DDT Vol. 8, No. 3 February 2003

Microarrays: the use of oligonucleotides and cDNA for the analysis of gene expression

J. Carl Barrett and Ernest S. Kawasaki

Completion of the human genome sequence has made it possible to study the expression of the entire human gene complement (>30,000 estimated genes). Aiding in this remarkable feat, DNA microarrays have become the main technological workhorse for gene expression studies. To date, detection platforms for most microarrays have relied on short (25 base) oligonucleotides synthesized *in situ*, or longer, highly variable length DNAs from PCR amplification of cDNA libraries. A third choice, long (50–80 base) oligonucleotide arrays, is now available and might eventually eliminate the use of cDNA arrays. The technology has advanced to such a point that researchers now demand microarrays that are cost-effective and have flexibility and quality assurance. Short- and long-oligonucleotide technologies offer such advantages, and could possibly become the major competing platform in the near future.

J. Carl Barrett
Center for Cancer Research
National Cancer Institute
31 Center Drive, MSC-2440
Bethesda
MD 20892, USA
Ernest S. Kawasaki*
NCI Advanced Technology Center
8717 Grovemont Circle
Gaithersburg
MD 20877, USA
*tel: +1 301 435 2891
fax: +1 301 402 3134
e-mail: kawasake@mail.nih.gov

▼ The millennium of 2000 has ushered in what could be considered as a new era for research into the molecular biology of humans. In early 2000, two research groups –the International Human Genome Sequencing Consortium and Celera – announced the sequencing of the entire human genome, comprising >3 billion base-pairs of nucleotides. The results of their studies were published simultaneously in the journals *Science* and *Nature* [1,2], and indicated that the number of genes in the human genome was much smaller than expected, that is, in the order of 30,000–40,000 genes compared with the original estimate of ~100,000 genes. This smaller number of genes was disappointing to those who believed humanity to be top of the evolutionary heap, especially because the genome of the modest nematode *Caenorhabditis elegans* was found to encode 19,000 genes [3], and the rice (*Oryza sativa*) genome topped the poll with up to 50,000 genes [4,5].

Historical perspective

This wealth of genomic information has enabled researchers to begin the study of the expression and function of every gene in the human body. But how does one begin to analyze the biology of >30,000 genes simultaneously in one experiment? One answer lies in microarray technology, which, although introduced in the 1990s, has only recently come into widespread use. Although the first gene expression microarray article was published in 1995 [6], the concept of miniaturized ligand-binding assays was developed in the mid- to late-1980s [7]. For example, microspot fluorescent immunoassays were described by Roger Ekins in 1989 [8], and the use of double-fluorescent labels for measuring protein analyte concentration was reviewed by the same author in 1989 [9]. Although not exactly equivalent, the use of fluorescent tags to measure nucleic acid hybridization is the mainstay of most array studies, and is a logical extension of earlier immunoassay studies. Thus, conceptually, the first microarrays were of the antigen–antibody type, an area of research that is also beginning to expand rapidly [10–14].

Technology overview

The microarray field is a good example of the assembly and convergence of several technologies, including automated DNA sequencing, DNA amplification by PCR, highly efficient oligonucleotide synthesis, nucleic acid labeling chemistries, and bioinformatics. A microarray can be thought of as a miniaturized gene-hybridization or -detection assay. Instead of measuring signals in assays at the macro level, such as in microtiter plates, membrane blots and test tubes, individual microarray assays

1359-6446/02/$ – see front matter Published by Elsevier Science Ltd. PII: S1359-6446(02)02578-3

or elements are measured in microns. For example, such 'micro-elements' range from 20 to 200 µm, as opposed to 5000–10,000 µm for the bottom-surface dimensions of microtiter plates. One platform used for printing microarrays is the common microscope slide, with dimensions of 25 mm × 75 mm. The latest robotic printers can easily fit 50,000 spots or elements onto a slide if the spots are 100 µm in diameter and spaced 50 µm apart. Thus, it is possible to place the entire human gene complement on one slide – a remarkable accomplishment.

Each of the elements contains a DNA sequence from one gene, and is used to measure the expression of its corresponding mRNA (mRNA) in a cell or tissue sample. Messenger RNA from control and experimental samples are used to synthesize fluorescently labeled cDNA or RNA probes, which are then hybridized to the microarray elements. The fluorescent signal of the hybridized probes is measured with a laser scanner capable of detecting emission from a variety of fluorescent dyes. The intensity of the signals from control and experimental samples are directly correlated with the original concentration of mRNA in the cell or tissue, and can therefore be used to deduce whether the expression of a particular gene is upregulated, downregulated, unchanged or absent. The sensitivity of the assay is high; microarrays have been reported to detect the presence of one mRNA per cell, that is, a concentration of one mRNA per ≥100,000 molecules. This is the basic method by which investigators are now able to study the expression of thousands of genes simultaneously in the cell or tissue type of their choice.

