Fig 7). EGFR is a member of the erbB receptor tyrosine kinase (RTK) family, activated by ligand binding at the cell surface triggering phosphorylation of the intracellular tyrosine kinase domain (56). Activation of the EGFR system has also been described in response to cellular stressors such as UV, osmotic and oxidative stress (57). MAC may cause analogous stress responses; indeed, there is evidence that it can induce expression of the EGF ligand and cause EGFR signalling activation without ligand binding (58). The response may involve G\textsubscript{i} protein activation independent of receptor, which is known to be activated by MAC (59). Indeed, our data showing an RGS16 expression response to MAC supports this assertion; RGS16 gene expression is induced as a feedback mechanism for G-protein signalling (60). MAC can transactivate several other RTKs, including fibroblast growth factor receptor-2 (FGFR2), and hepatocyte growth factor receptor (HGF) (58). Potentiation of EGFR activation may come from MMP cleavage and release of EGF family ligands such as HB-EGF at the cell surface, a pathway supported by our demonstration of AREG upregulation (61).

Increased expression of AREG, CXCL1, MMP3 and MMP13 is described at mRNA and protein levels in a number of human cancers such as breast, colorectal, ovarian and pancreatic (62-69). This increased expression of the four effector genes often correlates with tumour development and aggressivity, and can be predictive of patient prognosis (70-72). Their activities promote cell proliferation, activation and motility through various mechanisms. AREG contributes to tumourigenesis via its function as a growth factor, the development of autocrine or juxtacrine loops that promote cell proliferation and survival, and increased cell motility (63,71). CXCL1 acts through recruitment of myeloid-derived suppressor cells to the tumour microenvironment where they support tumour growth and metastasis and suppress the local immune response (70). MMP family members contribute through their role in regulating the ECM (73), promoting angiogenesis, tumour invasion and metastasis (72). In addition they cleave and activate molecules in ECM which promote proliferation, motility and induce alterations in adhesion (74). Increased expression of MMPs is associated with poor tumour differentiation, increased invasiveness,
poor prognosis, increased likelihood of metastasis and shorter survival time (72). In particular, MMP3 promotes tumour progression by releasing/activating E-cadherin, L-selectin, HB-EGF and TNFα and is described as a central mediator of mammary tumourigenesis (69). MMP3 is reported to induce a stable epithelial to mesenchymal transition, a process which is closely linked to tumour development (75). MMP13 contribution to tumour promotion is mainly through its pro-angiogenic activity, increasing vascular density in tumours (76). Together, alterations in the expression of these four genes represent a powerful influence on tumour development. Induction of expression of these genes in response to MAC may therefore indicate a tumour promoting role.

A role for C as a tumour promoting system has recently gained mainstream recognition (19). The work presented here represents a novel approach to uncover this relationship, using global expression data and systems biology analysis to explore both the mechanisms and the characteristics of such a response. The approach has provided evidence to suggest that MAC deposition which does not result in cell lysis is a potent tumour cell activator leading to significant changes in gene expression in several critical and interlinked pathways. These data fit well with published piecemeal studies in diverse cell types. The work not only sheds light on the signalling cascades and responsive transcription factor systems that respond to MAC but also reveals a downstream gene expression response to MAC which will alter tumour behaviour through induction of proliferative, migratory and survival pathways. Interestingly, a central role for the EGFR system was identified, although it was not clear whether this was activated directly by MAC or indirectly following MAC exposure. Overall, this work provides additional evidence implicating sublytic MAC in tumour cell activation and has implications not only for our understanding of the tumour promoting effects of C but also for new approaches to cancer therapy.

**EXPERIMENTAL PROCEDURES**

**Materials**- Pooled normal human serum (pNHS) was obtained from whole blood collected from consenting volunteers. Blood was placed in 20 mL glass vials and allowed to clot. Serum was separated by centrifugation, pooled, 0.22 µm filtered and stored in aliquots at -80°C. Sheep erythrocytes (ShE) in Alsever’s solution were purchased from TCS Biosciences (Buckingham, UK). Complement fixation diluent (CFD) was from Oxoid (Basingstoke, UK). Anti-ShE antiserum (Amboceptor) was from Siemens (Forchheim, Germany). CT26 mouse colon carcinoma and B16 mouse melanoma cell lines were from American Type Culture Collection (ATCC, Manassas VA, USA). RPMI 1640 medium (RPMI), fetal bovine serum (FBS) and additives and calcein-AM were obtained from Invitrogen (Paisley, UK). Complete medium comprised RPMI 1640 with 5% heat-inactivated FBS. All other chemicals were from Sigma Aldrich (Gillingham, UK).

