chapter twenty-two

Biomarkers of early effect in the study of cancer risk

Laura Gunn, Luoping Zhang, Matthew S. Forrest, Nina T. Holland, and Martyn T. Smith

Contents

I. Introduction
II. Molecular cytogenetics
   A. Fluorescence in situ hybridization (FISH)
   B. Applications of FISH
   C. Recent developments in FISH technology
III. PCR-based methods for detecting somatic mutations and aberrant gene methylation
   A. Conventional PCR detection of chromosome rearrangements
   B. Quantitative real-time PCR
   C. Another potential application of real-time PCR: measurement of aberrant gene methylation
   D. Multiplex real-time PCR with molecular beacons
   E. The use of Scorpion™ primers
   F. PCR-based detection of point mutations
IV. PCR-independent molecular technologies
   A. Rolling circle amplification
   B. The Invader® assay
V. The future application of arrays and nanobarcodes
VI. Conclusion
Acknowledgments
References
Abstract  Cancer is an abnormal genetic phenomenon, involving multiple steps of somatic mutation. Genetic damage can occur at the level of the gene (e.g., point mutations and deletions) or the chromosome (e.g., aneuploidy, translocations). During the last two decades, a wide spectrum of biomarkers of genetic damage has been developed to detect early mutational and chromosomal effects of carcinogenic exposure in humans. Historically, biomarkers have tended to measure mutations in surrogate genes or use cytogenetics to assess overall changes in chromosome structure and number. These biomarkers have been shown to be associated with a wide range of carcinogenic exposures, but they are not truly biomarkers of early effect as they are not on the causal pathway of disease. Identification of early causal genetic events in cancer has led to the recent development of novel biomarkers of early effect in high-risk populations. These novel biomarkers measure changes frequently observed among cancer patients, including point mutations in genes such as p53 and RAS, altered gene methylation, and aneuploidy (chromosome loss or gain), including monosomy 7 and trisomy 8, and specific chromosome rearrangements such as translocations. Future technologies will measure >50,000 endpoints from a drop of blood using proteomics and array technologies and identify all genetic polymorphisms related to susceptibility using high-throughput genomics. Application of these biomarkers to study individuals who may be at risk, but who do not yet have cancer, will result in improved early detection, as well as a better understanding of the risk factors for cancer itself.

I. Introduction

Carcinogenesis is a complex, multistage process, which involves the accumulation of a variety of mutations within a particular cell and its progeny. Although carcinogenesis depends on a number of different factors including environmental exposure, diet, genetics, and target tissue, certain general characteristics of cancers are known. The identification of the role of particular genes in cancer has opened a new avenue of research over the past two decades. Oncogenes and tumor suppressor genes have taken center stage with their respective roles in cancer. Alterations in these genes ranging from small insertions, deletions, point mutations, and aberrant methylation, to gross chromosomal aberrations, like translocations, and gene amplification, either enhance or inactivate the normal function of the gene and lead to abnormal proliferation, lack of cell cycle control, genomic instability, and eventually cancer. Mutations in these genes provide telltale signs of genetic changes or damage and possible cancer risk, often long before the onset of cancer. Particular genes, chromosomal regions or entire chromosomes are vulnerable to mutation at variable points in carcinogenesis. This suggests that certain mutations play a specific role in the ability of a cell to survive and continue to the next step of this multistep process, as well as potentially determining what the next mutation will be. These mutations, particularly
early events, may provide markers, which are indicative of genetic damage and potential cancer risk.

During the last two decades, a wide spectrum of biomarkers of genetic damage has been developed to detect mutational and chromosomal effects of carcinogenic exposure in humans. Historically, biomarkers have tended to measure mutations in surrogate genes, including hypoxanthine phosphoribosyltransferase and glycophorin A or use cytogenetics to assess overall changes in chromosome structure and number. These biomarkers have been shown to be associated with a wide range of carcinogenic exposures, but they are not true biomarkers of early effect as they are not on the causal pathway of disease. In addition, early biomarkers were not specific enough to identify small or gene-specific alterations, but rather indicated generalized chromosome damage or instability.