Microarray platforms

The following section describes several microarray platforms used for the study of gene expression. This is not meant to be an exhaustive review of all available

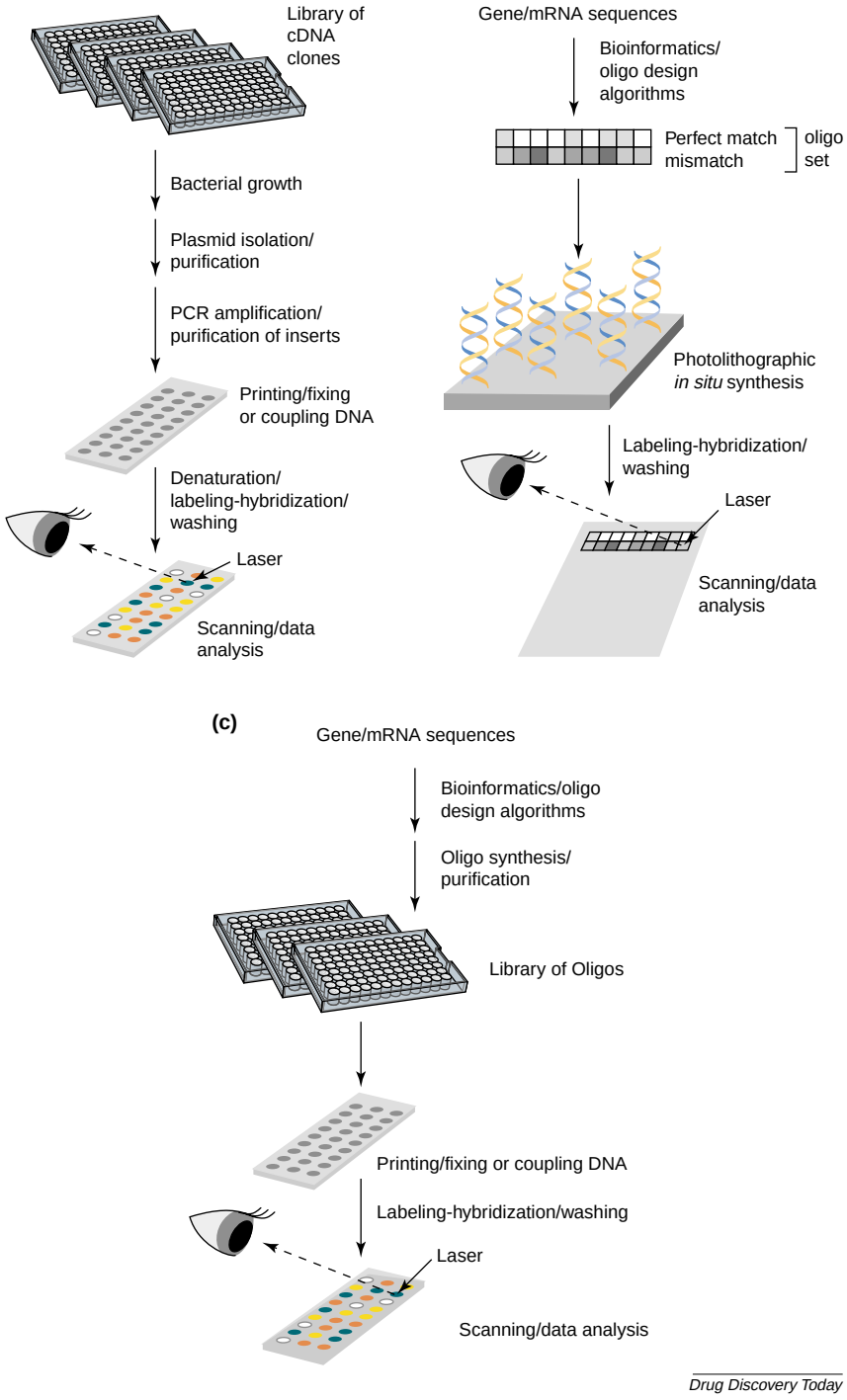

Figure 1. The three microarray platforms. **(a)** cDNA arrays. A library of cDNA clones contained in 96- or 384-well plates are cultured in bacteria. CDNA inserts from plasmids are amplified and products are spotted onto activated glass slides (25 mm x 75 mm). Probes from mRNA are labeled, hybridized to slides and scanned. **(b)** Short oligonucleotide arrays. Pairs of oligonucleotides with exact sequence and one mismatch are synthesized *in situ*. Probes are hybridized to the set that are actually scattered throughout chip and not in one row as shown. **(c)** Long oligonucleotide arrays. Algorithms are used to design long oligonucleotides of the same size, but with little sequence homology. Protocols for printing, hybridizing and analyses are similar to cDNA arrays.