**Preparation of antibody sensitised sheep erythrocytes (ShEA)-** ShE were washed into CFD and re-suspended at 4% (v:v) in CFD (10ml total volume) at 37°C. Amboceptor, diluted 1:2000 in 10ml of CFD, was mixed with the ShE suspension and incubated for 30 minutes at 37°C. The resultant ShEA were washed and diluted to 2% (v:v) in CFD.

**Haemolytic assay-** Haemolytic activity in pNHS was used as a measure of MAC formation, assessed by incubating (37°C for 60 min) triplicate serum dilutions in CFD with an equal volume of ShEA in wells of a 96-well plate. No serum (CFD alone) and 100% lysis (CFD containing 0.1% triton-X-100) controls in triplicate were included. Plates were spun, supernatant transferred to a flat bottomed 96 well-plate and absorbance measured at 410 nm using a FLUOstar OPTIMA plate reader (BMG Labtech, Aylesbury, UK). Percentage haemolysis was calculated using the equation:

\[
\% \text{ Lysis} = \left( \frac{A_{\text{complement release}} - A_{\text{spontaneous release}}}{A_{\text{detergent release}} - A_{\text{spontaneous release}}} \right) \times 100
\]

To titrate the effect of the C5 inhibitor Ornithodoros moubata C inhibitor (OmCI; gift of Dr. Miles Nunn) on MAC formation and haemolytic activity, aliquots of pNHS were pre-incubated with different doses of OmCI prior to measurement of haemolysis as above.

**Complement-directed cytotoxicity (CDC) assay-** We chose the well-described calcein release assay to measure tumour cell killing. The cell-
permeant calcein AM is taken into cells and trapped by de-esterification to calcein; release of calcein from the cells then correlates with lytic cell death. CT26 cells or B16 cells were grown as monolayers in complete medium to 80% confluence in 75 cm² TC flasks then washed in saline and harvested by incubation in 10mM EDTA in PBS (30min). Harvested cells were washed in RPMI, diluted to a density of 5x10⁵ cells per mL in complete medium, aliquoted at 100 µL/well into flat-bottomed 96 well-plates, and incubated for 16 hours at 37°C, 5% CO₂. Adherent cells were washed and 100 µL complete medium containing 2 µg/mL calcein AM was dispensed into each well. Plates were incubated for 1 hour at 37°C, 5% CO₂. Calcein-loaded cell monolayers were washed twice in RPMI, then pNHS dilutions (0-40% in 100 µL RPMI) dispensed directly into wells and incubated for a further 1 hour at 37°C, 5% CO₂. Supernatants were transferred to fresh 96-well plates and fluorescence measured (EX 485nm EM 520nm) in a Fluostar Optima plate reader. Remaining cells were lysed by addition of 100 µL 0.2 % Triton-X-100 in RPMI per well and released fluorescence measured as above. Percentage lysis was calculated using the following equation (FI = fluorescence intensity):

\[ \text{Percentage lysis} = 100 \times \left( \frac{FI_{\text{complement release}}}{FI_{\text{complement release}} + FI_{\text{detergent release}}} \right). \]

**Titrating sublytic C attack-** CT26 or B16 cell monolayers were washed with saline then incubated (37°C, 5% CO₂, 1h or 12 h) with pNHS at a dilution previously titrated to give less than 10% lysis at 1 hour, a timepoint when maximum lytic killing has been reached, either in the presence or absence of a dose of OmCI that completely inhibited haemolytic activity [10 µg/ml]. Monolayers were washed in RPMI and RNA harvested using the Genelute Mammalian Total RNA Miniprep Kit (Sigma-Aldrich).