As our understanding of the multistage process of carcinogenesis has improved over the years, biomarkers have evolved into more sensitive, specific, predictive markers of cancer risk. Current biomarkers measure changes frequently observed among cancer patients, including point mutations in key genes such as p53 and RAS, aberrant gene methylation, altered gene expression, aneuploidy (chromosome loss or gain), and specific chromosome rearrangements such as translocations.

The evolution of biomarker research has included the identification of causal genetic events as well as the means to study these events. A key contributing factor to the evolution of cancer biomarkers has been the parallel evolution of biotechnology. Technological advances, such as PCR and fluorescent microscopy, have revolutionized the field of cancer research by providing new levels of sensitivity and specificity. Technology will continue to fuel the development of genetic biomarkers as well as enabling the elucidation and analysis of complex phenotypes. These phenotypic markers will soon be analyzed by extremely high-throughput methods, requiring minimal amounts of biological material. This paper will review recent progress in the development of biomarkers of early effect for carcinogenesis and illustrate their use in our own work. Finally, we will describe recent technological advances that are likely to be responsible for the biomarkers of the future.

II. Molecular cytogenetics

A. Fluorescence in situ hybridization

Molecular cytogenetics employing fluorescence in situ hybridization (FISH) has been around for more than a decade and its application in biomarker research continues to grow. Recent technological developments, most significantly rolling circle amplification and automated spot counting, are likely to dramatically increase the utility of FISH in the next few years. FISH has several advantages over conventional cytogenetics, perhaps the most important of which is its rapid ability to detect specific chromosome changes on the
causal pathway to cancer. Interphase FISH, in particular, offers several advantages over classical cytogenetics. First, interphase FISH allows analysis of nondividing cells (Figure 22.1a). Second, a much larger number of cells, at least 1000 or more, may be analyzed. Third, the detection of aneuploidy is facilitated by simply counting the number of labeled regions representing a particular chromosome of interest within the isolated interphase nucleus. By contrast, metaphase FISH can readily detect structural rearrangements in addition to aneuploidy (Figure 22.1b). Furthermore, because metaphase FISH, like classical cytogenetics, analyzes dividing cells, the results from these two methods may be directly compared. A number of studies have determined that FISH is more sensitive and convenient than classical cytogenetics. Therefore, FISH appears to be the more suitable method for large-scale population biomonitoring. However, one of the drawbacks of FISH is its high cost due to a restrictive patent and monopoly of probe distribution.

B. Applications of FISH

One example of a specialized FISH assay primarily employed in radiation research is that developed by Tucker and coworkers. This assay uses single-color FISH by painting the chromosome pairs 1, 2, and 4 (or 3, 5 and 6) the same color, which allows for the detection of numerical and structural chromosome aberrations among these painted chromosomes and structural rearrangements between these and other untargeted chromosomes. This assay has been applied in vitro and in vivo in animal and human studies.
Since radiation is thought to cause equal levels of damage across all chromosomes, and chromosomes 1 through 6 (the largest chromosomes) make up 40% of the genome, it is hypothesized that measurement of damage in these large chromosomes can be extrapolated to the whole genome. This may not be true for chemical exposures as certain chemicals may have selective or preferential effects on certain chromosomes. For example, we showed that epoxide metabolites of 1,3-butadiene had more effect on certain chromosomes than others. Indeed, the hypothesis of equal levels of damage across the genome may not hold true even for low doses of radiation, as inversion of chromosome 10 has been shown to be highly sensitive to low intensity radiation exposure. Interestingly inv(10) rearranges the RET gene and is associated with thyroid cancer, potentially caused by linear energy transfer (LET) radiation.