types, but covers those most relevant to the subject of this article.

cDNA arrays: robotic printing

Before the availability of complete or near-complete eukaryotic genome sequences, genes expressed in cells, tissues or organs were identified through sequence analysis of cDNA banks. cDNA clones from cDNA banks of *Arabidopsis* and human peripheral blood lymphocytes were used in the construction of the first cDNA expression arrays [6,15]. The inserts of each plasmid were PCR amplified using universal primers corresponding to the vector sequences adjacent to the gene sequences. The amplified products were analyzed by gel electrophoresis, quantitated, arrayed by robotic printing onto surface-modified glass slides, fixed, and then used to study gene expression as described previously (Fig. 1a). Amplified DNAs can be fixed onto the slides electrostatically, or through crosslinking by heat or UV. Alternatively, the PCR products can be covalently attached at their 5′ ends to modified slides via an amine or other active group. This method, or minor variations thereof, has been the foundation of most gene expression research where the starting material is plasmids from cDNA banks [16–18].

Short oligonucleotide arrays: in situ synthesis

The first large-scale manufacturing of microarrays was developed by Affymetrix (http://www.affymetrix.com) using photolithographic methods similar to the production of computer chips [19] that were subsequently adapted to gene expression studies [20,21] (Fig. 1b). In brief, synthetic linkers with photolabile protecting groups are attached to a glass substrate, and a mask is used to direct light to predetermined areas on the substrate to remove the exposed groups. These de-protected groups are then available for reaction with bi-functional deoxynucleosides, resulting in chemical coupling. A new mask is used to direct coupling at other sites, and the step is repeated until the desired sequence and length of oligonucleotide is synthesized. The present format is a 1.28×1.28 cm chip containing up to 500,000 different oligonucleotide sequences of 25 bases in length, with each element being 18×18 μm in size. Eleven-to-twenty different oligonucleotides are made for each transcript or gene sequence to 'tile' or cover a portion of the 3′ end of the mRNA. Oligonucleotides are synthesized as perfect match (PM) and mismatch (MM) pairs [20,21]. The MM oligonucleotide has a one-base mismatch in the center position, and is used as a control to detect background noise and cross-hybridization from unrelated probes. The latest chip design contains the sequences of up to 20,000 genes, but there is space for more genes if fewer oligonucleotides are used to query each gene. A new approach for synthesizing oligonucleotides photolithographically using digitally controlled micro-mirrors [22–24] could simplify the production of arrays, but the technology is not yet widely available. Further information on similar technologies is available at company websites (febit, http://www.febit.de; Combimatrix; http://www.combimatrix.com).

Long oligonucleotide arrays: in situ synthesis

The photolithographic method is very efficient at producing thousands or millions of identical arrays, but is somewhat inconvenient and expensive for the creation of new arrays with added or different gene content. Technology has progressed to such a point that publication of a new genome sequence has become commonplace; therefore, the flexibility in designing and producing new arrays to capitalize on this wealth of sequence data is an important issue. The production of new arrays via the synthesis of long, 60-mer oligonucleotides by an ink-jet printing process addresses this point [25–27]. In this method, modified ink-jet pumps, similar to those used in printers, are used to dispense 100-picoliter reagent droplets onto a hydrophobic surface containing chemically active hydroxyl groups. The droplets contain phosphoramidite DNA monomers that react and are covalently bound. After washing and de-protection, the process is repeated until the desired oligonucleotide length is reached. The advantages of the *in situ* inkjet method are that no masks are required, synthesis is faster because each cycle attaches one base (four cycles per base are required with photolithography), and new arrays can be created by simply programming the computer with directions on how to synthesize the new set of oligonucleotide sequences. This versatile system can routinely produce arrays with >25,000 elements [25–27].

Long oligonucleotide arrays: robotic printing

As described previously, the flood of new sequence data enables researchers to design their own arrays *in silico*, completely bypassing the need to use cumbersome cDNA banks (Fig. 1c). As the cost of manufacturing oligonucleotides has been significantly reduced, obtaining a large gene-library of oligonucleotide sequences is now within the reach of most researchers. Long oligonucleotides are defined here, somewhat arbitrarily, as 40–80 bases in length. At this length they can be treated as 'short cDNAs' and there is much less chance of spurious cross-hybridization with unrelated sequences compared with short oligonucleotides. For those experienced in using cDNA arrays, working with long oligonucleotide arrays is a relatively easy transition. The oligonucleotides can be printed onto the same substrates (slides) as cDNAs, with the same printing device,

