Part 1: Biological and Scientific Aspects

1: An Update on Biotechnology in the Assessment of Prostate Cancer

Tim Lane, Jon Strefford and Tim Oliver

INTRODUCTION

Prostate cancer represents the second most common cancer in men (after lung cancer) in both the United Kingdom and the European Union and is the most common cancer in men in the United States. With over 20 000 new cases diagnosed in the United Kingdom every year (and 134 000 in Europe as a whole), the disease has a significant impact on the health of the nation. It is a large and growing problem, with the annual cost of treating prostate cancer in England and Wales alone likely to exceed €100 million [1].

Prostate cancers however differ widely in their aggressiveness and as a result their relative prognoses vary accordingly. Efforts aimed at understanding the underlying pathogenesis of the disease process serve to highlight both new and improved biological markers of disease initiation and progression, which may be of diagnostic or prognostic significance. In recent years, there has been a proliferation of innovative techniques that have served to probe malignant prostate growth at both the genomic and proteomic levels. Such advances, the majority of which have been pioneered by workers in the biotechnology industry, have resulted in both highly unique and specific tools for investigating the changes paralleling tumourigenesis. The techniques involved have inevitably become automated and advanced by improvements in microtechnology, which has further facilitated the rapid throughput of high-quality research data.

In this opening chapter, we seek to describe a variety of innovative tools alongside the improved understanding of prostate tumour biology, which have been accrued as a result.

PROSTATE CANCER GENOMICS

In the vast majority of prostate cancers, there are no inherited defects affecting high-penetrance susceptibility genes and malignant transformation results instead from a series of acquired somatic changes affecting many

genes on several chromosomes. Understanding the genetic changes involved in the development of prostate cancer is pivotal to a rational approach to both diagnosis and therapeutic intervention. It has been estimated that anywhere between 5 and 10 genes are deleted [putative tumour-suppressor genes (TSG)] before malignant transformation occurs and that a further series of gains (amplification of oncogenes) arise with the advent of metastases [2]. A number of techniques have been used to investigate these changes. While classical cytogenetics and loss of heterozygosity (LOH) studies have been the mainstay for investigators for the last few decades, these techniques have been increasingly superseded by a number of alternative strategies.

Fluorescent *In Situ* Hybridisation

Fluorescent *in situ* hybridisation (FISH) represents an established and widely used tool in the investigation of both DNA and RNA targets. Its use has become synonymous with the hybridisation of DNA probes to specific chromosomal regions and in the investigation of tumour genomes (Plate 1). It allows for a more comprehensive surveillance of chromosomal aberrations and with a higher degree of accuracy than standard GTG-banding techniques (which might otherwise be sufficient for constitutional applications). Such DNA *in situ* hybridisation has had a pivotal role in the elucidation of patterns of deletion and amplifications in tumour. It has been employed as a useful method in the detection of gene and chromosome anomalies in cell lines, touch preparations from fresh tissue, isolated nuclei from formalin-fixed paraffin-embedded tissue and routine histologic sections from paraffin blocks [3].

Given existing cytogenetic evidence indicating that the gain of chromosome 7 was associated with higher tumour grade, Jenkins *et al.* (1998) utilised FISH probes to investigate prostate specimens using a chromosome 7 centromere and five loci mapped to 7q31 [4]. In their study, they noted that gain of 7q31 was strongly associated with tumour Gleeson score. It was noted that the DNA probe for D7S522 spanned the common fragile site FRA7G at the 7q31 position and the authors postulated that the instability in this fragile site could explain both loss and gain of this region as noted by a number of authors. Bostwick *et al.* likewise investigated allelic imbalance at 7q31 and noted that its imbalance was greater in prostate cancer foci than in prostatic intraepithelial neoplasia (PIN) suggesting a role in the progression of precursor lesions to carcinoma [5].

