International Journal of Genomics and Proteomics
ISSN: 0976-4887 & E-ISSN: 0976-4895, Volume 3, Issue 1, 2012, pp.-50-56
Available online at http://www.bioinfo.in/contents.php?id=38

PROTEOMIC STUDY OF CONDITIONED MEDIA: CANCER BIOMARKER DISCOVERY

SHREEVE R.L., BANKS R.E., SELBY P.J. AND VASUDEV N.S.*

Cancer Research UK Centre, Leeds Institute of Molecular Medicine, St James's University Hospital, Leeds, UK.
*Corresponding Author: Email- N.Vasudev@leeds.ac.uk

Received: February 23, 2012; Accepted: March 09, 2012

Abstract- Cancer biomarkers are integral to current oncological practice and have the potential to impact all aspects of patient care. In particular, markers that allow accurate risk stratification and optimal choice of therapy for individual patients are emerging as central to improving patient outlook. Whilst proteomic technologies have improved significantly in recent years, serum proteomics continues to pose challenges in overcoming its vast dynamic range of protein concentrations. Profiling conditioned media (CM) of cancer cell lines provides an attractive alternative, since it is potentially enriched for tumour derived secreted proteins and, as such, represents a rich source of potential biomarkers. The approach has been widely adopted by researchers across all aspects of cancer marker discovery. Here, we highlight promising examples of such work and discuss the potential pitfalls associated with studying CM. Ultimately, it is imperative that candidate markers are adequately validated in annotated clinical specimens to render the research worthwhile.
Key words- Conditioned media; biomarker; cancer; secretome

Citation: Shreeve R.L., et al. (2012) Proteomic Study of Conditioned Media: Cancer Biomarker Discovery. International Journal of Genomics and Proteomics, ISSN: 0976-4887 & E-ISSN: 0976-4895, Volume 3, Issue 1, 2012, pp.-50-56

Introduction

Cancer is a multifaceted disease, usually with many somatic and some germline (in the case of familial cancers) genetic changes resulting in an abnormal proteome and phenotype. It is notoriously difficult to detect and treat, despite notable advances in recent years. Clinical cancer management is a multi-stepped process involving diagnosis, tumour staging, prognosis, treatment selection, response assessment and monitoring for disease recurrence. The potential for biomarkers to impact on all of these aspects of care is significant; most solid tumours are curable if detected at an early stage. In addition, oncological practice must be cost-effective and the ability to identify high-risk patients and / or those likely to respond to a given expensive treatment is important.

The search for biomarkers that would provide detailed information on diagnosis, prognosis, and disease monitoring has remained largely elusive. Relatively few cancer biomarkers are in current clinical use. Most, including proteins such as cancer antigen 125 [1], alpha-fetoprotein [2] and carcinoembryonic antigen (CEA) [3] are used in the longitudinal monitoring of patients for cancer recurrence and / or their response to treatment. Others, such as human epidermal growth factor receptor 2 (HER2) and the oestrogen receptor are used in treatment selection. However, very few markers have demonstrated sufficient sensitivity and specificity to be useful for screening at a population level.

The ideal tumour marker would be easily measured, economically viable and both highly sensitive and specific in its ability to discriminate for its target phenotype [4]. Traditionally, biomarker research has focussed on interrogation of tissue, using well-established and relatively straightforward techniques such as immunohistochemistry. Fluid-based biomarkers represent an attractive alternative, with major potential benefits [5,6]. Firstly, they are much easier to obtain, obviating the need for invasive biopsies. Secondly, biofluids are often in abundant supply and thirdly, fluid biomarkers may allow earlier detection of disease. Serum in particular, though less specific than proximal tissue fluids, may contain proteins from most if not all tissues and has

been estimated to contain tens of thousands of proteins [7,8]. Fluid biomarker research however is fraught with problems; for example the protein content in serum covers 10 orders of magnitude, with 14 proteins accounting for 94% of the total content. Nevertheless, the fact that cells secrete factors to communicate with and modify their surrounding environment is well established [9]; factors including cytokines, growth factors, extracellular components and others, may enter more accessible fluids and lead researchers into tapping their potential utility.

Presently the dynamic range of mass spectrometry is 10^5 orders of magnitude, considerably short of the range observed in serum, though sequential levels of fractionation can increase depth of coverage. Immunodepletion or enrichment techniques are popular in serum proteome analysis [10,11] although there is a risk of depleting desirable proteins along with high abundance proteins [12]. Notwithstanding methods to increase the dynamic range, our attempt to reach low abundance proteins is still largely thwarted by the dominance of a limited number of high abundant proteins [8,13]. Indeed, a recent study which combined 91 serum datasets generated a reference set of just 1929 proteins [14]. Despite including the largest number of datasets to date, it is likely this is still only a small fraction of the entire serum proteome.