the hybridization and washing conditions are similar, and no new analytical programs for expression analysis are required [28-37]. The design of the oligonucleotides, however, is of upmost importance. Commercial sources have proprietary algorithms for this purpose, although there are several programs in the public domain [38–41] if one wishes to create their own set of sequences. Basically, the oligonucleotides should have very similar melting temperatures or G–C (guanosine–cytosine) content, have very little homology with other oligonucleotides, be entirely contained within an exon, and have no repetitive- or hairpin-sequences. Oligonucleotides of any size can be printed onto appropriate substrates, and complex libraries from several commercial sources are now available (ClonTech, http://www.clontech.com; Compugen, http://www.cgen.com; MWG Biotech, http://www.mwg-biotech.com; Operon, http://www.operon.com). This platform is probably the most amenable for researchers who print their own arrays owing to its flexibility, and facile creation of both very high- and low-density arrays.

Comparison of expression experiments

A PubMed query of microarray literature from January 1995 to October 2002 yields a total of >2300 hits. Most of these appeared in the past two years, indicating an exponential increase in the number of microarray publications. The majority of these publications reported the use of short oligonucleotides or cDNAs; only a small number reported the use of long oligonucleotide arrays, as this technology is only now becoming available. Probably owing to cost and time, there have been very few attempts to replicate and/or compare expression data across different platforms. In addition, these types of studies are hindered by many array results not being confirmed by other methods such as northern blots [42] or quantitative RT–PCR [43,44]. In addition, there is a general lack of standardized experimental design or controls, so it is almost impossible to compare results across platforms (or even within platforms), and large-scale management and analysis of expression data becomes very difficult [45–51].

These problems are being addressed by the Microarray Gene Expression Data Group (MGED group) – an international consortium of microarray enthusiasts [52,53]. Their efforts have resulted in the journal *Nature* [54] having relatively strict guidelines for the submission of microarray data for publication [Minimum Information About a Microarray Experiment (MIAME) guidelines]. Hopefully, other journals will follow suit [55], as such rules will be extremely helpful in removing much of the confusion surrounding the interpretation and replication of microarray data.

cDNA arrays vs short oligonucleotide arrays

NCI cancer cell-line comparisons

Are expression results from different microarray platforms directly comparable? The answer at this point is 'yes and no', depending on the method of analysis. For example, results published from experiments using Stanford-type cDNA arrays and Affymetrix chips were used to examine the suitability of applying cross-platform data in measuring gene expression [56]. Messenger RNA measurements from 2895 matched genes in 56 cell-lines from the National Cancer Institute's standard panel of 60 cancer cell-lines (NCI 60) were used to compare the two platforms. In general, there was poor correlation between the two platforms in all measurements of similarity, such as clusters of genes and cell-lines, and the study suggested that cDNA array data cannot be combined with the short oligonucleotide array results. As a cautionary note, the authors state that the experiments were carried out at different times, at different laboratories, and using different materials and protocols, and this might explain the discordant results. However, if the results from the two platforms are so dependent on protocol, material and time, as far as the characterization of cancer lines goes, the data could be considered useless. By contrast, the optimist might say the data tell us something about the biology of the cell lines. The authors state that, without a third source of data, they could not recommend one platform above the other.

cDNA vs short oligonucleotide arrays

Incyte platform vs Affymetrix

Similar results were obtained with cDNA arrays from Incyte Genomics (http://www.incyte.com) versus the Affymetrix platform in a study of the differences between normal peripheral blood mononuclear cells and large granular lymphocytic leukemia [57]. In this case, experiments were performed in the same laboratory and with the same material, which could explain the poor correlation in the study described previously. However, there were differences in quantitation; for example, the gene encoding perforin showed a 103.0-fold change in the Affymetrix array, and only 3.8-fold change in the cDNA array. Their northern blot analyses gave a value somewhere between the two extremes. Large discrepancies such as these, spread throughout many genes, could easily explain why there seems to be little correlation between the platforms when applying clustering analysis and other algorithms to these data. Another study using cDNA arrays and short oligonucleotide arrays appears to confirm this explanation [58]. The study measured gene expression in a human neuroblastoma cell line, and after treatment of the cells, measurements from Affymetrix chips and Incyte cDNA arrays showed an increase

in the mRNAs of 218 genes and 4 genes, respectively. The four genes were a subset of the 218 RNAs. Nine genes out of the 218 were confirmed as upregulated by RT–PCR. In this case, the authors conclude that short oligonucleotide arrays are more reliable for measuring gene expression changes compared with data from cDNA microarrays. In several studies, the short oligonucleotide format appears to have a higher dynamic range than the cDNA format, which could explain why more genes are flagged when studying changes in mRNA levels. However, this is not always the case, as other studies show that the cDNA platforms perform as well or better than the short oligonucleotide arrays [59,60].