The 8p22 region has also been implicated in prostate tumourigenesis by FISH analysis and widely examined by a number of authors. For example, Emmert-Buck found loss of 8p12-21 in 63% of PIN and 91% of cancer

foci – work independently supported by Bostwick [6]. Given the number of authors highlighting regions of loss on the 8 *p-arm*, it is likely there are a number of putative tumour-suppressor genes located therein, one of which is likely to be the leucine zipper protein, FEZ1 and another the lipoprotein lipase gene (*LPL*) at the 8p22 position. In addition to losses on the 8 *p-arm*, a number of gains on the 8 *q-arm* have also been reported using FISH-based methodologies. Qian *et al.* [7] initially observed that gain of chromosome 8 was the most frequent chromosomal anomaly in metastatic foci and that the frequency was much higher in PIN and in carcinoma. Subsequent to this Jenkins *et al.* implicated the *c-myc* gene (8 *q-arm*) in 22% of metastatic foci which appeared at a higher frequency in these when compared to their primary (9%) counterparts [8]. It appears that an increase in copy number of the *c-myc* gene is associated with both systemic progression and early death of the patient. In this respect, it has been hypothesised that an over-expression of *c-myc* causes the breakdown of p27*kip*1 and a resulting activation of cyclin-dependent kinase 2 and cell proliferation. Of greater interest, however, was the demonstration that those with co-existent loss of 8p22 and gain of *c-myc* had the poorest outcomes overall [9].

Comparative Genomic Hybridisation

In contradistinction to FISH-based methodologies, comparative genomic hybridisation (CGH) represents a powerful molecular cytogenetic technique, which utilises tumour DNA to provide a genome-wide screen for sequence copy number aberrations in a single hybridisation reaction; a marked contrast to the single locus specific probes employed in classical FISH procedures (Plate 2). With tumour and reference DNA labelled differentially (with contrasting fluorophores), samples are mixed and co-hybridised onto normal metaphase chromosomes. By quantifying the hybridisation intensity variations between test and reference DNA samples, relative gains and losses of chromosomal regions can be assessed. Arguably, its principle advantage is that all chromosome regions can be screened for gains and losses simultaneously, giving an unparalleled data mining tool. Visakorpi and others have used this technique to examine in great detail the changes involved in prostate cancer initiation and progression [10]. Indeed, changes at the genomic level (in terms of sequence copy number changes) appear to affect some 74% of primary prostate cancers. The fact that regions of loss are almost five times as common as regions of gain in early disease, is taken by many to suggest the importance of TSG loss in the initiation of tumourigenesis. The regions of loss identified include 6q, 8p, 9p, 13q, 16q and 18q – regions of LOH also implicated in FISH-based studies. The application of CGH has additionally been employed to assess the progressive changes involved with the

transition from benign disease, through high-grade prostatic intraepithelial neoplasia (HGPIN), to localised and eventually metastatic hormone resistant prostate cancer. At the most basic level it appears clear, for example, that genetic aberrations in recurrent tumours are three times higher than in primary tumours, with the number of regions noted as gained almost five times greater than their primary counterparts. Of particular importance in relation to disease progression, however, appears to be the association of 5q loss and gains in 7p, 8q and Xq, in the progression to hormone resistant disease. The candidate genes involved in these regions are keenly sought. Unfortunately however, CGH, with its relatively low resolution (the method can only detect deletions greater than 10 Mb in size) does not allow for any degree of useful specificity – a deficiency that CGH-based microarrays have been able to address directly.

Genomic Microarrays

While chromosome-based CGH (with its whole-genome screening capability) is significantly faster and less laborious than methods examining for single-target dosage changes (Southern blotting, PCR and FISH), there are technical aspects that have limited its usefulness. Principle amongst these is its resolution. Because DNA within chromosomes is tightly coiled and condensed, a 10 Mb span is required to detect a deletion and a 2 Mb length to detect gains – both in part dependent on amplicon size. While providing a starting point for positional cloning studies, the regions contain far too many genes to localise sequences of interest. Another principle limitation is the need to identify individual chromosomes – something even experienced cytogeneticists may find difficult from the inverted DAPI images available. As a result of these difficulties (and with the recent advances in microarray technology), CGH-based microarrays have been developed, which use genomic DNA sequences as targets (Plate 3). DNA targets for these microarrays can be derived from yeast artificial chromosomes (YAC, 2 Mb), bacterial artificial chromosomes (BAC, 300 Kb) and cosmids (45 Kb) all of which are several magnitudes of size smaller than chromosome targets in conventional CGH. As such, they provide for an increase in resolution of copy number over traditional CGH.

