available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.ejconline.com

Profiling cancer stem cells using protein array technology

*Satoshi Nishizuka**

Molecular Translational Technologies, Molecular Therapeutics Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA
Laboratory of Proteomics and Analytical Technologies, SAIC-Frederick Inc., National Cancer Institute at Frederick, Frederic, MD 21702, USA

ARTICLE INFO

Article history:
Received 23 January 2006
Accepted 23 January 2006
Available online 27 April 2006

Keywords:
Protein array
Cancer stem cell
'Reverse-phase' protein lysate microarray
Proteomic profiling
Proteomics

ABSTRACT

Since cancer cells and somatic stem cells share the biological characteristics of self-renewal and proliferation, it has been suggested that the principles of stem cell biology can be applied to improve our understanding of cancer biology. Recent studies have shown that the majority of cancers appear to originate from a small subset of cells that have the ability of self-renewal and to proliferate, namely 'cancer stem cells'. The isolation of cancer stem cells has been demonstrated using cell surface markers in haematopoietic and non-haematopoietic malignancies. Advances in protein array technologies have enabled the use of minuscule amounts of biological materials to profile these cells at the molecular level. Using a combination of protein arrays and cancer stem cell isolation techniques, a higher resolution molecular profiling can be performed, which might improve therapies targeting the cancer stem cells.

Published by Elsevier Ltd.

1. Introduction

Self-renewal and proliferation are essential properties of both stem cells and cancer cells. Cancer cells appear to originate from a small subset of this population, namely, 'cancer stem cells'.[1–4] The origins of these cells have not yet been fully elucidated, but gaining an understanding of cancer stem cells and using them as a model system may allow cancer biology to be viewed from a different perspective.

The existence of cancer stem cells was first shown in an isolated subset of acute myeloid leukaemia patients, using surface markers.[5] This principle has since been applied to help identify other tissues of origin, such as multiple myeloma, and tumours arising from the breast and brain.[6–8] This has been of particular interest in the field of targeted therapeutics, since evacuation of cancer stem cells using chemotherapy might explain the mechanism underlying drug resistance.[9]

In order to clarify the roles of cancer stem cells in tumour biology, functional molecular profiling is necessary. Transcriptional or genomic microarrays are widely available and have contributed to the detailed molecular profiling of various malignancies. Recent advances in protein array technology have allowed researchers to characterise cancer cells at the proteomic level using only minuscule samples.[10–13] Although the means of isolating a large enough sample of pure cancer stem cells has not yet been established, it may be similar to the way in which subsets of cells are isolated from complex tissue components.[14–19] In the context of proteomic profiling of minuscule samples, cancer stem cells could be a type of cell population that needs to be isolated from the heterogeneous population of the tumour. In contrast to the array format utilised in transcriptional or genomic profiling, the mechanisms, applications and data output of protein arrays varies depending upon the format used.[10–13] Therefore, a more comprehensive study design is required, from posing the biological question to algorithm development for analysis. Among the currently available technologies, 'reverse-phase' protein lysate microarrays will be discussed as one of the potential methods with which it

* Tel.: +1 301 451 0238; fax: +1 301 480 6565.
E-mail address: nishizus@mail.nih.gov.
0959-8049/$ - see front matter. Published by Elsevier Ltd.
doi:10.1016/j.ejca.2006.01.042

may be possible to achieve high-resolution proteomic profiling of cancer stem cells.[20,21]

2. Cancer stem cells in tumours

Accumulation of multiple gene mutations may cause phenotypically diverse populations even in cases in which all cells originate from a single clone. In fact, it is known that most tumour cells do not form colonies when plated in soft agar or when injected into mice.[9,22,23] Those colonies arising from tissues of different origin have different growth characteristics and colony morphology.[4] It has hence been speculated that tumours could arise from a small subset of cells in a tissue-specific manner, and that these cells have the potential to differentiate or proliferate.[24] Such tissue-specific 'tumour initiating cells' have been isolated from a subset of acute myeloid leukaemias by tumour cell transplantation into severe combined immunodeficient (SCID) mice.[5,25] Certain subsets were capable of proliferating distinct from other clonal populations in the tumour. Those cells with proliferative potential were exclusively of the $CD34^{++}$ and $CD38^{-}$ phenotypes, which is similar to normal haematopoietic stem cells, whereas $CD34^{+}CD38^{+}$ leukaemia cells did not exhibit this type of tumour engraftment.[5] It has been speculated that it is the haematopoietic stem cells, rather than committed progenitor cells, that are likely to be the targets of malignant transformation of leukocytes.[3]

Identification of normal stem cells in solid tumours has been particularly difficult because of a lack of efficient molecular markers.[3,24] Tissue-specific stem cells for solid tumours that are composed of epithelial or mesenchymal cells have not been established as being of haematopoietic origin. However, tissue-specific stem cells have been isolated from neural systems[26,27] and it has been speculated that stem cells exist in most organs, with the possible exception of the heart.[28] The multipotent differentiation of adult normal stem cells in development has been demonstrated; they have the potential to differentiate into normal lineages other than the tissue of origin.[29–31] Houghton and colleagues reported that *Helicobacter*-induced chronic inflammation induced the migration of bone marrow-derived cells to the stomach, where they developed gastric cancer through metaplasia and dysplasia.[32] This suggests that normal stem cells have the potential to differentiate not only into normal lineages, but also to develop cancer. However, it is uncertain whether the endogenous supply of adult stem cells from other organs, normally used to ameliorate tissue defects, plays a major function in cancer progression.[29] If the tissue micro-environment causes the 'transdifferentiation' of the bone marrow derived cells, it suggests that the cell environment may determine the fate of multipotent cells, which could induce tumour development.[33] With regard to cancer initiation and progression, it may be reasonable to speculate that tissue-specific stem cells have a greater potential for the initiation of tumourigenesis. The primary target cells for proteomic profiling hence appears to be tissue-specific stem cells unless tissue-specific cancer stem cells can be identified. However, the effective isolation of these cells from different tissues is a challenging task, since the stem cell population is extremely small, usually comprising only 1–2% or less of the total.[28] Further progress in methods of identification of normal tissue-specific and cancer stem cells is required. The implications of success in this matter would be particularly great in anti-cancer drug studies[9]; it may be possible to develop more effective therapeutics if the therapy-resistant mechanisms associated with cancer stem cells are clarified for each type of tumour.