Long oligonucleotide array studies

Owing to their relative novelty, there are only a few papers comparing data from long oligonucleotide arrays with other formats. Data from the validation of these types of arrays can be found on the websites of companies producing oligonucleotide libraries (ClonTech, Compugen, MWG Biotech, and Operon) or pre-printed slides (Agilent, http://www.agilent.com). Recently, the assessment of the sensitivity of 50mer arrays compared with cDNA was described [61]. Here, the 50mers and PCR products were spotted on the same slide and their performances evaluated. In this case, no difference in sensitivity between the two were found, and both could detect the equivalent of ~10 mRNA copies per cell. The method was highly specific as cross-hybridization required that the probe have >75% homology with the oligonucleotide. At this level of homology, the probe is probably a related gene and would also have hybridized to its cDNA counterpart. An elegant study comparing short and long oligonucleotide arrays as well as cDNAs was reported by the Rosetta Inpharmatics group (http://www.rosettabio.com) [26]. Oligonucleotides were synthesized *in situ* with a range of 20–60 bases. The study demonstrated that the 60mer length gave the best combination of sensitivity and specificity. The sensitivity was reported as close to one in one-million, or ~0.1 mRNA copies per cell, assuming 100,000 transcripts per cell. This level is as good or better than those reported for short oligonucleotide or cDNA arrays. Comparing the 60mer arrays with cDNA arrays also gave good results. Using an array of one oligonucleotide per gene, they found a close correlation with results from the cDNA array ($r = 0.97$). That one long oligonucleotide per gene was sufficient for expression studies was an important finding as it significantly simplifies both the *in situ* synthesis and robotic printing approaches to making arrays. This platform was also used to show that a particular gene expression profile can be used to predict the clinical outcome of breast cancer [27]. The group studied samples from primary breast tumors and determined a gene expression signature strongly predictive of poor prognosis, in addition to establishing an expression signature of *BRCA1* carriers. These data, along with other publications [28–37], indicate that use of long oligonucleotides is an excellent approach to the study of gene expression.

Discussion

The microarray field is now entering a phase of widespread use, similar to the development of PCR technologies in the late 1980s. The availability of complete genomic sequences from many different organisms, advances in microarray instrumentation, and large-scale commercial involvement, gives the researcher a wide choice of platforms from which to choose when conducting gene expression studies. As was the case early in the PCR era, the microarray field is now suffering from enthusiastic use, and growing pains. With several thousand publications already in existence, it has become apparent to many microarray aficionados that a large number of studies cannot be reproduced or replicated, and that related studies from different laboratories, or even in the same laboratory, are not comparable. Establishment of the MIAME guidelines [52–55] for publication will go far to alleviate the confusion in the reporting of microarray experiments. These guidelines include details of: (1) how the experiment was designed, (2) the design of the arrays or the name and location of spots on arrays, (3) sample name, extraction and labelling, (4) hybridization protocols, (5) methods for image measurements, and (6) the controls used. Obviously, good experimental design is the first important step, and details of how to properly design microarray experiments are described in recent reviews [50,62–64].

There are many sources of variation in microarray studies, and the type of platform used is a major source of this variability. At present, arrays can contain oligonucleotides of 25, 30, 40, 50, 60, 65, 70 to 80 bases in length, and cDNAs of hundreds-of-bases to several kilobases in length. Although some of these platforms are being used for other purposes, such as SNP analysis and diagnostics, it would be helpful if the platforms for gene expression analysis were standardized. The use of short 25-base oligonucleotide and cDNA methods are prevalent, partly because they were the first to be used extensively, and because researchers are usually hesitant to change if a large part of their research programs are based on one particular platform.

Advantages and disadvantages

Short oligonucleotide arrays

The short oligonucleotide approach has been commercially available for several years and has a strong manufacturing

base, resulting in reagents and chips that are readily available. In addition, there are a large number of published studies using this system, and it is a generally accepted method for expression analysis by many laboratories and journals. With complex eukaryotic genomes, there might be problems with cross-hybridization of the oligonucleotides to unrelated probes because of the short length of the target, but the use of several oligonucleotides per gene appears to help eliminate this problem. One main disadvantage with this approach is that changes are not readily accommodated, for example, when new sequence information becomes available or if different arrays are desired.

cDNA arrays

The cDNA approach has the advantage that no sequence information is necessary before setting up the arrays. The PCR products from cDNA banks can be synthesized using universal primers, and interesting genes can be sequenced after array analysis. The large size of the PCR product is also helpful in enabling stringent hybridization conditions and lowering cross-hybridization of unrelated genes, although closely related gene families will still be able to anneal to some extent. Although PCR production of DNAs for microarrays is not difficult *per se*, the large number of DNAs required for complete coverage of a complex genome is taxing for most laboratories. To produce, analyze, purify, aliquot and keep track of 20,000–40,000 different PCR products is not easy, and is a difficult feat even for commercial sources dedicated to such a task.