Using conditioned medium (CM) as a model system may assist in the problem of identifying low to moderate abundance proteins that are most likely to include candidate biomarkers. Proteomic analysis of the media in which the cells have been residing may reveal proteins which have been secreted or shed [15]. For example interleukin-6 was shown to be present at concentrations 100-fold higher in adipose interstitial fluid as compared to plasma [16]. Additionally, a plethora of cell lines exist, making mimicry of a spectrum of disease phenotypes possible. Further, the CM is of a lower complexity than biological fluids, thus concentrating the proteins for a proteome of interest. Proteins may leave the cells and enter the surrounding fluids via a classical or a non-classical route (Figure 1). In addition, some proteins will be shed from the plasma membrane whilst others are likely to be present as a result of protein leakage through cell death. Between 10-25% of the genome has been estimated to code for secreted proteins [4,15] and approximately 17% of all human proteins in UniProt have a signal sequence [17]. Proteins coded with an N-terminal signal sequence exit via the classical secretion route with synthesis directed to the endoplasmic reticulum (ER), resulting in packaging through the endocytic pathway [18] (Figure 1). The proteins are then either secreted or inserted in the cell membrane; this may vary dependent on whether the signal peptide carries single nucleotide polymorphisms (SNPs) [19]. The mechanisms of non-classical secretion remain uncertain but it is apparent that the proteins, which lack a signal peptide, exit independently of the ER and golgi, appearing to follow one of four routes: direct translocation across the plasma membrane; exosome secretion; plasma membrane blebbing and endosomal fusion with the plasma membrane [20]. Non-classical secretion has been demonstrated by transport of green fluorescent protein out of Chinese hamster ovary cells [21], although it has not been confirmed if these routes of exit are physiological phenomena or the result of *in vitro* stress [22]. Exosomes, formed in multivesicular bodies, were previously thought to contain debris but more recently have been shown to contain proteins and microRNA, and may play a role in signalling and communication [23]. With multiple modes of cellular exit, the potential for a variety of protein species to co-exist in fluids is clear and CM may provide a simplified paradigm to aid biomarker discovery.

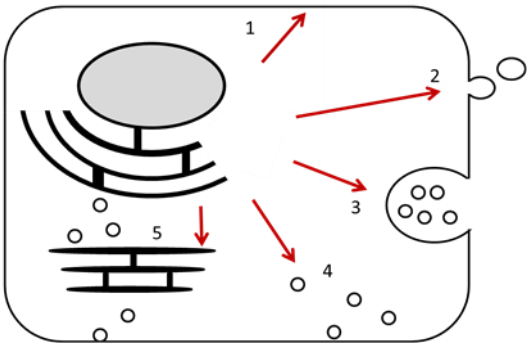

Fig. 1- Cellular secretion occurs through classical and non-classical methods. (1) direct translocation across the membrane; (2) plasma membrane blebbing; (3) endosomal fusion; (4) exosomal secretion; (5) classical secretion via the endoplasmic reticulum and golgi apparatus.

Preparation and Proteomic Analysis

No gold standard exists for the proteomic analysis of CM, with variation in the method at all stages and implications for comparisons between studies. The general workflow, outlined in Figure 2, consists of five main steps: collection; preparation; analysis; bioinformatics; validation. These steps should ideally be optimised for each individual cell line. For a more detailed description of the techniques involved the reader is directed to two recent reviews [24,25].

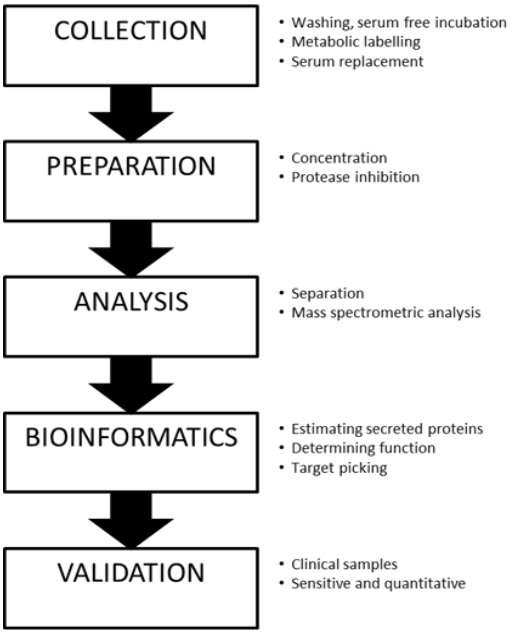

Fig. 2- Conditioned media analytical workflow. Key steps in the general workflow focus on collection, preparation, analysis, bioinformatics and validation. Due to the heterogeneity of cell lines, collection and preparation stages, in particular, should ideally be optimised for each cell line.

Collection

In addition to the secreted proteins, within the CM are the proteins already present within the cell media and potentially at much higher concentrations than endogenous proteins. The serum used to supplement cells contains xeno-components such as bovine serum albumin (BSA); the level of contamination by this protein alone may be significant [26] and may mask the proteins secreted by the cells. Thus CM is usually harvested following a period of serum starvation, typically between 24 and 48 hours. Methods to negate the effects of serum contamination focus largely on variations in washing steps and serum-free incubation, for which extensive optimisation studies have been performed [27,28]. Estimation of serum contamination has been performed by comparison of the 1D gel profile of varied fetal bovine serum concentrations with the CM profile, suggesting an estimated residual contamination of 0.004% following seven washes [29]. Due to the stress incurred by the cells during this period, it is necessary to estimate cell death to minimise contamination by cytosolic compounds. Commonly used approaches involve estimation of lactate dehydrogenase [30] or beta-tubulin [31] within CM. A novel approach to detect newly synthesised, genuinely secreted, proteins involved using a combination of *in vitro* metabolic labelling of cells and subsequent detection of proteins by fluorescence analysis and autoradiography [32], based on the premise that fluorescence detects all proteins present but autoradiography detects only those proteins which were synthesised by living cells during the metabolic labelling period. The data must, however, be interpreted with caution; CM was harvested after just six hours of incubation in labelled-medium and analysis would therefore exclude mature proteins synthesised before incubation. For example, GRP78 is a stable protein with a half-life longer than 24 hours [33] and has been shown to localise to the cell surface [34]. Another group used a roller bottle technique to culture prostate cancer cells in serum free media for a prolonged period (14 days) by using chemically defined CHO medium supplemented with glutamine [35]. Expression of two kallikrien markers, hK5 and hK6 were found to increase and plateau at 10 days, but were not found in cell pellets. A large proportion of intracellular proteins were identified suggesting a high proportion of cell death, however criteria for target picking was limited to extracellular or membrane proteins as determined by Gene Ontology (GO) annotations. This technique shows promise for the ability to negate masking and contamination by serum proteins but still requires optimising to limit cell death.