3. Acquiring cancer stem cells for analysis

Methods purifying the most relevant proliferative populations from heterogeneous tumours and of establishing cell culture models with inherent functional characteristics remain to be developed. The method of separating a pure population from a heterogeneous tumour may differ depending upon the tissue of origin.[3,24] Bulk tumours or peripheral blood from a cancer patient are composed of a heterogeneous cell population that includes some normal cells. Moreover, even within a single tumour with clonal expansion, the morphology, cell cycle, molecular signature, cellular response, self-renewal capability and degree of differentiation are not homogeneous.

Molecular surface markers have been used to show that the cells have the potential to proliferate more effectively than normal cells.[5–8,19,25,34,35] In a report by Bonnet and colleagues,[5] several surface markers were used in combination with flow cytometry to distinguish leukaemic tumour stem cells from the rest of the tumour cells, which had limited proliferative potential. The procedure yielded 10,000 cells for each human case, a sample size that permitted downstream assays including xenografts into mice. The isolation of cancer stem cells by surface markers has been applied successfully to solid tumours, including tumours of the breast and brain.[7,8] The phenotype of breast tumour cells, which was defined with the aid of surface markers and $CD44^{+}CD24^{-/low}$ $Lineage^{-}$ was identified as tumourigenic in nude mice, despite the relatively small number of cells injected.[7] Extensive in vitro studies have also been made possible by culturing isolation $CD133^{+}$ cells (neural stem cell marker) from brain tissues.[8]

With the lack of well-developed molecular markers, to isolate stem cells from the intestine has been a challenging task,[3,24] although the histological architecture and the self-renewal process in the colon epithelium has been well studied.[36–38] This knowledge has lead us to approach isolation and characterisation of colon epithelial stem cells according to their morphology. Stem cells are believed to be located in the base of the crypts immediately above the Paneth cell component.[37] Thus, the location of colorectal epithelial stem cells is relatively easily identifiable. However, where and when these stem cells acquire genetic alterations that accelerate proliferation remains unclear. In most of colorectal carcinomas, with the exception of well-differentiated ones, cell crypt organisation or polarity may not be well maintained. It is feasible if tumour stem cells can be identified from relatively well-differentiated tumours, such as adenomas, from which adenocarcinomas frequently arise.[39] The colorectal adenoma-carcinoma sequence has been widely accepted as a model of multi-step tumourigenesis.[40,41] With the aid of microdissected materials from both sporadic and familial adenomatous polyposis (FAP) adenomas, it has been proposed that adenomas arise from transformed stem cells at the crypt

bottom and grow by crypt fission with the 'bottom–up' fashion.[42] On the other hand, Shih and colleagues proposed a 'top–down' morphogenesis in which the dysplastic cells at the surface epithelium migrate and spread downwards to form monocryptal adenoma.[43] It is possible the two pathways of tumourigenesis both exist as there is evidence for both mechanisms of monoclonal adenoma morphogenesis.[38] The involvement of molecular components in Wnt/β-catenin signalling pathways has been implicated in colorectal stem cell tumourigenesis.[36–38] In particular, the adenomatous polyposis coli (APC) gene mutation is known to occur in the vast majority of colorectal adenomas and early stage adenocarcinomas.[39] If stem cell markers and somatic APC alteration can be detected in small subset of tumour cells and not any of the other genetic alterations that are known to be accumulated during tumour progression, then those cells are likely to be involved in the very initial phase of tumourigenesis and possibly represent cancer stem cells. The molecular signature involved in the initiation of colorectal tumourigenesis in the stem cells may be elucidated by studying each component of the pathways corresponding to each tumour stage.

Samples from solid tumours often are most accessible in only archived or frozen block. Even if fresh specimens are available, maintenance of the physiological conditions of the removed samples may not always be performed properly. There is always uncertainty in terms of blood supply or avoiding necrotic centres of the tumour in biopsy samples. The 'freshness' of the samples is often taken as being relative to the time of removal from its position in vivo, not from the timing of various stress events (e.g. hypoxia). The fresh tissues should be processed at ice cold temperature to avoid reactions such as protein degradation or dephosphorylation. A limited number of assays that utilise living cells from solid tumours are available[24]; this suggests the morphological approach should still be performed for primary information.

Although cases are still limited, colony-forming primary tumour graft permits detailed observations[4] since it is considered that only stem cells have the characteristics of self-renewal and proliferation.[7] It has been reported that cancer stem cells from brain tumours can be cultured.[8] However, reprogramming of normal stem cells in vitro raises the question of whether the cultured cells posses the same properties as the original stem cells. Normal stem cells and cancer cells share the entities in self-renewal and proliferation; if isolated stem cells tend to alter their properties due to their high potential for differentiation (or undifferentiation), this possibility should be taken seriously when considering isolation of cancer stem cells.[29,44–47] For the moment, however, culture or colony isolation is the most reasonable way to purify and 'amplify' the cancer stem cells. It remains necessary to obtain as many pure tumour stem cells as possible for use in proteomic profiling.

4. Protein array technologies

The principles of protein array technology were described in late 1980s by Roger Ekins, along with the broader range of the concept of 'ligand assays'.[48,49] The high potential of the miniaturised microarrays had already been shown with evidence of their sensitivity[50–52] and the application of sigmoid curves in ligand assays was also described.[53] There are many variations in protein array formats to date (Table 1). [20,54–68] Here, a protein array is defined as a microarray in which protein interaction is detected with the aid of miniaturised 'ligand' spotting technology[52]; however, it is important to design studies carefully based on the design of the experiments, the kind of output needed, and the kind of samples that are expected. In comparison to transcriptional arrays, protein arrays generally comprise fewer features per slide with different configurations. Furthermore, each format requires a different detection method, surface matrix and data analysis processes for each study design. Transcriptional arrays have introduced many different insights in biology by allowing the measurement of many genes simultaneously.

Table 1 – Miniaturized protein array technologies

Printed on matrix	Detection method	Matrix substrate	Key references (reference number)
Antibody	Antibody (ELISA)	NHS[a]	Mendoza, et al. [54]
Antibody	Analyte labeling	Nitrocellulose	Knezevic, et al. [55]
Antibody	Antibody (ELISA)	Scillicon	Wisse, et al. [56]
Antibody	Antibody (ELISA)	Aminosilanated	Woodbury, et al. [57]
Antibody/antigen	Analyte labeling	Poly-L-lysine	Haab, et al. [58]
Antibody/lysozyme	Direct imaging	Gold	Lee, et al. [59]
Antibody (microcantellever)	Energy transition	Gold	Wu, et al. [60]
Purified protein (human)	Analyte labeling	Nitrocellulose	Ge, et al. [61]
Purified protein	Analyte labeling	Aldehyde	McBeath, et al. [62]
Purified protien (yeast)	Analyte labeling	Aldehyde	Zhu, et al. [63]
Purified protein (yeast)	Antibody	Nitrocellulose	Michaud, et al. [64]
Fractionated protein	Antibody	Various[b]	Madoz-Girpide, et al. [65]
Whole cell lysate (human)	Antibody	Nitrocellulose	Paweletz, et al. [20]
Recombinant protein	Analyte labeling	Aldehyde/Nickel	Zhu, et al. [66]
Peptide	Multiple methods[c]	Gold	Houseman, et al. [67]
Various substrates	Analyte labeling	Poly-L-lysine	Robinson, et al. [68]

a N-hydroxysuccinimide.
b Poly-amine derivatized, poly-aldehyde, and nitrocellulose slides.
c Surface plasmon resonance, fluorescence, and phosphoimaging.