We have used commercially available (Vysis) CGH-based microarrays to explore loss of TSG in early stage prostate cancer and the variable gain in copy number of a wide variety of putative oncogenes in the progression to hormone resistant disease. Using prostate cancer cell lines, fresh frozen tissue and tumour material isolated by laser capture microscopy (LCM) from formalin-fixed paraffin-embedded tissue, we have identified a number of common regions of loss and gain in disease initiation and progression.

Using a series of 50 microdissected tumour specimens, we have identified a series of 60 regions that are consistently either amplified or deleted and known to harbour candidate genes. This minimum dataset appears to represent the beginnings of a genomic fingerprint, the significance of which has yet to be fully appreciated. It seems likely however that loss and gain of genomic material in pre-cancerous or early stage disease will identify putative TSGs in disease initiation, while the gain of additional material with progression to the androgen-resistant state will define novel targets for therapeutic intervention [11].

Multiplex FISH and Spectral Karyotyping

While relative loss and gain of genomic material can be assessed by FISH and CGH, it is clear that such approaches cannot reflect the complexity of chromosomal abnormalities in prostate cancer in their entirety. Of particular interest in this respect, is the growing realisation that chromosomal translocations may harbour influential fusion genes that have a significant impact on disease initiation and progression. This has most clearly been demonstrated in the field of haematological malignancy in which a translocation involving chromosomes 9 and 22 brings the *Abl* gene onto chromosome 22. As a result, the Philadelphia chromosome is formed and represents a derivative chromosome 22 : t (9;22). As a result, a fusion protein (termed Bcr-Abl) is formed, a product of chimeric mRNA arising as a consequence of the translocation. This Bcr-Abl product, functions as a constitutively activated tyrosine kinase, which is essential for the transforming function of the protein and plays a pivotal role in chronic myeloid leukaemia (CML). Following a high-throughput screen of chemical libraries searching for compounds with kinase inhibitory activity, a series of compounds originally optimised against the platelet derived growth factor (PDGF) receptor was shown to have equipotent activity against the *Abl* tyrosine kinase. ST1571 or Gleevec is the therapeutic agent that has resulted. Its rapid clinical success in the treatment of CML has already rendered existing treatment algorithms obsolete overnight [12].

Assessing such complex chromosomal translocations in solid tissue malignancies has traditionally been limited, both because of the restrictions imposed by traditional cytogenetics and the difficulties involved in establishing primary prostatic cultures. However, recently developed molecular cytogenetic techniques in the form of spectral karyotyping (SKY) and multiplex FISH (M-FISH) have greatly facilitated these processes. M-FISH is a combinatorial technique that allows the identification of human chromosomes by painting them with a spectrum of DNA probes labelled with a unique

combination of five fluorochromes, such that individual human chromosomes can be identified in 24 discrete pseudo-colours (Plate 4). The SKY methodology represents an essentially similar technique but in the latter case, instead of relying on a series of multiple excitation and emission filters, it relies on Fourier spectroscopy to identify individual chromosomes.

Pan *et al.* [13] were among the first to characterise chromosomal abnormalities in prostate cancer cell lines using SKY. In doing so, a number of novel chromosomal abnormalities were identified, which had remained previously unidentified by conventional cytogenetics – suitably demonstrating the increased sensitivity of the technique. Subsequent to this, Strefford *et al.* [14] reported M-FISH data on a similar group of commercially available malignant prostate cell lines and drew comparisons between those features identified in both SKY and M-FISH studies. One such chromosomal aberration that was identified as common to a number of studies was a reciprocal translocation involving chromosomes 1 and 15. Given the relative rarity of such translocations (especially ones that are recurrent) they postulated a potentially significant role for it in either the initiation or progression of the disease.