Preparation

Since large volumes of CM may prohibit direct in depth analysis, concentration is necessary to increase detection limits and a variety of techniques have been investigated including precipitation, ultracentrifugation, dialysis [28,36] and hexapeptide ligand libraries [37]. Two more novel methods to tackle this problem are nano-proteomics [38] and a hollow fibre culture (HFC) [39] system. Enrichment of the secretome using nanozeolites identified 1474 unique proteins from hepatocellular carcinoma secretome with a high degree of reproducibility [38]. Another group employed a hollow fibre culture (HFC) system used with nasopharyngeal carcinoma (NPC) cells adapted gradually to growth in serum-free medium [39]. The fibres provide a significant surface area of up to 2100cm^2 and contain 5kDa molecular weight cut off (MWCO) filters permitting movement of waste and nutrients into the fibres, but retaining most proteins and allowing concentration of the secretome. Cell lysis was estimated using Western blot detection of actin, G3PDH, HSP60 and tubulin and appeared to be very low (0.001-0.022%), however a large proportion (52%) of intracellular proteins were identified. This may be due to underestimation of proteins exiting via a non-classical route and those annotated with a dual localisation.

Analysis

Simplifying the proteome prior to proteomic analysis is important to allow low abundance proteins to be observed and some comparative optimisation has also been performed in this area [40]. Common methods include 1-Dimensional Polyacrylamide Gel Electrophoresis (1D PAGE), 2D PAGE and High-Performance Liquid Chromatography (HPLC). Problems with 2D PAGE, including low throughput and low dynamic range, have made this technique less commonly used. Differential gel electrophoresis (DIGE) provides a useful graphical representation but, as per standard 2D PAGE, contamination by albumin and transferrin can mask less abundant proteins. Separation by HPLC followed by MS/MS analysis (LC-MS/MS) allows for a more tailored experiment; reversed phase and strong cation exchange columns, individually or in combination, are commonly used in this field. Advancements in the sensitivity and precision of mass spectrometry, combined with the ability to detect post -translational modifications and to perform relative quantification are increasing the potential of this type of analysis. Alternatively, using an antibody array to identify secreted proteins has also been employed [41], but this technique must be used cautiously due to the heterogeneity of antibody specificity, and also the limitation on protein identification imposed by the requirement for an antibody.

Bioinformatics

The number of proteins reported to be identified in CM is frequently in the hundreds; delinearising the data to generate a more manageable list of potential targets is critical. Estimating the proportion of secreted proteins by using sophisticated prediction software such as SignalP, TMHMM and SecretomeP [42], which predict the presence of a signal peptide, transmembrane domains or secretion via a non-classical route respectively, can aid in shortening the list. For example, a combination of SignalP, TMHMM and protein databases used to detect secreted proteins, and then to exclude potential endoplasmic reticulum resident proteins, was used to shorten a list of 6255 potential targets to 319 [43]. Functional analysis and determination of subcellular location can also support target finding, by using a combination of GO annotations [44], database searching, literature searching and tools such as Ingenuity Pathways Analysis. In doing so, a plethora of information can be found, which also assists in determining biological relevance. Cross-referencing lists of identifications with published proteomes is a logical method of sorting [30,45,46], though interesting targets should not be eliminated entirely on this basis; it is possible they have not yet been detected in these proteomes. Other problems include a number of false positives and false negatives, for example, SignalP is not capable of determining those proteins that remain within the transport network [22]. Thus, reports in the percentages of observed secreted proteins vary considerably and may be linked to the prediction tools employed [47].

In short, these methods should be used with caution and not as a definitive source of inclusion or exclusion.

Validation

The tumour microenvironment is not fully represented in *in vitro* studies with a single cell type being examined and the potential for *in vitro* artefacts arising from culture ex vivo. Whether proteins observed in CM even enter biological fluids *in vivo* will depend upon multiple factors including their physical stability and sensitivity to endogenous proteases. To determine if a putative marker is clinically relevant, validation using clinical samples is critical. Western blotting, immunohistochemistry, multiple reaction monitoring (MRM) and particularly ELISA are sensitive and quantitative techniques suitable for this purpose and with the exception of MRM, these techniques can be performed routinely within the lab. Problems with validation mean this step is often omitted, thus stalling research on a target that may be of clinical significance. Insufficient samples, often of poor quality, are a key problem, as well as a lack of suitable antibodies or ELISAs.

Applications in Cancer Biomarker Discovery

Accessing proteins either secreted by the tumour or forming part of the tumour microenvironment may hold the key to cancer biomarker discovery. Conditioned media may be utilised in cancer research for either biomarker discovery or investigations into cancer specific pathways to better understand mechanisms of development, which in turn, will feed back into biomarker and therapeutic target discovery. Biomarkers may be diagnostic, prognostic or predictive and examples of where CM may deliver in each of these aspects are discussed briefly below.