They have also allowed biologists to challenge quantitative interpretations of many biological phenomena. Part of the theoretical approach associated with protein array technologies is common with that of transcriptional arrays[69] but it has larger complexity that makes the protein array technology less accessible. Transcriptional arrays are, in principle, based on hybridisation between specifically interacting probes and the target strands of nucleotides,[18] thus a relatively fair comparison between different probes is possible. Conversely, in protein arrays, the signal associated with the antibody–antigen interaction for instance, is determined by a more complex association between affinity/amount of antibodies and the epitopes on the target protein.[12,18] More importantly, the numerical output is quantitative, but in a relative and not an absolute sense. Therefore, a more sophisticated analytical process is required.

Before applying these protein array technologies for cancer stem cells, we should be aware of their limitations in terms of the purity and quantity of the samples required. One possible application is the detection of protein molecules in cancer stem cells using several probes that are spotted onto the solid surface. Protein arrays that immobilise antibodies (antibody arrays) may be useful when there are many molecules within a small set of samples. Knezevic and colleagues[55] have reported that they were able to profile squamous cell carcinomas in oral cavity from different histological compartments within a tumour using antibody arrays and with only a small volume of samples. Although further technical improvements are still required (only a small proportion of commercially available antibodies provide measurements that are sufficiently specific and accurate),[13,58] much progress has been made with regard to antibody immobilisation,[70,71] antibody specificity,[64] and label-free detection system.[60] Separation by cluster designation (CD) antigens can be used to isolate cancer stem cells.[5,7,8,25] More detailed CD profiling may be possible with antibody arrays using protein samples from different colonies after inoculating in soft agar. It has been demonstrated that for yeast, global proteomic profiling can be used to reveal protein–protein or protein–small molecule interactions.[63] Single slides containing a few thousand of purified human proteins are now commercially available[72]; this may enable the identification of the unique molecular interactions taking place in cancer stem cells.

With regard to preserving samples, the 'reverse-phase' protein lysate microarrays (RPAs) provide an alternative, having 'reverse' format arrays on which complex mixtures of proteins are spotted and detect one specific protein at a time.[20] The main advantage of the RPA is that all signals are developed under the same condition for all samples or probes on the array. RPAs have proved to be powerful tools for characterising many different types of cells at the proteomic level. For instance, the NCI-60 (a cancer cell line panel that has been used for anti-cancer drug screening) reverse-phase arrays[21] were designed to obtain quantitative output in order to relate with other quantitative databases. The numerical output of protein expression from the RPAs allows comparing quantitative mRNA data from transcriptional arrays and matched protein species. Subsequent functional classification has revealed that structure associated protein expression levels correlate well with the corresponding mRNA expression levels.[21] RPAs have also been used to identify molecular markers that are useful in the field of pathology, allowing the bridging of transcriptional and proteomic data.[73] They can also be useful for profiling cancer stem cells when there exist many comparable sets of populations, for example, stem cells from different tissues, or from different colonies based on the growth or morphology. Although there are many different formats of protein spotting technologies, RPAs will be described in more detail here as one of the most established experimental and analytical systems that may be applicable to cancer stem cell proteomic profiling.

5. Reverse-phase protein lysate arrays

5.1. *Background*

RPAs were first introduced by Paweletz and colleagues[20] to measure subtle quantitative changes in multiple classes of proteins in an individual cell type in tissues. In contrast to antibody or ligand arrays, the RPA immobilises whole cell protein lysates on nitrocellulose coated glass slides (Fig. 1). The mechanism of signal detection is in principle a tyramide-based amplified immunochemical method, like those used routinely in immunohistochemistry (Fig. 2(a)).[74] Protein expression levels are appropriately quantified by assessing the linear range in a serial dilution of the whole cell lysate (Fig. 2(b)).[21] Slides are probed with one antibody per slide, but hundreds of identical slides can be produced from one microtitre plate. Therefore, the RPA is a powerful method to use when: (i) many samples, (ii) many antibodies, and (iii) quantitative output, are available and desired. The experimental and analytical flow is shown in Fig. 3.

It is important to adjust lysis conditions and have prior knowledge of antibody specificity for quantitative analysis. Samples can be many clinical specimens, cell lines, or different pathological conditions of these materials. Before use, an antibody must be screened to establish its specificity for quantitation of the arrays. Most importantly, the RPA should be considered not as a substitute of Western or immunohistochemistry, rather as an additional technique in proteomics, providing quantitative data in a higher throughput.

5.2. *Sample complexity*

Whole protein lysate from a human cell is a complex mixture of the products of approximately 60,000 genes, resulting in an estimation of 10^9 protein molecules per cell.[75] The dynamic range of protein abundance in a biological sample has been estimated as high as 1,000,000.[76] Each protein is a unique entity in terms of three-dimensional conformations, charge, or number of residues, and each of them can be modified, forming a complex or changing the localisation, depending on the function required. Considering the dynamic range of protein expression levels among all species of protein within a cell, and between the cells (or conditions), it is perhaps not surprising that not all of the proteins of interest will be in the analysable range. A routine two-dimensional (2D) gel separation yields no more than 1500 proteins[76] because the resolution is limited to the more