**Global gene expression analysis-** RNA concentration and quality were measured using the Agilent 2100 Bioanalyzer (Agilent Technologies, Stockport, UK) and global gene expression analyses performed on the Illumina Microarray platform (Illumina, Saffron Walden, UK; Cardiff University Central Biotechnology Services). Amplification of material to generate cRNA and labelling was carried out according to manufacturer’s instructions. Hybridisation experiments were performed using the mouse ref8v2 BeadChips (2 X 8 samples) and analysed using the iScan Reader and Control Software (Illumina). GenomeStudio Expression Module software (Illumina) was used to convert signal intensity data into expression data. Data were normalized using the quantile method and log transformed [77]. Principal components analysis (PCA) and primary statistical analysis were performed using PartekGenomics Suite version 6.6 (build 6.13.0213, Partek Inc., Missouri, USA) and graphical representations obtained using GeneSpring 12.0 GX (Agilent Technologies). Pathway analysis was performed using MetaCore software (Thomson Reuters, London, UK).

**QPCR analysis of mRNA** - To validate differences in the relative expression of genes of interest implicated from gene expression analysis, extracted RNA samples (1 µg) were reverse transcribed using TaqMan Reverse Transcription Reagents (Applied Biosystems, Paisley, UK) following manufacturer’s instructions; resulting cDNA was stored at -80°C. For qPCR, sufficient cDNA for triplicate reactions of each primer pair diluted 1/10 in ultrapure H₂O, was mixed with 1x SYBR Green Jump Start Readymix (Sigma-Aldrich, Dorset, UK) according to manufacturer’s instructions. Reaction mixes were aliquoted into 48-well white PCR plates, sealed with optical flat 8-cap strips (Bio-Rad, Hertfordshire, UK) and placed in a MiniOpticon Real-Time PCR System (Bio-Rad) controlled using the Opticon Monitor 3.1 software. Thermocycling was adjusted from the manufacturer’s protocol (65°C annealing temperature and 40 cycles) to take account of relative expression, assessed using the ΔΔCₜ method where calculated Cₜ was the cycle number at which fluorescence crossed a threshold level selected as the point where PCR expansion was linear in all samples. The mean Cₜ values for the housekeeping genes β-actin and Polr2A were assessed, ΔΔCₜ calculated for each sample and results expressed as a percentage of the control.

**Acknowledgements:** We thank members of the Complement Biology Group for their support and input. This work was funded by a Cardiff University School of Medicine PhD Studentship to LDT.
Conflict of interest: The authors declare that they have no conflicts of interest relevant to the contents of this article.

Author contributions: BPM & TH conceived the idea for the project. LDT conducted most of the experiments and analyzed the results. RAW performed confirmatory experiments with additional cell lines. BPM and TH contributed to experimental design and data analysis at all stages. BPM, TH and LDT contributed to writing of the manuscript.

REFERENCES


Complement membrane attack and tumourigenesis


FOOT NOTES
The abbreviations used are: C, Complement; MAC, membrane attack complex; ECM, extracellular matrix; ShE, sheep erythrocytes; CFD, complement fixation diluent; pNHS, pooled normal human serum; CDC, complement dependent cytotoxicity; AN(tf), analyze network (transcription factor); NO, network object; EMT, epithelial mesenchymal transition, EGFR, epidermal growth factor; MMP, matrix metalloproteinase.
FIGURES LEGENDS

FIGURE 1. Optimization of sublytic complement conditions. A. Haemolytic activity testing lysis of ShEA by serum with or without 10 µg/mL OmCI (C5 blocker), titrated from 16% down to 0%. B. Susceptibility of CT26 cells to C lysis; cells in a monolayer were loaded with calcein AM then exposed to serum for 1h at 37°C. Lysis was calculated from the release of calcein into the supernatant and expressed as the percentage of the total entrapped calcein obtained by detergent lysis of the cells. Results are means of 4 separate experiments +/- SEM. C&D. Expression analysis of OPN in CT26 cells in response to exposure to sublytic C for 1, 6 and 12 hours. CT26 cells were exposed for 1 (i), 6 and 12 (ii) hours to 5% serum treated with or without a MAC-blocking dose of OmCI and OPN gene expression analysed by qPCR. Expression was calculated as % of untreated control (ii). Results are means of 3 determinations +/- SEM (*p<0.05, **p<0.01).