Diagnostic Markers

It is generally accepted that early diagnosis is key to an increased chance of long-term survival [48]. To this end, many groups have focussed on finding diagnostic markers that will identify the presence of a tumour earlier than by standard methods. The current usefulness of diagnostic markers is limited since many tumour markers lack sufficient specificity and sensitivity. For example, prostate specific antigen (PSA), used in the detection of prostate cancer suffers from both false positive and false negative results, and is not recommended for routine population screening [49]. Significant amounts of CM research focus on diagnostic marker discovery, with recent studies in a variety of cancers including glioblastoma [50], thyroid [51], colorectal [52], breast [53], pancreatic [54], head and neck [55] and lung cancer [56]. None have as yet identified a marker with sufficient sensitivity and specificity to form the basis of a screening test. However some findings may be of relevance in the context of differential diagnosis. A recent study identified five putative markers of pancreatic cancer that together outperformed CA19.9 in receiver operating characteristic (ROC) curve analysis [57]. The authors evaluated the secretomes of six pancreatic cancer cell lines and a normal cell line, and then compared the overlap with the proteome of two pancreatic juice samples pooled from six patients. The targets, identified based on GO annotation, tissue specificity, differential expression and integration of different fluids, were anterior gradient homolog 2, collagen alpha-1 (VI) chain, olfactomedin-4, polymeric immunoglobulin receptor and syncollin, and have been preliminarily validated in serum [57]. Furthermore, 76% overlap of the proteome of cell line derived CM and the pancreatic juice samples was reported, providing confidence in the use of this model in secreted protein analysis. Utilising a longitudinal approach, early markers of lung cancer were investigated. Conditioned media was collected at three increasing passages from the lung cancer cell line M-BE, which has early tumourigenic features at high passage [58]. A total of 47 proteins were found to be passage dependent although only one, cathepsin D, was validated. Using plasma from 104 lung squamous cell carcinoma (SCC) patients, cathepsin D was shown to be significantly elevated in SCC as compared to 36 normal donors and 15 patients with non-malignant lung disease ($P \leq 0.015$) and also discriminated patients with lymph node metastasis (LNM) ($P \leq 0.038$). Corroborating the observed results, cathepsin D was also found to be significantly elevated in a tissue microarray of SCC versus normal patients ($P \leq 0.001$) [58].

Prognostic markers

Following diagnosis and staging of a cancer, prognostic markers may aid in projecting disease outcome, thus guiding intensity of therapy or surveillance. A potential urinary marker of renal cell carcinoma (RCC) prognosis was recently identified by our laboratory [59]. In a pilot study, ten secreted proteins were identified as upregulated in the CM of an RCC cell line in comparison to pooled normal renal epithelial CM. Cathepsin D was further validated in preoperative urine of RCC patients, using 239 samples including healthy and benign subjects, and found to be significantly associated with overall survival (hazard ratio: 1.33; $P = 0.005$). Prediction of metastatic potential is also important in stratifying patient follow-up and treatment and a study on nasopharyngeal carcinoma found serum levels of cystatin A acted as a prognostic indicator of nodal metastasis [60]. The authors compared the secretome of an NPC bone metastatic cell line to the transcriptome of 19 paired NPC tissues to find overlapping patterns of expression; four targets, cystatin A, cathepsin B, manganese superoxide dismutase and MMP-2, were validated in serum. Cystatin A was found to predict nodal metastasis in a cohort of 84 patients ($P = 0.046$), as well as demonstrating significance as an indicator of overall survival ($P = 0.025$)[60].

Predictive markers

Many cancer therapies suffer from marked toxicity. Combined with a spectrum of patient response and high economic burden, the need to identify responding patients early in therapy is evident. Historically there has been some success in this field. Breast cancer patients with amplifications of the HER2 gene respond well to trastuzumab (Herceptin) therapy; similarly if the oestrogen or progesterone receptors are over-expressed, patients are selected for anti-oestrogen based therapy [61]. At a genetic level, in lung cancer patients, mutations in the kinase domain of EGFR predict sensitivity to erlotinib or gefitinib, whereas mutations in KRAS predict resistance [61]. The use of CM as an approach to identify predictive markers has, in fact, received little attention, despite the suitability of this system to drug studies and the demonstration of a role for the extracellular matrix (ECM) in drug resistance [62]. A comparison of the secretomes of a thyroid cancer cell line in the presence and absence of two tyrosine kinase inhibitors, RPI-1 and dasatinib, revealed proteomic changes in response to drug treat-

ment [63]. Using this approach 8, 16 and 32 proteins showed sensitivity to RPI-1, dasatinib or both drugs together respectively, with potential application as early markers of response. Alterations in signalling pathways were noted, with a particular decrease in adhesion pathways in line with the anti-proliferative effect of these drugs, and an increase in metabolic and antioxidant pathways [63]. Another predictive study compared the secretomes of two breast cancer cell lines, one sensitive and one resistant to the chemotherapeutic drug doxorubicin, to determine if a mediator of resistance could be found in the secretome [64]. From a list of 89 proteins demonstrating two fold changes, the majority of which were ECM components, the authors focussed on IL-18. This protein was upregulated in the resistant line, was novel in doxorubicin resistance, had links to metastasis and appeared to function as a key control point in network analysis. The study went on to demonstrate that recombinant IL-18 increased resistance, and anti-IL-18 antibodies increased sensitivity, to doxorubicin [64].