This work on M-FISH has recently been extended from work on prostate cancer lines to the analysis of primary cultures. Using a low calcium serum-free medium (PrEGM), we have established a number of malignant primary cell cultures from tissue harvested fresh from radical prostatectomy specimens. [15]. Though generally less complex in terms of karyotype than their immortalised cell counterparts, a number of novel chromosomal translocations have nevertheless been identified, some of which have also been previously identified in cell lines [16].

Some of the translocations identified by a variety of these authors appear common to a number of cell lines and cell cultures. In doing so, it raises the intriguing possibility that one or more breakpoint translocations may be giving rise to common fusion proteins involved in the initiation or progression of the disease. As such, it is possible that novel therapeutic interventions in the form of new drug treatments could herald innovative treatments in much the same way as ST1571 has revolutionised the management of haematological malignancies.

EPIGENETIC CHANGES IN PROSTATE CANCER

While the direct loss of a gene can have a fairly obvious effect on the functional phenotype of a cell, changes relating to promoter hypermethylation have been shown to represent an alternative to Knudson's 'two hit' hypothesis where TSGs are inactivated [17,18]. In essence, it reflects changes in a promoter region (which represents a regulatory DNA sequence upstream of a gene) that involves an increased binding of methyl groups to so-called

CpG islands. The modification of this critical region, which is thought to have significant regulatory effects, causes loss of gene expression insofar as this reversible and epigenetic event inhibits the transcription of genes into mRNA. While the mechanisms involved remain stubbornly obscure, the list of aberrant methylation genes in cancer is rapidly growing. For example, frequent methylation of *DAPK* and *RAR2* has been reported in both lung and breast cancer respectively and there has been widespread speculation about its diagnostic and prognostic implications. Likewise, there has also been growing interest in a possible aberrant promoter methylation profile in prostate cancer and its relationship to clinicopathological features.

Using a *methylation-specific polymerase chain reaction* (MSP) Maruyama and co-workers [19] examined a series of 100 prostate cancer specimens (and 32 benign controls) for promoter hypermethylation in a series of 10 preselected genes. Six of these genes, *RARβ*, *RASSF1A*, *GSTP1*, *CDH13*, *APC* and *CDH1* were found to be selectively methylated. Although in this study, the proportion of samples demonstrating *GSTP1* methylation was relatively low at 36%, other authors have demonstrated its much more pervasive involvement in prostatic malignancy. For example, Jeronimo *et al.* [20] using a fluorogenic and quantitative real-time MSP analysed cytidine methylation in the *GSTP1* promoter in a series of 69 patients with prostate cancer and 31 patients with benign prostatic hyperplasia. With the relative level of methylated *GSTP1* DNA in each case being determined by the ratio of MSP-amplified *GSTP1* to *MYOD1* (a reference gene), 91% of the malignant and 53% of the HGPIN specimens displayed *GSTP1* hypermethylation. Importantly however, some of the tissues from patients with BPH also displayed *GSTP1* hypermethylation. However, the distribution of the ratios of *GSTP1* : *MYOD1* differed significantly when plotted on a log scale. Using this fluorogenic quantitative approach, its sensitivity in the detection of prostate cancer was 85% and its positive predictive values was an impressive 100%. On the basis of these initial results, the authors investigated whether quantitative *GSTP1* methylation could be used to detect prostate cancer in small biopsy specimens. Using a cut-off value of 10 for the methylation ratio, the authors correctly predicted the histologic diagnosis of prostate cancer in 90% of the sextant biopsies and successfully excluded a diagnosis of malignancy in all the 10 patients whose biopsy specimens showed no evidence of malignancy.

The application of this fluorescent MSP technology to the detection of promoter hypermethylation of the *GSTP1* gene in DNA isolated from body fluids has more recently proved a potentially exciting and non-invasive tool for the detection of prostate cancer [21]. In this respect, Goessl and colleagues have isolated DNA from plasma, semen and post-prostatic massage urine in patients with prostate cancer and benign diseases alike. The authors

detected *GSTP1* promoter hypermethylation in 94% of tumours, 72% of plasma, 50% of ejaculate and 36% of urine in patients with prostate cancer and in doing so, proposed MSP as a specific tool for the molecular diagnosis of prostate cancer in bodily fluids.