Our laboratory is currently employing FISH to examine the cytogenetic changes in human blood cells caused by exposure to the established leukenogen, benzene. Our plan is to examine all 24 chromosomes and in particular to examine for chromosome changes associated with the development of leukemia. This study is being performed along with Drs. Rothman and Hayes of the National Cancer Institute (NCI), and Drs. Li and Yin at the Chinese Academy of Preventive Medicine in Beijing. We have already applied various FISH techniques in this collaborative study of 43 Chinese workers highly exposed to benzene (median exposure level = 31 ppm, 8 hr TWA) and 44 frequency-matched controls. To date, five chromosomes (1, 5, 7, 8, and 21) have been examined by metaphase FISH in these highly exposed Chinese workers and their matched controls. Frequencies of monosomy 5, 7, and 8, but not 1 or 21, increased with elevated exposure levels, whereas a significant trend was observed for trisomy of all five chromosomes. The most striking dose-dependent increases were found in monosomy 7 and trisomy 7, 8 and 21. The most common structural changes detected among chromosomes 1, 5, 7, 8, and 21 were t(8;21), t(8;?) (translocation between chromosome 8 and another unidentified chromosome), breakage of chromosome 8, and deletions of the long (q) arms of chromosomes 5 and 7. A significant trend was observed for all these changes. The loss and long arm deletion of chromosomes 5 and 7, two of the most common cytogenetic changes in therapy and chemical-related leukemia, were significantly increased in benzene-exposed workers over controls.

C. Recent developments in FISH technology

Recently, multicolor FISH (mFISH) and color banding have been developed. The mFISH method involves painting each of the 24 different chromosomes a different color using four or five fluorophores with combined binary ratio labeling, which allows the entire karyotype to be screened for chromosome aberrations (Figure 22.1c). Since the human eye cannot effectively distinguish the 24 colors, this method requires the use of an automated imaging system. In color banding, which is based on traditional banding techniques,
each chromosome is labeled by subregional DNA probes in different colors, resulting in a unique chromosome bar code (Figure 22.1d). This method allows the rapid identification of chromosomes and chromosome rearrangements. These techniques are at present relatively new and have not been employed as widely or extensively as FISH, but their potential is high.

Of great promise in its application to FISH is rolling circle amplification (RCA), which is a molecular methodology that allows visualization by FISH of small changes previously undetectable by microscopy. It relies upon the isothermal linear amplification of a single stranded circle of DNA. A variation of RCA uses a probe with two 3’ ends, a target specific sequence (40–50 nt) attached by a poly-T linker and (CH₂)₁₈ spacer to a primer (24–28 nt), which will initiate an RCA. When the probe is added to cells fixed onto microscope slides, the target specific sequence anchors the probe to a region of DNA and stops it from being removed during washing. Thus, when RCA occurs the product is attached to the genomic DNA; when decorator probes and biotin labeled dUTPs are added, the linear repeat can be collapsed by avidin to give a localized area of fluorescence. This technique means that much smaller fragments of DNA than would be apparent by traditional FISH can be visualized, including small deletions, duplications, point mutations, and even single nucleotide polymorphisms.

III. PCR-based methods for detecting somatic mutations and aberrant gene methylation

Chromosome translocations and other structural rearrangements produce novel fusion genes or products that can be detected at the DNA or RNA level by the polymerase chain reaction (PCR) or reverse-transcriptase PCR (RT-PCR) as well as by FISH. PCR holds a number of advantages over FISH, including: the ability to detect rare events (1 copy/10⁶–7 cells vs. 1/10⁴ cells by FISH) and the ability to study large numbers of people easily and at low cost. These potent advantages are accompanied, however, by two disadvantages. First, the high sensitivity of PCR makes it prone to false-positive results caused by sample contamination. However, contamination artifacts can be overcome with extremely rigorous lab procedures as well as by PCR and uracil glycosylase in PCR reactions to prevent carryover contamination. Second, until recently, quantitation was difficult, especially for RT-PCR. Quantitation has also become feasible through recent advances in exonuclease-dependent real-time PCR. This quantitative PCR assay, now generally called real-time PCR, allows for the absolute number of novel sequences to be quantified in a cell population. Real-time PCR is more sensitive than conventional PCR, where a sensitivity of 1 in 10⁷ can be reached if a stochastic multitube approach is taken. This technology has, therefore, paved the way for a new generation of biomarkers to be developed. Though no methods yet exist which employ PCR to measure rare aneuploidies or genome-wide structural damage, real-time and conventional PCR techniques which
measure specific chromosome rearrangements, such as translocations, inversions, and the methylation status of genes, have become available.