Cancer pathways and therapeutic targets

The general mechanisms underlying cancer development are important to understand and such research supports both biomarker and therapeutic target discovery. For most solid tumours, the development of metastatic disease remains fatal in the majority of patients. Greater insights into the underlying mechanisms behind this are therefore critical in advancing cancer management and many groups have focussed on exploiting conditioned media in this type of research [65-68]. Invasion and metastasis are believed to be heavily reliant on the actions of extracellular proteins [69,70], of which matrix metalloproteinases play an important role both in ECM degradation and the resulting release of proangiogenic factors such as VEGF [70]. A field which has arisen out of this pathway is 'degradomics', which involves identifying proteolytic targets of extracellular proteases involved in invasion [71,72]. This approach has been used to identify potential targets of MMP-9 in prostate cancer, by comparing the CM of a prostate cancer cell line, PC-3ML, with its MMP-9 knock down counterpart [73]. The study found differential expression of 69 proteins, six of which were largely confirmed by Western blotting. Two proteins, PN-1 and leukaemia inhibition factor, were demonstrated to be cleaved *in vitro* by MMP-9 using recombinant proteins, though this data must be regarded carefully; *in vitro* proteolysis may not mimic that which occurs *in vivo* [74]. Another key step in cancer metastasis is epithelial mesenchymal transition (EMT), a process whereby dedifferentiation of neoplastic cells increases their migratory capability [75]. By induction of this process in madin-darby canine kidney (MDCK) cells and comparison of the secretome with the parent cell secretome, 70 extracellular proteins were shown to demonstrate differential expression [76] indicating the impact the transition has on the secretome. The authors noted that the EMT resulted in reduced expression of basement membrane proteins and elevated expression of proteases and a subset of ECM proteins [76].

Problems and Pitfalls

Whilst the benefits of CM as a facet of proteomic biomarker discovery are evident, there are distinct disadvantages. Working with cell lines generates a multitude of problems. Diseased, infected or stressed cells can provide artefactual results and it is also estimated that 18% of cell line studies are performed with contaminated cell lines [77]. Cell lines should be checked to ensure they are not contaminated using marker expression or preferably Short Tandem Repeat (STR) profiling, which measures the length, and number, of DNA STRs across multiple loci. Further, serum free incubation is inherently stressful for the cells with potential alterations in the rate of proliferation [78,79], apoptosis [80] and glycosylation [81,82]. Skottmann and colleagues demonstrated the large impact of serum withdrawal on the transcriptome of human Embryonic Stem Cells (hESCs), reporting 1417 differentially regulated genes between hESCs grown in the presence or absence of serum [83]. Monitoring for cell death, in combination with bioinformatic analysis and stringent validation is currently the most popular method for overcoming this.

Model cell line systems provide the opportunity to answer biological questions which otherwise may not easily be answered. Advantages lie in the ability to focus on small scale changes, but this over-simplification can also be a distinct disadvantage. One of the key drawbacks of CM is the inability of this system to fully mimic the tumour surroundings. The concept of a 'tumour microenvironment' is well documented with the interdependence of the multiple cell types [84,85]. Tumour cells are capable of modifying the surrounding stroma to further support their development [9], and thus CM studies must take this into account. Further, the internal proteome of the cells may have changed during propagation *in vitro*, and it is possible alterations to the *in vitro* secretome may be even more pronounced due to the absence of communicatory signals from other cells. Methods to circumvent this problem include analysing the secretomes of a panel of cell lines from a particular cancer [29,57,86], though this would still lose valuable information on co-communication. Co-culturing cells is another potential approach. A study using this technique investigated the tumour microenvironment of breast cancer cells by co-culturing fibroblasts and the 8701-BC breast cancer cell line; cancer cell proliferation was shown to increase by approximately 58% [87]. The authors went on to use fibroblast CM as a chemoattractant for the 8701-BC cells, resulting in increased migration and invasion [87], and demonstrating the impact of the local environment on the tumour growth. A similar observation was reported by Toillon and colleagues, researching the pro-apoptotic effect of normal breast cells on MCF-7 breast cancer cells [88]. Conditioned media from normal cells applied to cancer cells increased apoptotic rate in a concentration dependent manner. Mass spectrometric analysis identified IGFBP-3 and maspin as potentially responsible and immunodepletion of these proteins inhibited the apoptotic effect [88]. Another approach involved using CM from an invasive bladder cancer cell line as a chemoattractant for another, comparatively less invasive, bladder cancer cell line. Increased motility was observed and analysis of the CM identified multiple proteins linked to invasiveness including SPARC and clusterin [67], thus demonstrating a role for secreted proteins in cancer invasiveness.

Concluding Remarks

Despite intensive efforts, success in the cancer biomarker field has been limited. There has been a significant amount of research into finding markers of diagnostic value using CM as a model system; however these are yet to yield strong candidates. A single unique biomarker, of sufficient sensitivity and specificity, may not be a realistic goal, but instead a panel of markers may go some way into tackling the difficulties imposed by inter-individual varia-

tion. Proteomic research into prognostic and predictive markers is still limited, despite their importance - the ability to individually tailor treatment programs is an important goal in cancer management.