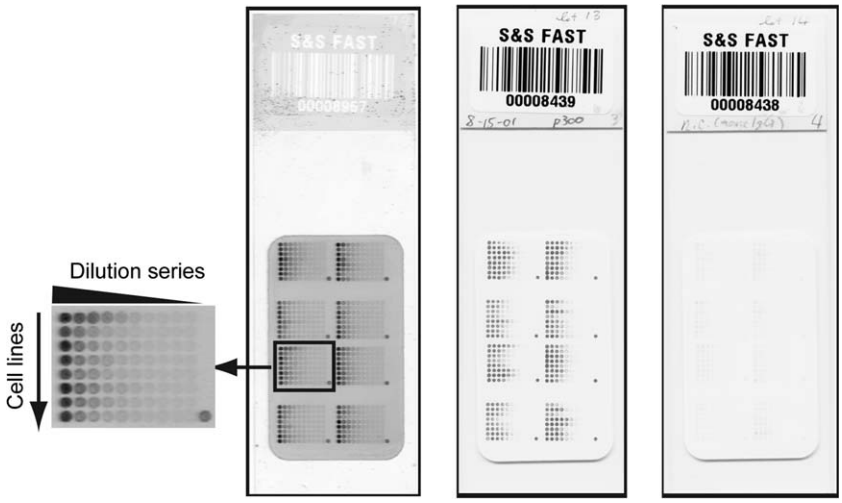

Fig. 1 – 'Reverse-phase' protein lysate microarrays. All NCI-60 cancer cell lines are spotted on a single slide. Left: staining with SYPRO ruby for total protein. Each row (see enlarged image at the left) consists of 10 twofold dilutions of an NCI-60 cell line or the control pool. A concentrated control pool was spotted at the bottom right corner of each field as a registration mark for scanning. Middle: catalysed signal amplification (CSA) staining (details in Fig. 2) for p300 protein. Right: negative control.

abundant proteins within the whole cell lysate.[75,77] Therefore, it may not be feasible to employ global proteomic monitoring to cover many proteins from a wide dynamic range with a single detection method. One of the greatest advantages of the RPAs is to permit using whole cell lysate enabling lower abundance proteins of interest using amplification technique.[74,78–80] Non-abundant proteins often play biologically important roles, and the loss of these potentially important proteins may be prevented by this ability to carry out whole cell lysate spotting without further fractionation of the sample. To ensure that as many proteins as possible are accessible to the antibodies, we have developed a procedure that keeps proteins fully denatured using a modified version of Pink Buffer[21] which was initially developed for 2D gel application by Anderson and colleagues.[81] This buffer system also allows the multiple processing of samples even after repeating thaw–freeze cycles.

The application of whole cell lysate spotting has protein-specific problems that are not often seen with nucleotide applications. Antibody-based detection systems largely depend on their affinities and the abundance of each protein species.[18,50] In order to achieve a large dynamic range of detection with the complex mixture, it is important to start the series of dilutions with a high concentration. In our previous studies, the estimated starting protein concentration on the array was 25–50 μg in estimate per well in 20 μl of buffer volume.[21] We have used a pin-and-ring microarrayer (GMS 417, Genetic Microsystems/Affymetrix, Boston, United States of America (USA)) to handle highly viscous samples in the modified Pink Buffer. Although whole cell lysate spotting has not yet been a major focus of the microarray industry, advances in microarray fabrication technologies now provide arrayers that are more capable of handling highly viscous samples and that can provide extremely high throughput. In

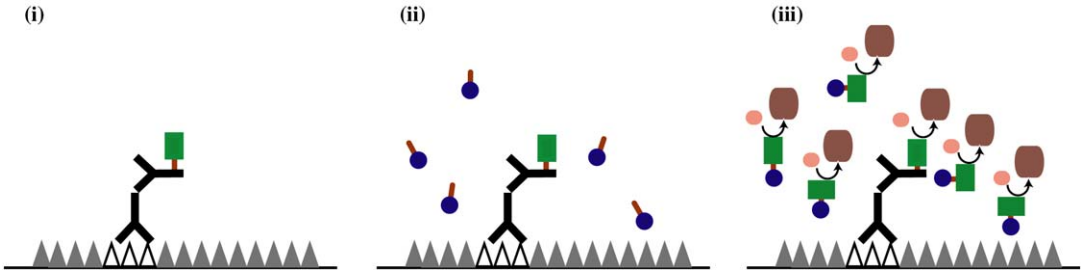

Fig. 2a – Catalysed signal amplification (CSA): (i) The specific primary antibody binds to the specific antigen (white triangle) represented in a complex mixture followed by secondary antibody incubation, which conjugated with biotin (brown bar) and horseradish peroxidase (HRP) (green square). (ii) Application of biotin conjugated tyramide (blue circle) amplification reagent. Tyramide is attracted by HRP-conjugated secondary antibody. Hence, there are many free biotins around the antigen. (iii) Application of additional HRP. The HRP molecules are conjugated with streptavidin (not shown) so bind tightly free biotins attached to the tyramide. The total amount of HRP is now amplified so that a more catalytic reaction by HRP takes place with the substrate, diaminobenzidine (DAB; pink dot). DAB molecules are oxidised by HRP at the presence of oxygen donor, hydrogen peroxidase (not shown) resulting in brownish precipitate (brown square).

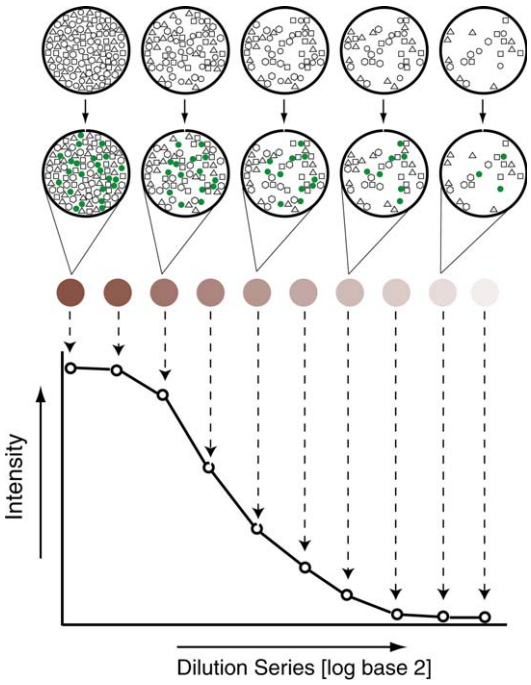

Fig. 2b – A conceptual view of 'Reverse-phase' protein lysate microarrays (RPAs). Each circle in the top row represents a serially diluted complex mixture of protein lysate. The direction of the dilution is shown from left (high concentration) to right (low concentration). The second row of five circles indicates that the immunostaining reaction has been performed, in each of which enzymatic reaction shown in Fig. 2 is ready to take place. The green dots represent the presence of horseradish peroxidase (HRP) as a result of specific primary antibody binding to the antigen. The enzymatic reaction yields a brownish precipitation on each of the ten dilution series. The intensity of the precipitation is not necessarily linearly associated with the dilution (a plateauxing of the intensity is often seen). After the enzymatic reaction is complete, the intensity of each spot is scanned and the resulting image is converted to pixel values. An intensity-dilution 'dose–response' curve is demonstrated. The protein expression level for the proteins tested is calculated from the linear range.