**A. Conventional PCR detection of chromosome rearrangements**

RT-PCR and PCR have previously been used to detect a number of translocations including t(14;18), t(8;21), t(9;22), and t(4;11). Using these techniques, t(9;22) and t(14;18) have been detected in unexposed individuals of different ages and in smokers. Both translocations were found to increase with age and the t(14;18) translocation was increased in cigarette smokers. Studies from our laboratory showing detectable t(8;21) by RT-PCR in an otherwise healthy benzene exposed worker clearly demonstrate the potential of RT-PCR for monitoring specific aberrations in populations exposed to suspected or established leukemogens. Because many of these translocations have multiple breakpoints or translocation partners, multiplex assays have also been developed to detect multiple or unknown rearrangements. Despite recent improvements in sensitivity and applicability, conventional PCR methods remain semiquantitative. However, with the recent advent of real-time PCR, quantitation is no longer an obstacle. Now that quantitation problems can be overcome, a new avenue of biological monitoring for early detection of cancer has been opened. PCR-based procedures therefore hold further promise for detecting specific chromosome aberrations, especially when used in combination with FISH.

**B. Quantitative real-time PCR**

Real-time PCR is comparable to conventional PCR in that it uses sense and antisense primers to amplify a targeted sequence of DNA. However, real-time PCR techniques such as TaqMan, molecular beacons, and Scorpion probes, rely upon fluorescent signals to quantitate the rate of amplification as the reaction progresses. TaqMan employs an additional, nonextendable oligonucleotide probe, which is positioned between the two primers during the annealing phase of amplification. The oligonucleotide probe is labeled with a fluorescent reporter dye, such as FAM (6-carboxyfluorescein), at the 5’ end and a quencher fluorescent dye, such as TAMRA (6-carboxy-tetramethyl-rhodamine), at the 3’ end. When the probe is intact, fluorescence resonance energy transfer (FRET) to TAMRA quenches the FAM emission. During the extension phase of amplification, the Taq polymerase extends the primer to the region of the probe, at which point the 5’ exonuclease property of Taq cleaves the reporter dye from the probe. This results in an increase in fluorescent signal that is proportional to the amount of amplification product. The increase in reporter molecules is measured in real time by the ABI Prism 5700 or 7700 Sequence Detection Systems (PE Applied Biosystems). After each cycle, a fluorescent signal is measured resulting in an amplification plot, in which the point at which the fluorescence crosses a defined threshold, Ct, that is proportional to the starting copy number. Cts
of positive control samples are used to generate a standard curve. From this standard curve, it is possible to calculate copy number of unknown samples. Methods for the quantitative detection of translocations using the above TaqMan technology have recently been reported. For example, methods for the analysis of t(14;18), t(8;21), t(9;22) and other translocations have been presented or published.25,32–35

We have further developed and refined these real-time RT-PCR methods so that they permit the quantification of t(8;21), t(15;17), inv(16), t(11q23), t(14;18), t(12;21), and t(1;19) fusion gene expression at low levels (1 transcript per 100,000 cells). This has enabled us to examine the levels of these translocations in adult and cord blood of control individuals, in workers exposed to benzene and in patients given the flavonoid quercetin therapeutically. We have found that a significant number of control individuals express measurable levels of the specific translocations and that mitogen-stimulated blood culture enhances the expression of several translocations. Data from benzene-exposed workers and quercetin treated patients are currently under analysis, but in vitro studies in cord blood have demonstrated the ability of

Figure 22.2 Diagram of TaqMan technology in quantitative PCR. (Adapted from PCR Applications, 1999.31)
the benzene metabolite hydroquinone to produce t(14;18), providing a mechanistic basis for the production of lymphoma by benzene.\textsuperscript{36}