A candidate biomarker may be derived either from the tumour cell, surrounding cells or arise as a systemic response. Mining CM from cell lines provides an opportunity to focus specifically on the first of these. However *in vitro* models are, by their nature, open to artefacts and steps to address this are fundamental in advancing the usefulness of this research. In depth studies investigating the overlap between *in vitro* and *in vivo* systems would provide valuable information on the validity of CM studies as a model [89], although the 'holy grail' would be collection of the *in vivo* secretome itself. Proximal fluids collected directly from the site of origin may help bridge the gap. Analysis of nipple aspirate fluid revealed 64 proteins including many unique to this fluid [90]. Another group collected fresh tumour interstitial fluid from surgically resected breast tissue to identify proteins secreted by breast adipocytes, identifying 359 unique proteins [91]. In pursuit of the 'true' secretome, Huang and colleagues inserted capillary ultrafiltration (CUF) probes into the tumour mass of a live mouse to directly collect the secretome of the tumour during the progressive and recessive stages. Five proteins were identified at each stage [92,93].

Although current standards of research are high, there is a strong need for increased standardisation of the techniques used. Analysis of multiple cancer cell lines by individual groups are useful [63,94], since they lessen the effect of experimental variation, and have also indicated that there exists a degree of putative biomarker overlap between different cancers. More uniform methods to eliminate the masking and contaminating effects of serum proteins, or the toxic effects brought about by serum-free incubation, would enhance reproducibility and study compatibility, and would also allow deeper mining. Improved validation is also urgently required; a step which is often skipped due to a lack of clinical samples of sufficient quality, yet is of critical standing. The impact of this stretches into sample processing and collection which is beyond the scope of this review, but adequate infrastructure must in place within the laboratory to allow for this [95].

The advent of major advances in genomics presents an opportunity to refine approaches to the use of CM and proteomic approaches to identify novel shed biomarkers. Whole genomic sequencing of malignancies at various stages of evolution towards the invasive and metastatic phenotype provides precise indications of the somatic genetic changes which underpin the phenotype. Using such information to characterise and/or generate appropriate cell line models has the potential to elevate the value of the proteomic analysis of CM from such cell lines and will be an important step in the generation of novel candidate biomarkers for subsequent validation.

References

[1] Bast R.C., Feeney M., Lazarus H., Nadler L.M., Colvin R.B., et al. (1981) *J Clin Invest*, 68, 1331-1337.

[2] Abelev G.I., Perova S.D., Khramkova N.I., Postnikova Z.A., Irlin I.S. (1963) *Transplantation*, 1, 174-180.

[3] Gold P., Freedman S.O. (1965) *J Exp Med*, 121, 439-462.

[4] Kulasingam V., Diamandis E.P. (2008) *Nat Clin Pract Oncol*, 5, 588-599.

[5] Good D.M., Thongboonkerd V., Novak J., Bascands J-L., Schanstra J.P., et al. (2007) *J Proteome Res*, 6, 4549-4555.

[6] Hanash S.M., Pitteri S.J., Faca V.M. (2008) *Nature*, 452, 571-579.

[7] Anderson N.L., Anderson N.G. (2002) *Mol Cell Proteomics*, 1, 845-867.

[8] Archakov A.I., Ivanov Y.D., Lisitsa A.V., Zgoda V.G. (2007) *Proteomics*, 7, 4-9.

[9] Rønnov-Jessen L., Petersen O.W., Bissell M.J. (1996) *Physiol Rev*, 76, 69-125.

[10] Smith M.P.W., Wood S.L., Zougman A., Ho J.T.C., Peng J., et al. (2011) *Proteomics*, 11, 2222-2235.

[11] Millioni R., Tolin S., Puricelli L., Sbrignadello S., Fadini G.P., et al. (2011) *PloS One*, 6, e19603.

[12] Björhall K., Miliotis T., Davidsson P. (2005) *Proteomics*, 5, 307-317.

[13] Vasudev N.S., Ferguson R.E., Cairns D.A., Stanley A.J., Selby P.J., et al. (2008) *Proteomics*, 8, 5074-5085.

[14] Farrah T., Deutsch E.W., Omenn G.S., Campbell D.S., Sun Z., et al. (2011) *Mol Cell Proteomics*, M110.006353.

[15] Pavlou M.P., Diamandis E.P. (2010) *J Proteomics*, 73, 1896-1906.

[16] Sopasakis V.R., Sandqvist M., Gustafson B., Hammarstedt A., Schmelz M., et al. (2004) *Obes Res*, 12, 454-460.

[17] Wu C.H., Apweiler R., Bairoch A., Natale D.A., Barker W.C., et al. (2006) *Nucleic Acids Res*, 34, D187-D191.

[18] van Vliet C., Thomas E.C., Merino-Trigo A., Teasdale R.D., Gleeson P.A. (2003) *Prog Biophys Mol Biol*, 83, 1-45.

[19] Jarjanazi H., Savas S., Pabalan N., Dennis J.W., Ozcelik H. (2008) *Proteins*, 70, 394-403.

[20] Nickel W., Seedorf M. (2008) *Annu Rev Cell Devel Biol*, 24, 287-308.

[21] Tanudji M., Hevi S., Chuck SL. (2002) *J Cell Sci*, 115, 3849-3857.

[22] Hathout Y. (2007) *Expert Rev Proteomics*, 4, 239-248.

[23] Simpson R.J., Lim J.W., Moritz R.L., Mathivanan S. (2009) *Expert Rev Proteomics*, 6, 267-283.