Long oligonucleotide arrays

The completion of numerous genomic sequences and increased efficiency of production has combined to make the use long oligonucleotide (50–80 bases) for gene expression studies a highly attractive alternative to short oligonucleotide or cDNA arrays. There are several advantages to this platform over the others. Their longer length enables more specificity in hybridization than shorter lengths, and they can therefore be treated similarly to cDNA arrays. Every oligonucleotide is of the same length and has almost the same melting temperature and G–C content, enabling more consistent hybridization conditions for every gene on the array. Compared with cDNA arrays, it is much easier to deposit the same concentration of DNA per spot with oligonucleotides. The oligonucleotides are single stranded and do not require a denaturation step as with the cDNA format; this also eliminates the problem of re-naturation, which can decrease hybridization efficiency. The molar concentration of oligonucleotides in arrays of long oligonucleotides is much higher than the short oligonucleotide or cDNA versions, with an ~100-fold difference in concentration. This enables more probe molecules to be annealed, and more than makes up for the shorter length relative to the cDNA arrays. Through appropriate design, long oligonucleotides can distinguish between alternatively spliced mRNAs [65,66] – something that is not possible with cDNA arrays. Finally, additions, changes or modifications in the arrays are more straightforward compared with photolithographic methods or cDNA arrays. This is a major advantage to the researcher when new sequence information becomes available or when experimental results require more focused arrays.

Concluding remarks

The microarray field is in great need of standardization to take full advantage of the tremendous wealth of genomic data. Reducing the number of microarray platforms would be a good first step, and the long oligonucleotide approach is a good candidate to replace cDNA arrays, leaving the researcher with two basic array methods for the study of gene expression. Although there are other, non-array methods for analyzing gene expression, such as SAGE [67–70], the simplicity of the oligonucleotide approach makes it the most attractive option for the general research community.

The 'final' human genome sequence is scheduled to be released in the spring of 2003, in time for the 50th anniversary of the publication of the structure of DNA. For those engaged in drug development, this means, in essence, that every possible drug target encoded in the genome will be available for testing. Therapeutic drug discovery will no longer be hampered by a shortage of targets but, rather, hindered by an excess of targets. Microarrays will play an essential role in overcoming this obstacle in both target identification and in the long road of drug discovery and development [71–76]. Standardized oligonucleotide array platforms will be crucial in helping the pharmaceutical industry to take full advantage of this genomic treasure chest.

References

1 Venter, J.C. *et al.* (2001) The sequence of the human genome. *Science* 291, 1303–1351
2 Lander, E.S. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
3 The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018
4 Yu, J. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Indica*). *Science* 296, 79–92
5 Goff, S.A. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Japonica*). *Science* 296, 92–100
6 Schena, M. *et al.* (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470
7 Ekins, R. and Chu, F.W. (1999) Microarrays: their origins and applications. *Trends Biotechnol.* 17, 217–218