collaboration with Aushon BioSystems (Lexington, MA, USA),[82] we are in the process of developing a large scale RPA system using a solid-pin arrayer designed for spotting viscous samples. The arrayer is capable of depositing protein samples with high consistency and at high speed in the face of inherent problems such as evaporation. The arrayer can handle up to 100 slides and 30 or more microtitre plates at a single run under highly controlled environmental conditions.[82]

5.3. *Specificity determination*

The samples spotted on the array contain whole protein fractions that can be separated with higher dimensional methods. The signals obtained from RPAs are considered to be the sum of specific and non-specific bands generated by an antibody.[21] To achieve high-quality quantitation, the antibodies used for probing need to be screened using the Western blot prior to the array experiment with the same samples to ensure that they produce a single predominant band. In a previous study, we screened more than 200 antibodies against a pool of sample cell lines prior to application of RPAs.[21] Although only about 70% of mouse monoclonal antibodies showed a single predominant band and about 70% of the screened antibodies produced an analysable signal on arrays[83], this ensured that the quality of the signal (i.e. its specificity) was more reliable. In cases, where other methods of measuring protein expression are available (e.g. immunohistochemistry), the use of both RPA and immunohistochemisty may provide information about any associations between the semiquantitative and quantitative results, or between cellular localisation and the quantitative results, without the necessity for prior antibody screening. It should be noted, however, that the way target proteins form in the tissue/cell and in nitrocellulose membranes may be considerably different, and this may or may not affect the signal even if the same antibody is used.

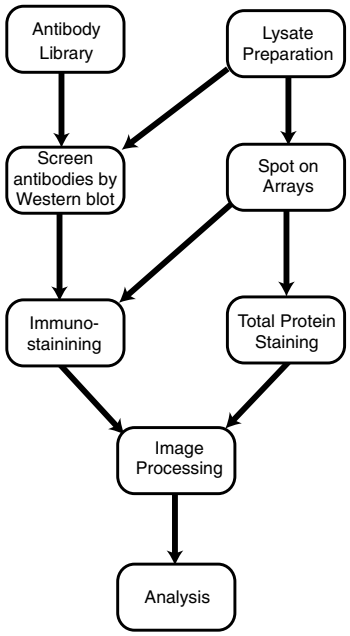

Fig. 3 – Experimental and analytical flow of 'reverse-phase' protein lysate microarray (RPA) system.

5.4. *Surface matrix*

To spot complex mixture, the surface matrix has to satisfy various requirements. In general, the greater the surface area the greater the number of proteins that can be captured; this is particularly important for those proteins whose abundance is unknown. Nitrocellulose membranes have a great capacity for protein capture when the proteins are presented in high concentration in liquid, and retain substantial fraction after stringent washing procedures. The actual mechanical forces involved in protein molecule capturing process remain unknown, although there are several hypotheses, such as hydrophobic interactions, hydrogen bonding, and electrostatic interactions.[84] For RPAs, we have been using nitrocellulose-

coated glass slide (Grace BioLabs, Bend, Oregon, USA).[85,86] These nitrocellulose membranes are particularly suitable when the protein lysing or buffer solutions contain a high concentration of chemical components. The main disadvantages of the nitrocellulose membranes are that it produces a high background and production lot differences.[87] With regard to the former issue, it is important to choose antibodies that are known to produce high signals or to eliminate from analyses slides that have a low signal-to-noise ratio. Post-processing of the slide image will also be considered. To minimise the influence of the lot differences, it is useful to use control slides with the same sample and antibodies for every experiment so that consistency between lots can be estimated. The best scenario would be to generate all of the data required in a single run, if practical.

5.5. *Signal detection*

Even if antibodies are screened prior to their use in arrays, the intensity of the signals achieved will not be known until it is developed on arrays. This is because these signals are the product of complex chemical reactions that are affected by various factors, including the antibody, sample and the affinity of the antibody.[18,50] Therefore, it would be preferable to use detection methods with wider dynamic range between samples with a protein and within a sample with different proteins.

The tyramide-based technique has been reported to produce the same signal intensity for 2.5–500-fold diluted primary antibodies compared with conventional immunohistochemical techniques.[74] Hence, applications that require many different antibodies should take advantage of this technique since it is likely that more target antigens would lie in the detectable range. For nitrocellulose membranes, we generally use colourimetric detection with a tyramide-based amplification system.[20,21] The resulting signals can be treated simply as black and white intensity like ink spotted onto paper. With the combination of tyramide signal amplification, or using a non-autofluorescent spectrum, such as infrared, several of successful fluorescent applications have been reported with RPA format.[86,88,90] Both colourimetric and fluorescent signal detection have advantages and disadvantages with respect to background issues. It is important to choose a signal detection system that generates high signal-to-noise ratio for each study.

'High resolution' proteomic profiling using RPAs requires many quality antibodies, which often introduces the problem of cost. This problem might be alleviated if institutions set up a central 'antibody bank', whereby many different antibodies could be collected from different laboratories within a particular institution. In addition, the use of a good standard sample, not a positive control for a particular antibody, would allow the testing of many antibodies in a higher throughput manner, providing information on whether particular antibodies of interest are likely to work. Some antibody companies offer a small-volume sampler kit. The volume of the antibody provided in the kit is sufficient for multiple Western blotting, which would be more than sufficient for use in many protein array experiments as well as antibody screening on Western blots (one slide requires 0.5–0.8 ml of total antibody incubation solution). Antibody screening is important for quantitative studies because it not only provides information about the antibody's specificity, but also allows us to determine whether the sample–antibody combination is appropriate for quantitative analysis.

5.6. *Image analysis*

Like other microarray techniques, the acquired images of RPAs need to be digitised for quantitation. An ordinary flatbed scanner can be used for image processing for colourimetric data, removing the need for special equipment such as laser scanners. We have used an ordinary 'stationary-grade' optical flatbed scanner (Epson Perfection 1200) for the slide scan with a resolution of 600 dpi, and 8-bit gray-scale settings.[21] The 500 μm solid pins of the GMS arrayer produce spots of approximately 600 μm in diameter, so the 600 dpi resolution is sufficient (about 200 pixels per spot). The size of the resulting TIFF file is only 1 MB so it can be processed without much frustration. We were able to achieve a wide dynamic range in previous studies.[21,73] More recent flatbed scanners have reached at the non-interpolate resolutions of 4800 dpi, which is equivalent to 0.5 μm pixel scale scanning (Perfection 4870). It is a high-end photo-scanning device, but the cost is about 100 times less than a laser microarray scanner. However, such scanners are not designed for image quantitation, and the built-in scanning driver often skews the signal intensity of the image. The association between the optical density and digital readout therefore often needs to be calibrated. This may be resolved using calibration density strips, such as of the type we have developed.[89] Obtaining accurate quantitative data requires an understanding of the characteristics of each scanner and the image processing programs. The resolution of the scanned images is determined by the spot size, and background subtraction has to be determined by extensive optimisation studies.