[24] Dowling P., Clynes M. (2011) *Proteomics*, 11, 794-804.

[25] Skalnikova H., Motlik J., Gadher SJ., Kovarova H. (2011) *Proteomics*, 11, 691-708.

[26] Pellitteri-Hahn M.C., Warren M.C., Didier D.N., Winkler E.L., Mirza S.P., et al. (2006) *J Proteome Res*, 5, 2861-2864.

[27] Xue H., Lu B., Lai M. (2008) *J Transl Med*, 6, 52.

[28] Makridakis M., Vlahou A. (2010) *J Proteomics*, 73, 2291-2305.

[29] Schiarea S., Solinas G., Allavena P., Scigliuolo G.M., Bagnati R. et al. (2010) *J Proteome Res*, 9, 4376-4392.

[30] Sardana G., Jung K., Stephan C., Diamandis E.P. (2008) *J Proteome Res*, 7, 3329-3338.

[31] Mbeunkui F., Fodstad O., Pannell L.K. (2006) *J Proteome Res*, 5, 899-906.

[32] Zwickl H., Traxler E., Staettner S., Parzefall W., Grasl-Kraupp B., et al. (2005) *Electrophoresis*, 26, 2779-2785.

[33] Satoh M., Nakai A., Sokawa Y., Hirayoshi K., Nagata K. (1993) *Exp Cell Res*, 205, 76-83.

[34] Delpino A., Piselli P, Vismara D., Vendetti S., Colizzi V. (1998) *Mol Membrane Biol*, 15, 21-26.

[35] Sardana G., Marshall J., Diamandis E.P. (2007) *Clin Chem*, 53, 429-437.

[36] Cao J., Shen C., Zhang J., Yao J., Shen H., et al. (2011) *Sci China Life Sci*, 54, 34-38.