EXPRESSION PROFILING USING MICROARRAYS

Genetic alterations, regardless of the mechanisms that bring them about, often result in changes in mRNA expression levels which in turn have an impact on the protein expression pattern of a cell and its subsequent phenotype. The use of cDNA arrays allows the quantitative measurement of mRNA in many thousands of genes in a single biological sample. Its conceptual basis is the hybridisation of a complex probe derived from tissue RNA (Complementary DNA, cDNA) to DNA fragments that represent target genes arrayed on a glass slide. The tissue-derived probes are produced by the reverse transcription of RNA accompanied by simultaneous labelling. It is a complex probe, insofar as it contains in solution many different sequences of cDNA in various amounts, corresponding to the number of copies of the original mRNA species extracted from the tissue specimen. This labelled probe is then allowed to hybridise to the array, which may contain thousands of targets derived from PCR-amplified cDNA inserts or oligonucleotides – each representing a gene of interest (or part of a gene). The amount of labelled probe binding to arrays is directly proportional to the level of RNA in the original sample and as such provides a measure of its expression in tissues.

One of the most widely used commercial systems is manufactured by Affymetrix (Santa Clara, CA, United States). In this particular system, manufacture depends on the utilisation of hundreds of thousands of oligonucleotides synthesized directly *in situ* on glass chips by means of a photochemical reaction combined with an innovative masking technology. Each target gene under investigation is represented by a series of oligonucleotides in addition to appropriately mismatched sequences, which provide necessary internal controls. With test and reference probes labelled with different fluorescent markers (in much the same way as with CGH-based microarrays) expression is assessed directly by high-resolution scanners and sophisticated imaging software. Given that, it is now routinely possible to spot many thousands of targets of interest on less than a square centimetre of such arrays, the information provided by such systems are immense and has spawned a rapidly diversifying bioinformatics industry.

One of the most frequently employed systems of data interpretation is based around a hierarchical clustering algorithm [22] originally described by Eisen (1998). With a step-wise analysis of gene expression levels (involving

the establishment of a similarity metric), all genes are incorporated into a dendrogram that connects genes (or nodes) generated by the clustering. Importantly, the length of each branch in the dendrogram reflects the degree of similarity between connected nodes or genes. Finally, software analysis allows the representation of correlated genes, which share similar expression patterns over large groups of specimens analysed, to provide a visual display of similarities between what might have originally appeared as a disparate group of specimens.

A number of research groups have applied this emerging technology to the investigation of both initiation and progression in prostate cancer. Luo *et al.* (2001) performed gene expression profiling on both benign and malignant prostate tissue in an attempt to identify fundamental differences [23]. Using cDNA microarrays consisting of 6500 genes, a series of 210 genes were identified as distinguishing benign and malignant disease. When ranked according to the ability to differentiate benign prostatic hyperplasia (BPH) from malignancy – the number one ranked gene was *hepsin*, a gene that encodes for a transmembrane serine protease, previously implicated in cell growth. Subsequent reverse transcriptase-polymerase chain reaction (RT-PCR) analysis was used to determine expression level of the *hepsin* gene in a number of benign and malignant samples. This analysis confirmed the high expression of *hepsin* in tumour specimens with low or minimal signal in benign specimens. While many of the other identified genes were not independently verified by PCR-based methods, the authors noted that a database search revealed other confirmatory evidence of differential expression of many of the genes identified by microarray.