C. Another potential application of real-time PCR: measurement of aberrant gene methylation

In addition to the different types of genetic damage involved in carcinogenesis, epigenetic mechanisms, such as DNA methylation, have gained attention as potential key players in certain cancer types. Aberrant methylation, which may be induced by environmental exposures, may result in altered carcinogen metabolism, cell cycle regulation, and DNA repair. For example, in leukemia and lymphoma, translocations cause the formation of novel fusion genes that produce excessive growth,\textsuperscript{37–38} and other genes undergo transcriptional silencing by methylation, which causes aberrant cell cycle control.\textsuperscript{39} Aberrant methylation and transcriptional silencing appears to be an early event in both solid tumors, including lung,\textsuperscript{40} colon,\textsuperscript{41} hepatocellular,\textsuperscript{42} and bladder,\textsuperscript{43} as well as hematologic malignancies.\textsuperscript{39} A number of different methods have been developed to detect aberrant methylation of genes, including the use of methylation sensitive restriction enzymes, bisulfite sequencing, and methylation-specific PCR.

Perhaps one of the most interesting targets of aberrant methylation is the tumor suppressor gene p16\textsuperscript{INK4a}, which is a key component in the G1/S cell cycle checkpoint and has been shown to be involved in colon cancer, leukemia, and lung cancer. Recently, Lo et al. have developed a real-time methylation-specific PCR protocol\textsuperscript{44} and applied it to bone marrow samples of patients with multiple myeloma as well as cell lines with known methylation status. The authors demonstrated that the real-time method had high concordance with the conventional method, however, with the added sensitivity and specificity of the real-time technology. In addition, the authors correlated methylation status with p16 mRNA expression and observed transcription was inversely correlated with methylation status. As with other real-time methods, this application shows great potential for future studies involving methylation of key genes in carcinogenesis as well as other biological processes.

D. Multiplex real-time PCR with molecular beacons

Yet another novel application of real-time PCR technology is the multiplex amplification of DNA by Vet et al.\textsuperscript{45} Using molecular beacons rather than TaqMan probes, the authors successfully performed multiplex PCR reactions which were monitored in real-time. Although similar to TaqMan probes, molecular beacons do not require the 5’ exonuclease properties of Taq polymerase. Instead, when unbound, the fluorescent moiety is kept in physical proximity to the quencher in a loop conformation. When the probe anneals to a target sequence, the loop is linearized and the reporter and quencher are separated in space, and the reporter emits fluorescence. In a multiplex
reaction, different colored fluorescent moieties are used for each amplicon. The potential applications of this assay are promising. Because research potential is often limited by the amount of material available for analysis, the possibility of real-time multiplex PCR provides the opportunity to examine multiple points of interest simultaneously with the same amount of material previously required for one point.

E. The use of Scorpion™ primers

An alternative to molecular beacons and TaqMan are Scorpion primers. While TaqMan and molecular beacons rely upon a bimolecular approach, with separate probes and primers, Scorpion primers use a unimolecular method with an integral tail that is used to probe an extension product of the primer. Scorpion primers consist of a normal PCR primer linked to a nonamplifiable monomer, such as hexethylene glycol, which prevents copying of elements further in the 5’ direction, a nonfluorogenic (dark) quencher of the fluorophore, a probe element flanked by self-complementary stems (hairpin), and a fluorophore. In its unextended form, the Scorpion is non-fluorescent (the fluorophore is quenched). During PCR, the primer element of the Scorpion is extended at its 3’ end and the Scorpion becomes a full PCR product. The recognition sequence of the Scorpion then hybridizes to its complementary target sequence within the same strand of the PCR product and fluorescence is emitted. A recent paper suggests that this means Scorpions perform better than TaqMan or molecular beacons, especially under fast cycling conditions.

F. PCR-based detection of point mutations

The use of PCR technology has vastly improved detection and identification of mutations in cancers. Increased sensitivity and reproducibility has provided the possibility of utilizing these mutation assays as biomarkers of early effect, detection of minimal residual disease, or precursors to relapse. Because of the low frequency of many of these mutations in the normal population, the normal background levels and variability have not yet been established. Recently, a number of assays have been developed which improve sensitivity orders of magnitude over previously used methods. Many of these assays employ methods to selectively amplify the relative number of mutants in a large pool of wildtype in order to increase the sensitivity of detecting rare mutant alleles, a method referred to as genotypic selection.