FIGURE 2. Primary microarray data analysis. A. PCA plot of top three principle components. Three dimensional plot showing the top 3 principle components of the microarray data as calculated using principle component analysis (PCA). Contributing principle components (PC) are labelled on each axis alongside the calculated % contribution to overall variation. Each sample from the experiment is represented by a coloured sphere; red=control, green=pNHS, blue=pNHS+OmCI. A centroid sphere shows how these samples are grouped according to their experimental conditions; black=control, pale blue=OmCI at 1hr, darker blue=OmCI at 12hrs, light green=pNHS at 1hr and darker green=pNHS at 12 hours. B. Scatter plot comparisons between samples exposed to pNHS and pNHS+OmCI at 1 (i) and 12 (ii) hours. Log2 transformed, median baseline adjusted data. Expression is presented as distribution around a median that represents equal gene expression in the two conditions. The parallel flanking lines represent gene expression changes of +/- 1.3-fold change; data points falling outside these lines are considered to be differentially expressed. Data points are coloured according to their expression levels (median baseline adjusted) upon exposure to pNHS; green=below median, red=above median.

FIGURE 3. Network analysis of overlap gene list. List includes genes upregulated by sublytic MAC at both time points. The network was generated in MetaCore using the following options: ‘Shortest Path’ network building algorithm with a maximum of 2 steps, and inclusive of canonical pathway; this latter option allows sequences of interactions that occur frequently in the cell to be counted as single steps in the shortest path. The network describes the interconnected regulation of upregulated genes and highlights four key downstream effector genes. The network is organised so that nodes are organized by the sub-cellular localization of their products, from extracellular to nuclear. Nodes present in the input list are in blue circles. Thick light blue lines highlight the various canonical pathways of signal transduction and transcription regulation. Seed nodes are circled in navy blue, lines represent interactions, either transcriptional regulation or protein-protein associations; red=inhibition and green=activation.

FIGURE 4. qPCR validation of statistically significant hits. Microarray=original microarray data, Primary Validation=RNA extracted in parallel with that used in microarray, Secondary validation=RNA extracted in a fresh sublytic attack experiment. RNA was reverse transcribed and FAM110C, RGS16, IRF1 and HBB-BH1 gene expression analysed by qPCR and calculated as expression relative to housekeeping genes β-actin and Pol2ra using the ΔΔCt calculation then presented as % of untreated control. Results are means of 3 determinations +/- SEM (*p<0.05, **p<0.01, ***p<0.001).

FIGURE 5. qPCR validation of network identified hits. Microarray=original microarray data, Primary Validation=RNA extracted in parallel with that used in microarray, Secondary validation=RNA extracted in a fresh sublytic attack experiment. B16 validation= RNA extracted from fresh sublytic attack experiment using the B16 mouse myeloma cell line (MMP3 message was not significantly detected in this cell line). In all cases RNA was reverse transcribed and MMP3, MMP13, CXCL1 and AREG gene expression analysed by qPCR and calculated as expression relative to housekeeping genes β-actin and
Pol2ra using the ΔΔCt calculation then presented as a % of untreated control. Results are means of 3 determinations +/- SEM (*p<0.05, **p<0.01, ***p<0.001).

**FIGURE 6.** Network analysis of collated gene list. Network describing the interconnected regulation of the four key downstream effector genes as well as the 8 statistically significant genes. The network was generated in MetaCore using the following options: ‘Shortest Path’ network building algorithm with a maximum of 2 steps, excluding canonical pathways. Network is organised to show cellular localisation from extracellular (top) to nucleus (bottom). From the list, ITPRIP, FAM110C and HBB-BH1 are not represented on the network due to lack of connectivity. Seed nodes are circled in navy blue, lines represent interactions either transcriptional regulation or protein-protein associations; red=inhibition and green=activation.

**FIGURE 7.** Network of transcriptional regulation and ligand receptor signalling. This network uses a greater number of genes, including all those identified as significantly changed by combining significantly differentially expressed genes with those significantly upregulated at both time points. The network was generated in MetaCore using the following options: ‘Analyze Network (Transcription Factors)’ network building algorithm with ‘Add ligands and TF targets’ selected. The algorithm generates a list of possible networks, with scores based on the number of seed nodes to non-seed nodes and the presence of canonical pathway threads. The network with the highest score for these two factors was selected and manually organised, first showing the four validated genes, and then the most highly connected objects regardless of functional type, to their right. The remaining objects were sorted by function so that TFs, kinases, phosphatases, generic proteins and binding proteins were from left to right. Ligands and receptors were placed to the far left. Seed nodes are circled in navy blue, predicted receptor trigger is highlighted in green, predicted controlling TF is highlighted in red, lines represent interactions either transcriptional regulation or protein-protein associations; red=inhibition and green=activation.
**Tables.**

**TABLE 1.** Gene list of top most significant expression changes when comparing pNHS with pNHS+OmCI at 1 hour and 12 hours. List criteria: FDR adjusted p-value <0.05 A. pNHS 1 hour relative to pNHS+OmCI 1 hour, fold-change (FC)> 2 (upregulated) and B. FC< -2 (down regulated), C. pNHS 12 hour relative to pNHS+OmCI 12 hour, FC> 1.5. Entities are sorted by p-value and FC displayed (pNHS relative to pNHS+OmCI).