Dhanasekaran *et al.* [24] in a seminal paper of the same year similarly applied microarray technology to the investigation of benign and malignant prostatic disease. Using a 9984 element cDNA microarray they analysed a variety of benign, hyperplastic and malignant specimens in addition to a number of cell lines with the principal aim of establishing a molecular classification of prostate cancer. Using several methods of gene selection to create a more limited set of targets for further analysis, the authors selected 200 genes (those with the largest effect sizes). From these, a number of candidate genes were identified including *HPN* (*hepsin*), *PIM1*, *LIM*, *TIMP2*, *HEVIN*, *RIG* and *THBS1* (thrombospondin 1) – amongst others. Selected genes identified by microarray were corroborated by northern analysis. *Hepsin*, for example, was 4.3-times up-regulated by microarray and 11.3-fold up-regulated by northern analysis. Subsequent *hepsin* immunohistochemistry on a total of 738 arrayed tissue samples using an affinity purified *hepsin* peptide antibody demonstrated preferential staining of malignant over benign tissue specimens. A similar approach was undertaken with *PIM1* kinase checking expression on high-throughput tissue microarrays;

PIM1 expression being observed as moderate or high in over half of the prostate cancer specimens.

The approach taken by both these sets of researchers, amply demonstrates the ability of DNA microarrays to identify new and potentially useful markers of disease for which diagnostic and prognostic utility appears likely. Their use is set to increase and with larger numbers of genes being analysed with the latest generation of DNA microarrays, further candidate genes are likely to come to the fore.

PROTEOMICS AND PROSTATE CANCER

While there are considerable number of researchers who are content to limit their work to the transcript level, most would agree that any meaningful interpretation of their data ultimately requires an extrapolation of that information to the protein level. An unquestioning acceptance of the validity of this extrapolation is clearly not sufficient. Principal amongst a number of widespread concerns relating to this assumption is the apparent discordance of mRNA expression measurements and co-existent protein levels [25]. As a result, it is crucial to verify such findings at the transcription level by assessing protein expression directly because the phenotype of a given cell is ultimately determined by the composition and activation status of its proteins. In addition, there are a host of post-translational modifications, which can only be determined by proteomic methodologies. As a result there has been growing enthusiasm for both qualitative and quantitative proteomic measurements.

One of the principal tools in the emerging field of proteomics, curiously is not new by itself, in that two-dimensional electrophoresis has been widely used as an investigative tool for more than two decades. Essentially, it involves the resolution of proteins on both their isoelectric point as well as molecular weight. Given its current level of refinement, however, two-dimensional gels rarely resolve more than 1000 proteins. As a result, it has become frequent to apply a series of affinity-based purification strategies to crude lysates to isolate a desired set of proteins before performing electophoresis. In this sense, two-dimensional electrophoresis is not too dissimilar from DNA microarrays – the difference being that instead of giving a transcript expression pattern, a protein expression pattern is achieved instead. A variety of computer software packages are currently available to align gel images and assign cluster indices and gauge relative abundance of proteins at select spots. It is, however, in the combination of two-dimensional electrophoresis and mass spectrometry that proteomics achieves its greatest resolution. Proteins are at first separated by gel electrophoresis and subsequently digested by sequence specific proteases

(endo-peptidase), such that a peptide mixture can be eluted for further processing. Using matrix-assisted laser desorption and ionisation (MALDI) a mass spectrum or peptide-mass fingerprint is obtained. By fragmenting individual peptides to gain sequence information and submitting masses to a database search, individual matches can be made. If the protein is not identified (because of its novel nature or post-translational modification) it can be analysed by tandem mass spectroscopy. A sequence tag can be used for a more specific search or sometimes identified directly from the tandem mass spectroscopy. An alternative strategy involves using surface-enhanced laser desorption ionisation (SELDI) spectrometry. It is a highly sensitive, specific and high-throughput technology for the study of protein lysates. It differs from the conventional MALDI in that it does not rely on pre-clearing complex biological mixtures by high performance liquid chromatography or gas chromatography.