[37]Colzani M., Waridel P., Laurent J., Faes E., Rüegg C., et al. (2009) *J Proteome Res*, 8, 4779-4788.
[38]Cao J., Hu Y., Shen C., Yao J., Wei L., et al. (2009) *Proteomics*, 9, 4881-4888.
[39]Wu H-Y., Chang Y-HY-C., Liao P-C. (2009) *J Proteome Res*, 8, 380-389.
[40]Piersma S.R., Fiedler U., Span S., Lingnau A., Pham T.V., et al. (2010) *J Proteome Res*, 9, 1913-1922.
[41]Mustafa S.A., Hoheisel J.D., Alhamdani M.S.S. (2011) *Mol BioSystems*, 7, 1795-1801.
[42]Emanuelsson O., Brunak S., von Heijne G., Nielsen H. (2007) *Nat Protoc*, 2, 953-971.
[43]Bortoluzzi S., Scannapieco P., Cestaro A., Danieli G.A., Schiaffino S. (2006) *Proteins*, 62, 776-792.
[44]Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., et al. (2000) *Nat Genet*, 25, 25-29.
[45]Gunawardana C.G., Kuk C., Smith C.R., Batruch I., Soosaipillai A., et al. (2009) *J Proteome Res*, 8, 4705-4713.
[46]Wang C-L., Wang C-I., Liao P-C., Chen C-D., Liang Y., et al. (2009) *J Proteome Res*, 8, 4428-4440.
[47]Vasudev N.S. (2008) PhD. The Application of Proteomics to Biological Fluids for Biomarker Discovery in Renal Cell Carcinoma, University of Leeds.
[48]Etzioni R., Urban N., Ramsey S., McIntosh M., Schwartz S., et al. (2003) *Nat Rev Cancer*, 3, 243-252.
[49]Heidenreich A., Bellmunt J., Bolla M., Joniau S., Mason M., et al. (2011) *Eur Urol*, 59, 61-71.
[50]Polisetty R.V., Gupta M.K., Nair S.C., Ramamoorthy K., Tiwary S., et al. (2011) *J Proteomics*, 74, 1918-1925.
[51]Kashat L., So AK-C., Masui O., Wang X.S., Cao J., et al. (2010) *J Proteome Res*, 9, 5757-5769.
[52]Alvarez-Chaver P., Rodríguez-Piñeiro A.M., Rodríguez-Berrocal F.J., García-Lorenzo A., Páez de la Cadena M., et al. (2011) *J Proteomics*, 74, 874-886.
[53]Gromov P., Gromova I., Bunkenborg J., Cabezon T., Moreira J.M.A., et al. (2010) *Mol Oncol*, 4, 65-89.
[54]Chang Y-T., Wu C-C., Shyr Y-M., Chen T-C., Hwang T-L., et al. (2011) *PloS One*, 6, e20029.
[55]Ralhan R., Masui O., Desouza L.V., Matta A., Macha M., et al. (2011) *Proteomics*, 11, 2363-2376.
[56]Planque C., Kulasingam V., Smith C.R., Reckamp K., Goodglick L., et al. (2009) *Mol Cell Proteomics*, 8 (12), 2746-2758.
[57]Makawita S., Smith C., Batruch I., Zheng Y., Rückert F., et al. (2011) *Mol Cell Proteomics*, 10, M111.008599
[58]Lou X., Xiao T., Zhao K., Wang H., Zheng H., et al. (2007) *J Proteome Res*, 6, 1083-1092.
[59]Vasudev N.S., Sim S., Cairns D.A., Ferguson R.E., Craven R.A., et al. (2009) *Br J Cancer*, 101, 1175-1182.
[60]Chang K-P., Wu C-C., Chen H-C., Chen S-J., Peng P-H., et al. (2010) *Proteomics*, 10, 2644-2660.
[61]Sawyers C.L. (2008) *Nature*, 452, 548-552.
[62]Sethi T., Rintoul R.C., Moore S.M., MacKinnon C., Salter D., et al. (1999) *Nature Med*, 5, 662-668.
[63]Caccia D., Zanetti Domingues L., Miccichè F., De Bortoli M., Carniti C., et al. (2011) *J Proteome Res*, 10, 4196-4207.
[64]Yao L., Zhang Y., Chen K., Hu X., Xu L.X. (2011) *PloS One*, 6, e24684.
[65]Xue H., Lü B., Zhang J., Wu M., Huang Q., et al. (2010) *J Proteome Res*, 9, 545-555.
[66]Chiu K-H., Chang Y-H., Wu Y-S., Lee S-H., Liao P-C. (2011) *J Proteome Res*, 10, 1110-1125.
[67]Makridakis M., Roubelakis M.G., Bitsika V., Dimuccio V., Samiotaki M., et al. (2010) *J Proteome Res*, 9, 3243-3259.
[68]Ween M.P., Lokman N.A., Hoffmann P., Rodgers R.J., Ricciardelli C., et al. (2011) *Int J Cancer*, 128, 1570-1584.
[69]Nguyen D.X., Bos P.D., Massagué J. (2009) *Nat Rev Cancer*, 9, 274-284.
[70]Brooks S.A., Lomax-Browne H.J., Carter T.M., Kinch C.E., Hall D.M.S. (2010) *Acta Histochemica*, 112, 3-25.
[71]López-Otín C., Overall C.M. (2002) *Nat Rev Mol Cell Biol*, 3, 509-519.
[72]Doucet A., Butler G.S., Rodríguez D., Prudova A., Overall C.M. (2008) *Mol Cell Proteomics*, 7, 1925-1951.
[73]Xu D., Suenaga N., Edelmann M.J., Fridman R., Muschel R.J., et al. (2008) *Mol Cell Proteomics*, 7, 2215-2228.
[74]Karagiannis G.S., Pavlou M.P., Diamandis E.P. (2010) *Mol Oncol*, 4, 496-510.
[75]Thiery J.P. (2002) *Nat Rev Cancer*, 2, 442-454.
[76]Mathias R.A., Chen Y-S., Wang B., Ji H., Kapp E., et al. (2010) *J Proteome Res*, 9, 1007-1019.
[77]MacLeod R.A., Dirks W.G., Matsuo Y., Kaufmann M., Milch H., et al. (1999) *Int J Cancer*, 83, 555-563.
[78]Cooper S. (2003) *FASEB J,* 17, 333-340.
[79]Shin J-S., Hong S-W., Lee S-LO., Kim T-H., Park I-C., et al. (2008) *Int J Oncol*, 32, 435-439.
[80]Hasan N.M., Adams G.E., Joiner M.C. (1999) *Int J Cancer*, 80, 400-405.
[81]Gawlitzek M., Valley U., Nimtz M., Wagner R., Conradt H.S. (1995) *J Biotechnology*, 42, 117-131.
[82]Watson E., Shah B., Leiderman L., Hsu Y.R., Karkare S., et al. (1994) *Biotechnol Prog*, 10, 39-44.
[83]Skottman H., Strömberg A-M., Matilainen E., Inzunza J., Hovatta O., et al. (2006) *Stem Cells*, 24, 151-167.
[84]Hanna E., Quick J., Libutti S.K. (2009) *Oral Dis*, 15, 8-17.
[85]Albini A., Sporn M.B. (2007) *Nat Rev Cancer*, 7, 139-147.
[86]Srisomsap C., Sawangareetrakul P., Subhasitanont P., Chokchaichamnankit D., Chiablaem K., et al. (2010) *J Biomed Biotechnol*, 2010, 437143.
[87]Cancemi P., Albanese N.N., DiCara G., Marabeti M.R., Costantini F., et al. (2010) *Connect Tissue Res,* 51, 88-104.
[88]Toillon R-A., Lagadec C., Page A., Chopin V., Sautière P-E., et al. (2007) *Mol Cell Proteomics*, 6, 1239-1247.
[89]Paulitschke V., Kunstfeld R., Mohr T., Slany A., Micksche M., et al. (2009) *J Proteome Res*, 8, 2501-2510.
[90]Varnum S.M., Covington C.C., Woodbury R.L., Petritis K., Kangas L.J., et al. (2003) *Breast Cancer Res Treat*, 80, 87-97.
[91]Celis J.E., Moreira J.M., Cabezón T., Gromov P., Friis E., et al. (2005) *Mol Cell Proteomics*, 4, 492-522.
[92]Huang C-M., Ananthaswamy H.N., Barnes S., Ma Y., Kawai M., et al. (2006) *Proteomics*, 6, 6107-6116.
[93]Huang C-M., Nakatsuji T., Liu Y-T., Shi Y. (2008) *Recent Pat AntiCancer Drug Discov*, 3, 48-54.
[94]Wu C-C., Hsu C-W., Chen C-D., Yu C-J., Chang K-P., et al. (2010) *Mol Cell Proteomics*, 9, 1100-1117.
[95]Jackson D.H., Banks R.E. (2010) *Proteomics Clin Appl*, 4, 250 -270.