8 Ekins, R. *et al.* (1989) High specific activity chemiluminescent and fluorescent markers: their potential application to high sensitivity and 'multi-analyte' immunoassays. *J. Biolumin. Chemilumin.* 4, 59–78
9 Ekins, R.P. (1989) Multi-analyte immunoassay. *J. Pharm. Biomed. Anal.* 7, 155–168
10 Hayward, R.E. *et al.* (2001) Probing the proteome – protein arrays and their applications. *Drug Discov. Today* 6, 1263–1264
11 Mitchell, P. (2002) A perspective on protein microarrays. *Nat. Biotechnol.* 20, 225–229
12 Templin, M.F. *et al.* (2002) Protein microarray technology. *Trends Biotechnol.* 20, 160–166
13 Petricoin, E.F. *et al.* (2002) Clinical proteomics: translating benchside promise into bedside reality. *Nat. Rev. Drug Discov.* 1, 683–695
14 MacBeath, G. (2002) Protein arrays and proteomics. *Nat. Genet.* 32 (Suppl.), 526–532
15 Schena, M. *et al.* (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10614–10619
16 Duggan, D.J. *et al.* (1999) Expression profiling using cDNA microarrays. *Nat. Genet.* 21 (Suppl), 10–14
17 Winzeler, E.A. *et al.* (1999) Fluorescence-based expression monitoring using microarrays. *Methods Enzymol.* 306, 3–18
18 Lou, X.J. *et al.* (2001) Expression monitoring using cDNA microarrays. *Methods Mol. Biol.* 175, 323–340
19 Fodor, S.P.A. *et al.* (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science* 251, 767–773
20 Lockhart, D.J. *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14, 1675–1680
21 Lipshutz, R.J. *et al.* (1999) High density synthetic oligonucleotide arrays. *Nat. Genet.* 21 (Suppl), 20–24
22 Singh-Gasson, S. *et al.* (1999) Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat. Biotechnol.* 17, 974–978
23 Cerrina, F. *et al.* (2002) Biological lithography: development of a maskless microarray synthesizer for DNA chips. *Microelectron. Eng.* 61-62, 33–40
24 Nuwaysir, E.F. *et al.* (2002) Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Res.* 12, 1749–1755
25 Blanchard, A.P. *et al.* (1996) High-density oligonucleotide arrays. *Biosens. Bioelectron.* 11, 687–690
26 Hughes, T.R. *et al.* (2001) Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* 19, 342–347
27 van Veer, A. *et al.* (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530–536
28 Chambers, J. *et al.* (1999) DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. *J. Virol.* 73, 5757–5766
29 Beaucage, S.L. (2001) Strategies in the preparation of DNA oligonucleotide arrays for diagnostic applications. *Curr. Med. Chem.* 8, 1213–1244
30 Dudley, A.M. *et al.* (2002) Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7554–7559
31 Relogio, A. *et al.* (2002) Optimization of oligonucleotide-based DNA microarrays. *Nucleic Acids Res.* 30, e51
32 El Atifi, M. *et al.* (2002) Long oligonucleotide arrays on nylon for large-scale gene expression analysis. *BioTechniques* 33, 612–618
33 Wang, R-F. *et al.* (2002) Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples. *FEMS Microbiol. Lett.* 213, 175–182
34 Wagner, E.K. *et al.* (2002) Practical approaches to long oligonucleotide-based DNA microarray: lessons from Herpesviruses. *Prog. Nucleic Acid Res. Mol. Biol.* 71, 445–491
35 Stingley, S.W. *et al.* (2000) Global analysis of herpes simplex virus type I transcription using an oligonucleotide-based DNA microarray. *J. Virol.* 74, 9916–9927
36 Yang, W.C. *et al.* (2002) General and specific alterations in programming of global viral gene expression during infection by VP16 activation-deficient mutants of herpes simplex virus type I. *J. Virol.* 76, 12758–12774
37 Wang, D. *et al.* (2002) Microarray-based detection and genotyping of viral pathogens. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15687–15692
38 Li, F. and Stormo, G.D. (2001) Selection of optimal DNA oligos for gene expression arrays. *Bioinformatics* 17, 1067–1076
39 Rouillard, J-M. *et al.* (2002) OligoArray: genome-scale oligonucleotide design for microarrays. *Bioinformatics* 18, 486–487
40 Chen, H. and Sharp, B.M. (2002) Oliz, a suite of Perl scripts that assist in the design of microarrays using 50mer oligonucleotides from the 3′ untranslated region. *BMC Bioinformatics* 3, 27
41 Wright, M.A. and Church, G.M. (2002) An open-source oligomicroarray standard for human and mouse. *Nat. Biotechnol.* 20, 1082–1083
42 Taniguchi, M. *et al.* (2001) Quantitative assessment of DNA microarrays – comparison with Northern blot analyses. *Genomics* 71, 34–39
43 O'Connell, J. (2002) RT-PCR in biomedicine. Opportunities arising from the new accessibility of mRNA. *Methods Mol. Biol.* 193, 3–17
44 Bustin, S.A. (2000) Absolute quantitation of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* 25, 169–193
45 Aach, J. *et al.* (2000) Systematic management and analysis of yeast gene expression data. *Genome Res.* 10, 431–445
46 Quackenbush, J. (2001) Computational analysis of microarray data. *Nat. Rev. Genet.* 2, 418–427
47 Nadon, R. and Shoemaker, J. (2002) Statistical issue with microarrays: processing and analysis. *Trends Genet.* 18, 265–271
48 Brazma, A. *et al.* (2002) Microarray data representation, annotation and storage. *Adv. Biochem. Eng. Biotechnol.* 77, 113–139
49 Holloway, A.J. *et al.* (2002) Options available – from start to finish – for obtaining data from DNA microarrays II. *Nat. Genet.* 32 (Suppl.), 481–489
50 Churchill, G.A. (2002) Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* 32 (Suppl.), 490–495
51 Quackenbush, J. (2002) Microarray data normalization and transformation. *Nat. Genet.* 32 (Suppl.), 496–501
52 Brazma, A. *et al.* (2001) Minimum information about a microarray experiment (MIAME) – towards standards for microarray data. *Nat. Genet.* 29, 365–371
53 Stoeckert, C.J., Jr *et al.* (2002) Microarray databases: standards and ontologies. *Nat. Genet.* 32 (Suppl.), 469–473
54 No author (2002) Microarray standards at last. *Nature* 419, 323
55 Ball, C. *et al.* (2002) Standards for microarray data. *Science* 298, 539
56 Kuo, W.P. *et al.* (2002) Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* 18, 405–412
57 Kothapalli, R. *et al.* (2002) Microarray results: how accurate are they? *BMC Bioinformatics* 3, 22
58 Li, J. *et al.* (2002) Differential gene expression patterns revealed by oligonucleotide versus long cDNA arrays. *Toxicol. Sci.* 69, 383–390
59 Yuen, T. *et al.* (2002) Accuracy and calibration of commercial oligonucleotide and custom cDNA arrays. *Nucleic Acid Res.* 30, e48
60 Rhodes, D.R. *et al.* (2002) Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway disregulation in prostate cancer. *Cancer Res.* 62, 4427–4433
61 Kane, M.D. *et al.* (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acid Res.* 28, 4552–4557
62 Kerr, M.K. and Churchill, G.A. (2001) Experimental design for gene expression microarrays. *Biostatistics* 2, 183–201
63 Yang, Y.H. and Speed, T. (2002) Design issues for cDNA microarray experiments. *Nat. Rev. Genet.* 3, 579–588
64 Simon, R. *et al.* (2002) Design studies using DNA microarrays. *Genet. Epidemiol.* 23, 21–36