Proteins of interest are directly applied to a surface, utilising a defined chemical chromatographic characteristic such that it allows the construct of reproducible protein profiles. Using an essentially two-dimensional gel approach to the analysis of benign and malignant prostatic disease, Guevara *et al.* (1986) were one of the first groups to demonstrate the utility of a proteomic approach to the identification of new and informative biomarkers in prostate cancer [26]. In their study, a series of nine proteins were identified as being common to a number of malignant samples (and absent from benign tissue) and a further three were only identified in benign tissue (and significantly absent from the malignant cases identified). Present day investigators, who have continued this work, now benefit from recent advances in the biotechnology industry such that they can now begin to accurately identify the proteins identified in these early pioneering studies. For example, Alaiya *et al.* (2001), using a two-dimensional gel electrophoresis followed by gel digestion and MALDI mass spectrometry, compared tissue harvested from patients with benign prostatic disease and those with prostatic carcinoma [27]. In addition to a 40 kDa protein (identified as prostatic acid phosphatase) that decreased two-fold between benign and malignant disease, the authors reported the increased expression of heat shock protein 70 and a decreased expression of tropomyosin 1 in malignant tissue – the significance of which was uncertain.

Larger studies have now demonstrated the ability to differentiate normal, pre-malignant and malignant prostatic tissue on the basis of proteomics alone. Cazares *et al.* (2002) using SELDI spectrometry examined benign, malignant and high-grade PIN tissue obtained from radical prostatectomy specimens [28]. In turn, they determined several small molecular mass peptides and proteins which were up-regulated in prostate cancer and pre-malignant lesions. While there was no single protein alteration observed in all

PIN and prostate cancer specimens, it became readily apparent that the combination of a number of these markers had utility in distinguishing between tissue pathologies. Using a logistic regression analysis involving seven differentially expressed proteins resulted in a predictive equation that correctly distinguished pathological specimens with a sensitivity and specificity of 93.3% and 93.8% respectively.

A particularly innovative application of this proteomic fingerprinting has more recently been applied to the serum of patients with prostate cancer by Bao-Ling *et al.* [29]. The authors noted that while prostate-specific antigen (PSA) measurements had played a major part in increasing awareness and improving disease management, its lack of specificity had seriously limited its usefulness. The authors reasoned that the evaluation of a proteomic signature for prostate cancer could be established from the patient's sera using the protein profiling technologies detailed above. Using SELDI mass spectroscopy combined with an artificial learning algorithm, protein profiles of sera from 167 prostate cancer patients were examined along with age-matched normal controls. Using a nine-protein mass pattern decision tree, the authors correctly assigned 96% of the samples. A subsequent blinded series demonstrated a sensitivity of 83% and specificity of 97% with a positive predictive value of 96% for correctly identifying patients with prostate cancer.

Such proteomic-based approaches (which appear to be growing in popularity for the identification of occult disease in a number of solid tissue malignancies) rely not on individual biomarkers but alternatively on complex patterns or signatures of disease. Given the heterogeneous nature of prostatic malignancy, this approach appears to have at last offered the potential for a reliable and non-invasive test on which the selection of patients for histological sampling might reasonably be based – something which a range of individual biomarkers has so conspicuously failed to do over the last decade.

CONCLUSION

For decades it has been clear that the limitations imposed by standard scientific methodologies on data acquisition, represented the single most important barrier to the elucidation of a variety of clinically significant markers in both disease initiation and progression. Over a period of less than 5 years, however, this stance has radically changed. In contrast to previous decades, scientific and translational researchers are now faced with a deluge of information. The challenge, therefore, appears to be not how much data can be acquired, but how much can be deemed significant above the inevitable hiss of background noise. In this respect, it is likely that researchers will no longer focus on individual markers but instead examine the pattern of changes (whether at a genomic or proteomic level) associated with malignant

transformation. And as is now becoming increasingly clear, the challenge of the future is one of data handling, a role that the emerging field of bioinformatics will have to grasp firmly if we stand any chance of benefiting from the rapid advances of recent years.