One recently published assay used genotype selection for the detection of mutations in the H-RAS gene. By combining two previously published methods, the Mut-Ex + ACB-PCR technique is one of the most sensitive genotypic selection methods. This assay begins with the denaturation of a heterogeneous sample of mutant and wildtype double stranded DNA. When reannealing, mutant DNA forms heteroduplex DNA with normal strands, while normal DNA strands form homoduplexes. Mut S, a thermostable
protein, is added. This binds to the mispaired sequence of the heteroduplex, which protects the short sequence of mutant DNA from digestion from 3’–5’ exonuclease activity of T7 DNA polymerase, whereas the wildtype DNA is digested. This Mut-Ex step results in a 1000-fold enrichment of mutant alleles relative to wildtype. To further increase sensitivity, the next step utilizes an additional selection technique, ACB-PCR: (allele-specific competitive blocker PCR). This genotypic selection method is based on preferential amplification by allele specific primers. The first primer has more mismatches to wildtype than mutant, resulting in preferential amplification of mutant DNA. The second primer is a blocker primer which preferentially anneals to the wildtype sequence, but is modified with a 3’ dideoxyguanosine residue, which prevents extension. ACB-PCR therefore results in preferential amplification of mutant DNA with a sensitivity of as few as ten mutant alleles detected in the presence of $10^8$ copies of the wildtype allele.

As one of the most sensitive methods available for mutation detection, the MutEX+ ACB-PCR technique has many potential applications. This method is based on increasing the ratio of mutant DNA relative to wildtype and is therefore a sensitive method for the detection of rare mutations. However, this method is not appropriate for unknown mutations, as the sequence of the mutated region is necessary for the design of ACB-PCR primers.

Genotypic selection methods have also been applied to $p53$ mutation detection. Sites which are commonly mutated in the $p53$ gene, referred to as mutational hotspots, have been targeted as potential biomarkers of early effect. Assays utilizing allele specific PCR have been designed to detect and preferentially amplify mutations in these hotspots. These assays either used alone or in combination with single strand conformational polymorphism and sequencing result in a considerable improvement in sensitivity over conventional methods of mutation detection.\textsuperscript{31}

\section*{IV. PCR-independent molecular technologies}

Traditional mutation detection and genotyping, be it RFLP, sequencing (normal, mini-, or pyro-), SSCP, etc., have relied upon one technology, PCR. However there are two new methodologies which break free from the restrictions of PCR: rolling circle amplification and the Invader$^\text{\textregistered}$ assay.

\subsection*{A. Rolling circle amplification}

Rolling circle amplification has a variety of potential uses, all of which rely upon the isothermal linear amplification of a single stranded circle of DNA. Padlock probes use rolling circle amplification to identify small sequences, polymorphisms, or mutations. A linear probe is designed to have target-complementary segments of 15–20 nucleotides at both ends, separated by a linker of approximately 50 nucleotides. The linker can include sequence elements that allow for amplification or identification of individual probes. Upon hybridization to a target DNA or RNA sequence, the two ends of the
probe become juxtaposed and can be joined by a DNA ligase, ensuring a similar level of specificity to that of PCR. Circularized probes can then be detected in a number of ways. For example, the primer can be designed to hybridize to the padlock with which primer extension creates a linear tandem repeat of the circle that can be probed with labeled decorator oligos. Hyperbranched RCA occurs if a second primer is added which copies the RCA product, resulting in faster replication than normal RCA. Digestion with a restriction enzyme, known to cut the padlock only once, allows the resolution of the product on an agarose gel. It is also possible to use PCR to amplify a padlock probe if the primers used amplify across the ligation point.

This method can also be used for in situ allelic discrimination. For allelic discrimination, there are two probes which only have 20 nucleotide target specific sequences which differ at the 3’ nucleotide. A single anchor probe is designed to hybridize adjacent to the rolling circle probes. When ligase is added, the anchor probe will ligate to the allele discriminating probe only if the 3’ nucleotide matches the target sequence. Amplification can then proceed as described above. Genotype can then be determined by using different labels for each decorator probe.