5.7. *Interpretation on dose–response curves*

The final output is provided as numerical values converted from digital images. Each step possesses the potential for error, which may be the cause of lack of perfect reproducibility. Possible error sources include, among others, slide surface inconsistency, staining variation, sampling variations. The size of each error source will also differ between experiments. Quantitative signal development in which antigen, antibody and enzymes react is the consequence of a complex process. The rationale behind generating numbers that may represent protein expression levels is to find out whether there is an observable robust association among proteins, or cells. We previously introduced a 25% dose interpolation (DI_{25}) algorithm to measure protein expression levels.[21,73] Briefly, using numerical values from *P*-scan[91,92] readout from each spot on the array, the algorithm employs monotonic linear spline fit to the serial dilution curve. An extensive optimisation analysis revealed that approximately 25% of the intensity range of a given slide tends to yield the expression value (DI_{25}) that lies within a linear range. In principle, the algorithm generates one single number per cell per protein so that various statistical methods can be applied thereafter. The pattern of two-dimensional hierarchical clustering using NCI-60 cancer cell

line panel revealed that a biological replicates and technical replicates come together in the same cluster, suggesting that the system, including the DI_{25} algorithm is technically reliable and biologically reasonable.[21] However, we do not yet fully understand the level of noise and signals that the system generates because the numerical data obtained from any technology may arise from a mixture of biological significance and technical noise. Nonetheless, quantitative data sets generated by the DI_{25} algorithms have been used for various statistical applications, as described previously.[21,73] It is particularly important to design study flow and the array format only after analytical programs are constructed. Each of the steps (Fig. 3) are tightly associated each other, so the system should be considered as a continuous process that is designed to answer particular biological questions.

6. Application of RPA for cancer stem cell biology

Recent progress in tissue-specific stem cell discovery has shown a broader application of the concept of stem cells in haematopoietic systems.[29] These stem cells seem to be the target of cancer initiation and are considered to be the only subset of cells that have a significant potential for self-renewal and proliferation. It is particularly important to study the fraction of this subset if cancer is interpreted as a stem cell disorder.[3] However, the cancer stem cell fraction appears to be extremely small, which severely limits the number of methods that can be applied for molecular profiling.[5,25,28] Identification of the cancer stem cells by morphological approach has also yet to be established. Perhaps the first approach should be characterising tissue-specific normal stem cells and comparing them with cancer cells that have the ability to proliferate. RPA technology would be useful here, by allowing the accommodation of many samples on one array, which may have been collected from different parts of different tumours or individuals. Although the minimum number of cells necessary for an assay is determined by the method of collection, the high throughput array format permits the use of many antibodies over many slides produced at the same time. The antibodies should be chosen based on relevance in the first round, their use being expanded for detailed profiling, using many CD markers,[93] for instance. Although major efforts are still required to establish efficient methods of collection of pure cancer stem cells and to develop corresponding analytical algorithm, the molecular profiling of cancer stem cells at the proteomic level remains critical to our understanding of the mechanisms underlying the development and progress of cancers, and to the development of therapeutic target identification.

7. Conclusion

To characterise a small subset of cells is still a challenge. However, it would be of incalculable benefit in providing new insights in cancer biology and therapeutics. Integration of the concept and molecular profile of cancer stem cells, which is becoming ever more detailed, will be of particular importance.

Conflict of interest statement

None declared.

Acknowledgements

The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organisations imply endorsement by the US Government. This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract Number NO1-CO-12400.

The author appreciates the contribution of the following people: Daisaku Morita, Sylvia Major, Frank Washburn, Dan Asin, Ellen Stevens, Lynn Young, Lu Charboneau, Ginny Espina, Emanuel Petricoin, Lance Liotta, and John Weinstein, for experimental and analytical support; Peter Munson, Michael Dean, and Martina Rudelius for critical comments on the manuscript. The author also appreciates Jennipher Grudzien of Grace BioLabs for providing high quality nitrocellulose membranes, and Pete Honkanen and John Austin of Aushon BioSystems for excellent advice on microarray engineering.

REFERENCES

1. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature* 2001;**414**(6859):105–11.
2. Clarke MF. Neurobiology: at the root of brain cancer. *Nature* 2004;**432**(7015):281–2.
3. Huntly BJ, Gilliland DG. Leukaemia stem cells and the evolution of cancer-stem-cell research. *Nat Rev Cancer* 2005;**5**(4):311–21.
4. Hamburger AW, Salmon SE. Primary bioassay of human tumor stem cells. *Science* 1977;**197**(4302):461–3.
5. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 1997;**3**(7):730–7.
6. Matsui W, Huff CA, Wang Q, et al. Characterization of clonogenic multiple myeloma cells. *Blood* 2004;**103**(6):2332–6.
7. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci USA* 2003;**100**(7):3983–8.
8. Singh SK, Clarke ID, Terasaki M, et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res* 2003;**63**(18):5821–8.
9. Dean M, Fojo T, Bates S. Tumour stem cells and drug resistance. *Nat Rev Cancer* 2005;**5**(4):275–84.
10. Sakanyan V. High-throughput and multiplexed protein array technology: protein–DNA and protein–protein interactions. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005;**815**(1–2):77–95.
11. Templin MF, Stoll D, Schrenk M, Traub PC, Vohringer CF, Joos TO. Protein microarray technology. *Drug Discov Today* 2002;**7**(15):815–22.
12. Liotta LA, Espina V, Mehta AI, et al. Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* 2003;**3**(4):317–25.
13. Utz PJ. Protein arrays for studying blood cells and their secreted products. *Immunol Rev* 2005;**204**:264–82.