<table>
<thead>
<tr>
<th>Probeset ID</th>
<th>Symbol</th>
<th>Transcript</th>
<th>p-value (pNHS * 1 vs. OmCI * 1)</th>
<th>Fold-Change (pNHS * 1 vs. OmCI * 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_2623983</td>
<td>Egr2</td>
<td>ILMN_212209</td>
<td>4.57E-11</td>
<td>2.08117</td>
</tr>
<tr>
<td>ILMN_2600744</td>
<td>Rgs16</td>
<td>ILMN_209950</td>
<td>3.03E-09</td>
<td>2.74759</td>
</tr>
<tr>
<td>ILMN_2995794</td>
<td>Itprip</td>
<td>ILMN_242056</td>
<td>6.49E-08</td>
<td>2.16089</td>
</tr>
<tr>
<td>ILMN_2662926</td>
<td>Egr1</td>
<td>ILMN_215729</td>
<td>9.47E-08</td>
<td>2.37517</td>
</tr>
<tr>
<td>ILMN_2757634</td>
<td>Fam110c</td>
<td>ILMN_222981</td>
<td>1.89E-07</td>
<td>2.07236</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probeset ID</th>
<th>Symbol</th>
<th>Transcript</th>
<th>p-value (pNHS * 1 vs. OmCI * 1)</th>
<th>Fold-Change (pNHS * 1 vs. OmCI * 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_2834777</td>
<td>Irf1</td>
<td>ILMN_209850</td>
<td>5.15E-10</td>
<td>-2.19138</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probeset ID</th>
<th>Symbol</th>
<th>Transcript</th>
<th>p-value (pNHS * 12 vs. OmCI * 12)</th>
<th>Fold-Change (pNHS * 12 vs. OmCI * 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_1260323</td>
<td>Akr1c18</td>
<td>ILMN_215518</td>
<td>4.16E-07</td>
<td>1.61755</td>
</tr>
<tr>
<td>ILMN_1225816</td>
<td>Hbb-bh1</td>
<td>ILMN_216390</td>
<td>1.99E-07</td>
<td>1.58676</td>
</tr>
</tbody>
</table>

**TABLE 2.** Gene list entities. Areg, Mmp13, Mmp3, Cxcl1 are sorted by p-values at 1 hour and 12 hours and fold change displayed (pNHS relative to pNHS+OmCI).

<table>
<thead>
<tr>
<th>Probeset ID</th>
<th>Symbol</th>
<th>Transcript</th>
<th>p-value (pNHS * 1 vs. OmCI * 1)</th>
<th>Fold-Change (pNHS * 1 vs. OmCI * 1)</th>
<th>p-value (pNHS * 12 vs. OmCI * 12)</th>
<th>Fold-Change (pNHS * 12 vs. OmCI * 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILMN_1238547</td>
<td>Areg</td>
<td>ILMN_217903</td>
<td>2.31E-06</td>
<td>1.43899</td>
<td>0.00185142</td>
<td>1.18086</td>
</tr>
<tr>
<td>ILMN_2737685</td>
<td>Mmp13</td>
<td>ILMN_210384</td>
<td>1.06E-06</td>
<td>1.52378</td>
<td>0.0100312</td>
<td>1.14512</td>
</tr>
<tr>
<td>ILMN_2753809</td>
<td>Mmp3</td>
<td>ILMN_219123</td>
<td>9.89E-05</td>
<td>1.23906</td>
<td>0.00367792</td>
<td>1.08968</td>
</tr>
<tr>
<td>ILMN_2763245</td>
<td>Cxcl1</td>
<td>ILMN_223377</td>
<td>8.74E-03</td>
<td>1.55636</td>
<td>0.0215031</td>
<td>1.45109</td>
</tr>
</tbody>
</table>
Figure 1.
Figure 2.

A.  

B. (i)  

(iii)
Figure 3.
Figure 4.
Figure 5.
Figure 6.
Figure 7.