B. The Invader assay

The Invader assay (Third Wave Technologies, Inc., Madison, WI) is a platform-independent homogenous single-tube assay that does not involve PCR, restriction digests, or gel electrophoresis. Invader assays rely upon the creation of a unique substrate for Cleavase®, a structure-specific 5’ nuclease, by the annealing of a probe and an upstream oligonucleotide to a target sequence. The probe contains two regions, an analyte-specific region that forms a duplex with the target and a noncomplementary 5’ arm region, which serves as a reporter molecule precursor. Cleavage of the probe only occurs when the probe and upstream oligonucleotide overlap; therefore, two target DNAs differing by a single nucleotide may be discriminated. The cleaved arm of the probe then forms cleavage structures with FRET (fluorescence resonance energy transfer) cassettes, labelled with reporter and quencher dyes. Cleavage results in the release of the reporter dye generating an increase in fluorescent signal. Therefore, Invader can be used as an alternative to real-time PCR as well as for allelic discrimination and mutation detection.

V. The future application of arrays and nanobarcodes

A wide variety of cDNA microarrays are now available for studying gene expression. To date, these have found only limited application in biomarker research. This is probably due to at least two reasons. The first is expense. Academic researchers have been hindered in the application of this technology because of its high cost. However, competition, consortia, and new technologies (such as Operon’s 70mer oligonucleotides) are driving prices downward.
The second is the need for large amounts of high-quality RNA. cDNA microarrays rely on the RT-PCR reaction, which we have recently shown is relatively inefficient using commercial enzymes (Curry J., McHale C., and Smith M.T., manuscript in preparation). Thus, large quantities of RNA are required as starting material, somewhere in the order of 50 micrograms. Again new advances are driving this amount down, but a second problem is that most epidemiologists do not collect biological material in such a fashion that RNA is preserved. RNA is inherently a highly unstable molecule and thus its isolation and storage are challenging. We are currently working on protocols for the field collection of RNA from samples obtained in biomarker studies.

Investigators also realize that gene expression is not necessarily phenotype. Thus, protein arrays and proteomics have become of particular interest. Protein arrays hold great potential in biomarker research as they may identify a pattern of protein expression associated with a particular exposure or the early onset of disease. There are also large banks of biological material available for analysis by such arrays.

Possibly of even greater potential is the future application of nanotechnology, including Nanobarcodes™. A company called Surromed (California) is attempting to measure over 50,000 endpoints in a single drop of blood. In order to do this, it is developing several new technologies including Surro-scan to replace flow cytometry and Nanobarcodes labelled with antibodies so that thousands of proteins can be recognized in a drop of plasma. Nanobarcodes are freestanding, cylindrically shaped metal nanoparticles whose composition can be varied along length, such that more than seven metals can be used to produce a nanoscale supermarket-like bar code. Their identification is based on differential reflectivity using microscopy and they can be functionalized with proteins, oligonucleotides, and organics. They can also be made magnetic. This allows for a wide variety of applications including multiplexed analysis of gene expression and multiplexed immunoassays with 10–20,000 flavors. Molecular epidemiologists would benefit by recognizing this coming technology and begin storing small amounts of cryopreserved whole blood as well as the usual serum, buffy coat, etc., that are conventionally stored.

VI. Conclusion

This is an exciting time in biomarker research. Technologies are emerging which will make a significant impact on the field. The potential for improved early detection of cancer is apparent. In addition to enhanced sensitivity and specificity, many new technologies emphasize functional aspects of mutation which require RNA, intact cryopreserved cells, and undamaged proteins. It is therefore important to recognize that biological samples will have to be processed in a more sophisticated manner to ensure their preservation and utility in future biomarker research. Further, it is clear that a significant effort is necessary to lower the cost of these new technologies if they are to be applied in developing countries where they are most needed.
Acknowledgments

Our work on biomarkers is supported by the National Institute of Environmental Health Sciences through grants P42ES04705, P30ES01896 and RO1ES06721 and by the National Foundation for Cancer Research. LG is supported by training grant T32ES07075.

References

29. Fuscoe, J.C. et al., Quantification of t(14;18) in the lymphocytes of healthy adult humans as a possible biomarker for environmental exposures to carcinogens, *Carcinogenesis*, 17, 1013–1020, 1996.