65 Clark, T.A. *et al.* (2002) Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* 296, 907–910
66 Maniatis, T. and Tasic, B. (2002) Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* 418, 236–243
67 Liang, P. (2002) SAGE Genie: a suite with panoramic view of gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11547–11548
68 Evans, S.J. *et al.* (2002) Evaluation of Affymetrix gene chip sensitivity in rat hippocampal tissue using SAGE analysis. *Eur. J. Neurosci.* 16, 409–413
69 Patino, W.D. *et al.* (2002) Serial analysis of gene expression: technical considerations and applications to cardiovascular biology. *Circ. Res.* 91, 565–569
70 Soulet, D. and Rivest, S. (2002) Perspective: how to make microarray, serial analysis of gene expression, and proteomic relevant to day-to-day endocrine problems and physiological systems. *Endocrinololgy* 143, 1995–2001
71 Somogyi, R. and Greller, L.D. (2001) The dynamics of molecular networks: applications to therapeutic discovery. *Drug Discov. Today* 6, 1267–1277
72 Clarke, P.A. *et al.* (2001) Gene expression microarray analysis in cancer biology, pharmacology, and drug development: progress and potential. *Biochem. Pharmacol.* 62, 1311–1336
73 Ulrich, R. and Friend, S.H. (2002) Toxigenomics and drug discovery: will new technologies help us produce better drugs? *Nat. Rev. Drug Discov.* 1, 84–88
74 Hamadeh, H.K. *et al.* (2002) Gene expression analysis reveals chemical-specific profiles. *Toxicol. Sci.* 67, 219–231
75 Petricoin, E.F., III *et al.* (2002) Medical applications of microarray technologies: a regulatory science perspective. *Nat. Genet.* 32 (Suppl.), 474–479
76 Gerhold, D.L. *et al.* (2002) Better therapeutics through microarrays. *Nat. Genet.* 32 (Suppl.), 547–552

Drug Discovery Today will protect your identity!

Here is an unrivalled opportunity to get off your chest those things that really irritate you!

…and to be able to tell the relevant people what really irritates you without them knowing it is you!

Send in your letters and get some real debate going!

Please send all contributions to Dr Joanna Owens
e-mail: joanna.owens@elsevier.com

Publication of letters is subject to editorial discretion

Contributions to *Monitor*

We welcome recommendations of papers for review within *Monitor*, in the fields of combinatorial chemistry, pharmacogenomics, pharmacoproteomics, bioinformatics, new therapeutic targets, high throughput screening, new drug delivery technologies and other promising lines of research.

Details of recent papers or those *in press* should be directed to Dr Steve Carney, Editor, *Drug Discovery Today*, Elsevier Science London, 84 Theobald's Road, London, UK WC1X 8RR. tel: +44 207 611 4132, fax: +44 207 611 4485, e-mail: s.carney@elsevier.com

Contributions to *Profiles*

We welcome contributions for the *Profiles* series, which gives a commentary on promising lines of research, new technologies and progress in therapeutic areas. Articles should provide an accurate summary of the essential facts together with an expert commentary to provide a perspective. Brief outlines of proposed articles should be directed to the *Monitor* Editor (see below). Articles for publication in *Monitor* are subject to peer review and occasionally may be rejected or, as is more often the case, authors may be asked to revise their contribution. The *Monitor* Editor also reserves the right to edit articles after acceptance.

All suggestions or queries relating to *Monitor* should be addressed to Dr Steve Carney, Editor, *Drug Discovery Today*, Elsevier Science London, 84 Theobald's Road, London, UK WC1X 8RR. tel: +44 207 611 4132, fax: +44 207 611 4485, e-mail: s.carney@elsevier.com