14. Ahram M, Flaig MJ, Gillespie JW, et al. Evaluation of ethanol-fixed, paraffin-embedded tissues for proteomic applications. *Proteomics* 2003;**3**(4):413–21.
15. Paweletz CP, Liotta LA, Petricoin 3rd EF. New technologies for biomarker analysis of prostate cancer progression: Laser capture microdissection and tissue proteomics. *Urology* 2001;**57**(4 suppl 1):160–3.
16. Zangar RC, Varnum SM, Covington CY, Smith RD. A rational approach for discovering and validating cancer markers in very small samples using mass spectrometry and ELISA microarrays. *Dis Markers* 2004;**20**(3):135–48.
17. Petricoin EF, Zoon KC, Kohn EC, Barrett JC, Liotta LA. Clinical proteomics: translating benchside promise into bedside reality. *Nat Rev Drug Discov* 2002;**1**(9):683–95.
18. Liotta L, Petricoin E. Molecular profiling of human cancer. *Nat Rev Genet* 2000;**1**(1):48–56.
19. Handgretinger R, Gordon PR, Leimig T, et al. Biology and plasticity of CD133+ hematopoietic stem cells. *Ann NY Acad Sci* 2003;**996**:141–51.
20. Paweletz CP, Charboneau L, Bichsel VE, et al. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* 2001;**20**(16):1981–9.
21. Nishizuka S, Charboneau L, Young L, et al. Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *Proc Natl Acad Sci USA* 2003;**100**(24):14229–34.
22. Gioanni J, Farges MF, Duplay H, et al. In vitro clonogenicity in relation to kinetic and clinicopathological features of breast cancer. *Bull Cancer* 1988;**75**(3):285–90.
23. Price JE, Tarin D. Low incidence of tumourigenicity in agarose colonies from spontaneous murine mammary tumours. *Differentiation* 1989;**41**(3):202–7.
24. Pardal R, Clarke MF, Morrison SJ. Applying the principles of stem-cell biology to cancer. *Nat Rev Cancer* 2003;**3**(12):895–902.
25. Lapidot T, Sirard C, Vormoor J, et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 1994;**367**(6464):645–8.
26. Morrison SJ, White PM, Zock C, Anderson DJ. Prospective identification, isolation by flow cytometry, and in vivo self-renewal of multipotent mammalian neural crest stem cells. *Cell* 1999;**96**(5):737–49.
27. Uchida N, Buck DW, He D, et al. Direct isolation of human central nervous system stem cells. *Proc Natl Acad Sci USA* 2000;**97**(26):14720–5.
28. Preston SL, Alison MR, Forbes SJ, Direkze NC, Poulsom R, Wright NA. The new stem cell biology: something for everyone. *Mol Pathol* 2003;**56**(2):86–96.
29. Anderson DJ, Gage FH, Weissman IL. Can stem cells cross lineage boundaries? *Nat Med* 2001;**7**(4):393–5.
30. Jiang Y, Jahagirdar BN, Reinhardt RL, et al. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* 2002;**418**(6893):41–9.
31. Krause DS, Theise ND, Collector MI, et al. Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell* 2001;**105**(3):369–77.
32. Houghton J, Stoicov C, Nomura S, et al. Gastric cancer originating from bone marrow-derived cells. *Science* 2004;**306**(5701):1568–71.
33. Bissell MJ, Labarge MA. Context, tissue plasticity, and cancer: are tumor stem cells also regulated by the microenvironment? *Cancer Cell* 2005;**7**(1):17–23.
34. Singh SK, Hawkins C, Clarke ID, et al. Identification of human brain tumour initiating cells. *Nature* 2004;**432**(7015):396–401.
35. Potten CS, Booth C, Tudor GL, et al. Identification of a putative intestinal stem cell and early lineage marker; musashi-1. *Differentiation* 2003;**71**(1):28–41.
36. Radtke F, Clevers H. Self-renewal and cancer of the gut: two sides of a coin. *Science* 2005;**307**(5717):1904–9.
37. Pinto D, Clevers H. Wnt, stem cells and cancer in the intestine. *Biol Cell* 2005;**97**(3):185–96.
38. Brittan M, Wright NA. Stem cell in gastrointestinal structure and neoplastic development. *Gut* 2004;**53**(6):899–910.
39. Powell SM, Zilz N, Beazer-Barclay Y, et al. APC mutations occur early during colorectal tumorigenesis. *Nature* 1992;**359**(6392):235–7.
40. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *New Engl J Med* 1988;**319**(9):525–32.
41. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;**61**(5):759–67.
42. Preston SL, Wong WM, Chan AO, et al. Bottom–up histogenesis of colorectal adenomas: origin in the monocryptal adenoma and initial expansion by crypt fission. *Cancer Res* 2003;**63**(13):3819–25.
43. Shih IM, Wang TL, Traverso G, et al. Top–down morphogenesis of colorectal tumors. *Proc Natl Acad Sci USA* 2001;**98**(5):2640–5.
44. Kondo T, Raff M. Oligodendrocyte precursor cells reprogrammed to become multipotential CNS stem cells. *Science* 2000;**289**(5485):1754–7.
45. Palmer TD, Markakis EA, Willhoite AR, Safar F, Gage FH. Fibroblast growth factor-2 activates a latent neurogenic program in neural stem cells from diverse regions of the adult CNS. *J Neurosci* 1999;**19**(19):8487–97.
46. Donovan PJ. Growth factor regulation of mouse primordial germ cell development. *Curr Top Dev Biol* 1994;**29**:189–225.
47. Matsui Y, Zsebo K, Hogan BL. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* 1992;**70**(5):841–7.
48. Ekins R. A shadow over immunoassay. *Nature* 1989;**340**(6231):256–8.
49. Ekins RP. Multi-analyte immunoassay. *J Pharm Biomed Anal* 1989;**7**(2):155–68.
50. Ekins RP, Chu FW. Multianalyte microspot immunoassay – microanalytical 'compact disk' of the future. *Clin Chem* 1991;**37**(11):1955–67.
51. Ekins RP, Chu F. Developing multianalyte assays. *Trends Biotechnol* 1994;**12**(3):89–94.
52. Ekins RP. Ligand assays: from electrophoresis to miniaturized microarrays. *Clin Chem* 1998;**44**(9):2015–30.
53. DeLean A, Munson PJ, Rodbard D. Simultaneous analysis of families of sigmoidal curves: application to bioassay, radioligand assay, and physiological dose–response curves. *Am J Physiol* 1978;**235**(2):E97–E102.
54. Mendoza LG, McQuary P, Mongan A, Gangadharan R, Brignac M, Eggers M. High-throughput microarray-based enzyme-linked immunosorbent assay (ELISA). *Biotechniques* 1999;**27**(4):778–80.. 782–6.
55. Knezevic V, Leethanakul C, Bichsel VE, et al. Proteomic profiling of the cancer microenvironment by antibody arrays. *Proteomics* 2001;**1**(10):1271–8.
56. Wiese R, Belosludtsev Y, Powdrill T, Thompson P, Hogan M. Simultaneous multianalyte ELISA performed on a microarray platform. *Clin Chem* 2001;**47**(8):1451–7.
57. Woodbury RL, Varnum SM, Zangar RC. Elevated HGF levels in sera from breast cancer patients detected using a protein microarray ELISA. *J Proteome Res* 2002;**1**(3):233–7.
58. Haab BB, Dunham MJ, Brown PO. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol* 2001;**2**(2):RESEARCH0004.
59. Lee KB, Park SJ, Mirkin CA, Smith JC, Mrksich M. Protein nanoarrays generated by dip-pen nanolithography. *Science* 2002;**295**(5560):1702–5.