REFERENCES

1 Cancer Research Campaign. Scientific Year Book. 2000–2001. CRC.
2 Fearon E *et al.* A genetic model for colorectal tumorigenesis. *Cell* 1990; **61**: 759–67.
3 Schwarzacher and Heslop-Harrison. *Practical in-situ hybridisation*. Bios Scientific Publishers Ltd. (2000).
4 Jenkins R *et al.* A molecular cytogenetic analysis of 7q31 in prostate cancer. *Cancer Res* 1998; **57**: 524–31.
5 Bostwick D. Independent origin of multiple foci of PIN. *Cancer* 1998; **83**: 1995–2002.
6 Emmert-Buck MR *et al.* Allelic loss on chromosomes 8p12-21 in microdissected PIN. *Cancer Res* 1995; **55**: 2959–62.
7 Qian J *et al.* Determination of gene and chromosome dosage in PIN and carcinoma. *Annal Quant Cytol Histol* 1998; **20**: 373–80.
8 Jenkins R et al. Detection of *c-myc* oncogene amplification and chromosome anomalies in metastatic prostate carcinoma by FISH. *Cancer Res* 1997; **57**: 524–31.
9 Sato L *et al.* Clinical significance of alterations of chromosome 8 in HGPIN and prostate carcinoma. *J Natl Cancer Ins* 1999; **91**: 1574–80.
10 Visakorpi T *et al.* Genetic changes in primary and recurrent prostate cancer by comparative genomic hybridisation. *Cancer Res* 1995; **55**: 342–7.
11 Lane TM *et al.* Genomic microarrays in the analysis of prostate cancer progression. *J Med Genet* 2002; **39**: S59.
12 Druker B *et al.* Perspectives on the development of a molecularly targeted agent. *Cancer Cell* 2002; **1**: 31–6.
13 Pan Y *et al.* 5q11, 8p11 and 10q22 are recurrent chromosomal breakpoints in prostate cancer cell lines. *Genes, Chromosomes Cancer* 2001; **30**: 187–95.
14 Strefford JC *et al.* The use of multicolour fluorescent technology in the characterisation of prostate carcinoma cell lines: a comparison of multiplex fluorescent in-situ hybridisation and SKY data. *Cancer Genet Cytogenet* 2001; **124**: 112–21.
15 Hudson DL *et al.* Proliferative heterogeneity in the human prostate: evidence for epithelial stem cells. *Lab Invest* 2000; **80**: 1243–50.
16 Lane TM *et al.* Cytogenetics of primary prostate cultures. *Prostate and Prostatic Diseases* (2002) *in press*.
17 Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *PNAS* 1971; **68**: 820–23.
18 Jones P *et al.* Cancer Epigenetics comes of age. *Nat Genet* 1999; **21**: 163–7.
19 Maruyama R *et al.* Aberrant promoter methylation profile of prostate cancers and its relationship to clinicopathological features. *Clini Cancer Res* 2002; **8**: 514–9.
20 Jermino C *et al.* Quantitation of *GSTP1* methylation in non-neoplastic prostatic tissue and organ confined prostate adenocarcinoma. *J Nat Cancer Inst* 2001; **93**: 1747–52.
21 Goessl C *et al.* DNA alterations in body fluids as molecular tumour markers for urological malignancies. *Eur Urol* 2002; **41**: 668–76.
22 Eisen M *et al.* Cluster analysis and display of genome wide expression patterns. *PNAS* 1998; **95**: 14863–8.
23 Lao J et al. Human prostate cancer and benign prostatic hyperplasia: molecular detection by gene expression profiling. *Cancer Res* 2001; **61**: 4683–8.
24 Dhanasekaran S *et al.* Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001; **412**: 822–26.
25 Anderson L *et al.* A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 1997; **18**: 533–7.

26 Guevara J *et al.* Distinctive protein pattern in two-dimensional electrophoretograms of cancerous prostatic tissues. *Cancer Res* 1986; **46**: 3599–604.
27 Alaiyet AA *et al.* Identification of proteins in human prostate tumour material by two-dimensional gel electrophoresis and mass spectrometry. *Cell Mol Life Sci* 2001; **58**: 307–11.
28 Cazares H *et al.* Normal, benign, pre-neoplastic and malignant prostate cells have distinct protein expression profiles resolved by surface enhanced laser desorption-ionisation spectroscopy. *Clin Cancer Res* 2002; **8**: 2541–52.
29 Bao-Ling A *et al.* Protein finger-printing coupled with a pattern-matching algorithm distinguishes prostate cancer from BPH and healthy men. *Cancer Res* 2002; **62**: 3609–14.