60. Wu G, Datar RH, Hansen KM, Thundat T, Cote RJ, Majumdar A. Bioassay of prostate-specific antigen (PSA) using microcantilevers. *Nat Biotechnol* 2001;**19**(9):856–60.
61. Ge H. UPA, a universal protein array system for quantitative detection of protein–protein, protein–DNA, protein–RNA and protein–ligand interactions. *Nucleic Acids Res* 2000;**28**(2):e3.
62. MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science* 2000;**289**(5485):1760–3.
63. Zhu H, Bilgin M, Bangham R, et al. Global analysis of protein activities using proteome chips. *Science* 2001;**293**(5537):2101–5.
64. Michaud GA, Salcius M, Zhou F, et al. Analyzing antibody specificity with whole proteome microarrays. *Nat Biotechnol* 2003;**21**(12):1509–12.
65. Madoz-Gurpide J, Wang H, Misek DE, Brichory F, Hanash SM. Protein based microarrays: a tool for probing the proteome of cancer cells and tissues. *Proteomics* 2001;**1**(10):1279–87.
66. Zhu H, Klemic JF, Chang S, et al. Analysis of yeast protein kinases using protein chips. *Nat Genet* 2000;**26**(3):283–9.
67. Houseman BT, Huh JH, Kron SJ, Mrksich M. Peptide chips for the quantitative evaluation of protein kinase activity. *Nat Biotechnol* 2002;**20**(3):270–4.
68. Robinson WH, DiGennaro C, Hueber W, et al. Autoantigen microarrays for multiplex characterization of autoantibody responses. *Nat Med* 2002;**8**(3):295–301.
69. Eckel-Passow JE, Hoering A, Therneau TM, Ghobrial I. Experimental design and analysis of antibody microarrays: applying methods from cDNA arrays. *Cancer Res* 2005;**65**(8):2985–9.
70. Kwon Y, Han Z, Karatan E, Mrksich M, Kay BK. Antibody arrays prepared by cutinase-mediated immobilization on self-assembled monolayers. *Anal Chem* 2004;**76**(19):5713–20.
71. Lee Y, Lee EK, Cho YW, et al. ProteoChip: a highly sensitive protein microarray prepared by a novel method of protein immobilization for application of protein–protein interaction studies. *Proteomics* 2003;**3**(12):2289–304.
72. Available from: <http://www.invitrogen.com>.
73. Nishizuka S, Chen ST, Gwadry FG, et al. Diagnostic markers that distinguish colon and ovarian adenocarcinomas: identification by genomic, proteomic, and tissue array profiling. *Cancer Res* 2003;**63**(17):5243–50.
74. von Wasielewski R, Mengel M, Gignac S, Wilkens L, Werner M, Georgii A. Tyramine amplification technique in routine immunohistochemistry. *J Histochem Cytochem* 1997;**45**(11):1455–9.
75. Vuong GL, Weiss SM, Kammer W, et al. Improved sensitivity proteomics by postharvest alkylation and radioactive labelling of proteins. *Electrophoresis* 2000;**21**(13):2594–605.
76. Simpson R. *Proteins and proteomics*. Woodbury, NY, USA: Cold Spring Harbor Laboratory Press; 2003.
77. Richards P, Lees J. Functional proteomics using microchannel plate detectors. *Proteomics* 2002;**2**(3):256–61.
78. Bacarese-Hamilton T, Mezzasoma L, et al. Detection of allergen-specific IgE on microarrays by use of signal amplification techniques. *Clin Chem* 2002;**48**(8):1367–70.
79. Bobrow MN, Harris TD, Shaughnessy KJ, Litt GJ. Catalyzed reporter deposition, a novel method of signal amplification. Application to immunoassays. *J Immunol Methods* 1989; **125**(1–2):279–85.
80. Bobrow MN, Shaughnessy KJ, Litt GJ. Catalyzed reporter deposition, a novel method of signal amplification. II. Application to membrane immunoassays. *J Immunol Methods* 1991;**137**(1):103–12.
81. Anderson NL, Esquer-Blasco R, Hofmann JP, Anderson NG. A two-dimensional gel database of rat liver proteins useful in gene regulation and drug effects studies. *Electrophoresis* 1991;**12**(11):907–30.
82. Available from: http://aushon.com/.
83. Major SM, Nishizuka S, Morita D, et al. AbMiner: A bioinformatic resource on available monoclonal antibodies and corresponding gene identifiers for genomic, proteomic, and immunologic studies. *BMC Bioinformatics* 2006;**7**:192.
84. Jones KD. Troubleshooting protein binding in nitrocellulose membranes. *IVD Technology Magazine*:26.
85. Available from: <http://www.gracebio.com/>.
86. Stillman BA, Tonkinson JL. FAST slides: a novel surface for microarrays. *Biotechniques* 2000;**29**(3):630–5.
87. Bacarese-Hamilton T, Gray J, Crisanti A. Proteome-scale analysis of the immune response against pathogenic microorganisms. In: Schena M, editor. *Protein microarrays*. Sudbury, MA, USA: Jones and Bartlet; 2004. p. 387–400.
88. Calvert VS, Tang Y, Boveia V, et al. Development of multiplexed protein profiling and detection using near infrared detection of reverse-phase protein microarrays. *Clin Proteom J* 2004;**1**:81–9.
89. Nishizuka S, Washburn NR, Munson PJ. Evaluation method of ordinary flatbed scanners for quantitative density analysis. *BioTechniques* 2006;**40**:442–8.
90. Chan SM, Ermann J, Su L, Fathman CG, Utz PJ. Protein microarrays for multiplex analysis of signal transduction pathways. *Nat Med* 2004;**10**(12):1390–6.
91. Available from: <http://abs.cit.nih.gov/pscan/>.
92. Carlisle AJ, Prabhu VV, Elkahloun A, et al. Development of a prostate cDNA microarray and statistical gene expression analysis package. *Mol Carcinog* 2000;**28**(1):12–22.
93. Liu AY, Roudier MP, True LD. Heterogeneity in primary and metastatic prostate cancer as defined by cell surface CD profile. *Am J Pathol* 2004;**165**(5):1543–56.