reviews

Oligomer-chip technology

Jörg D. Hoheisel

Oligonucleotide microarrays have caught the attention of a wide scientific audience because of a collection of recent publications that demonstrate the technique's huge potential in a development that seems to be having similar consequences for molecular genetics as the arrival of the semiconductor had for electronics. From today's still rather confined basis, the methodology is bound to expand and diversify, radically transforming not only molecular genetics but also numerous other fields of application in medicine and pharmaceutical research.

The main emphasis of current genome analysis is still deciphering the basic sequence information of both the coding and noncoding regions of genomic DNA. However, technological development has already moved on to analytical procedures for the cellular effects and functional consequences of the DNA-encoded information. In recent years, DNA-chip technology has emerged as a prime candidate for the performance of many such analyses in molecular biology and medicine, particularly given the stimulus provided by the rapid increase in the amount of genetic information needing interpretation. This is because one of the main features of this technique is its high degree of parallelism, and the capacity to assay a large number of samples in a short time will make it almost indispensable. Several different methods for the creation of DNA arrays are being used, but the basic principle remains the same: hybridisation of an unknown sample to an ordered array of immobilised DNA molecules of known sequence produces a specific hybridisation pattern that can be analysed or compared to a given standard (Fig. 1). The array elements consist of either synthetic oligonucleotides or longer DNA fragments, usually PCR products or isolated cloned DNA.

The basic principle of using oligonucleotide arrays was first proposed in the late 1980s, when four groups independently developed the concept of determining a DNA sequence by hybridisation to a comprehensive set of oligonucleotides, such as all the possible 65 536 octamer sequences (sequencing by hybridisation, SBH)[1–4]. As indicated by the designation of the early scientific gatherings in this field[5], the initial aim was the establishment of SBH, in preference to any other application. Only with time did the true potential of oligomer arrays become clearer and the technical barriers to the (experimentally complex) process of SBH more obvious, shifting the emphasis to more immediately addressable applications.

J. D. Hoheisel (j.hoheisel@dkfz-heidelberg.de) is at Molecular-Genetic Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, D-69120 Heidelberg, Germany.

Basic aspects of oligonucleotide microarrays

Array production

Fabrication of the arrays is quite obviously of crucial importance to oligonucleotide-chip technology. Initial *in situ* synthesis techniques[7] would have been sufficient for the generation of comprehensive arrays, but for the combinatorial limitations of the channel system, which failed to synthesise entirely independent sequences at the individual grid positions. With the advent of a procedure that allowed the control of oligomer synthesis on the chip by photolithography[8,9], this shortcoming was resolved; in addition, a new dimension in the miniaturisation and complexity of the arrays unfolded.

Alternative production methods do exist, and these can produce features unobtainable by photolithography. Transferring individual DNAs onto a solid support[10], for example, seems to be a very tedious process, but it does permit a variation in the amount of material placed at each grid position; with modern high-throughput robotics, large-scale chip production by this method is well within our reach. Individual-transfer arrays possess an advantage over photolithographic arrays in their ability to be constructed with varying duplex stabilities of the oligonucleotides, merely by varying their concentration on the chip surface[11]; also, standard synthesis chemistry is well established for a number of nucleotide derivatives for which no light-inducible monomer equivalents are available. The critical drawback, however, is still the need for the external synthesis and storage of very many different oligonucleotides prior to actual chip production. Control of chip production by electrochemical patterning of surfaces using electrode arrays seems to be a viable alternative to photolithography in terms of size, speed and ease of synthesis[12]; whether it produces arrays of similar or even better quality remains to be seen.

 0167 – 7799/97/$17.00. PII: S0167-7799(97)01123-2

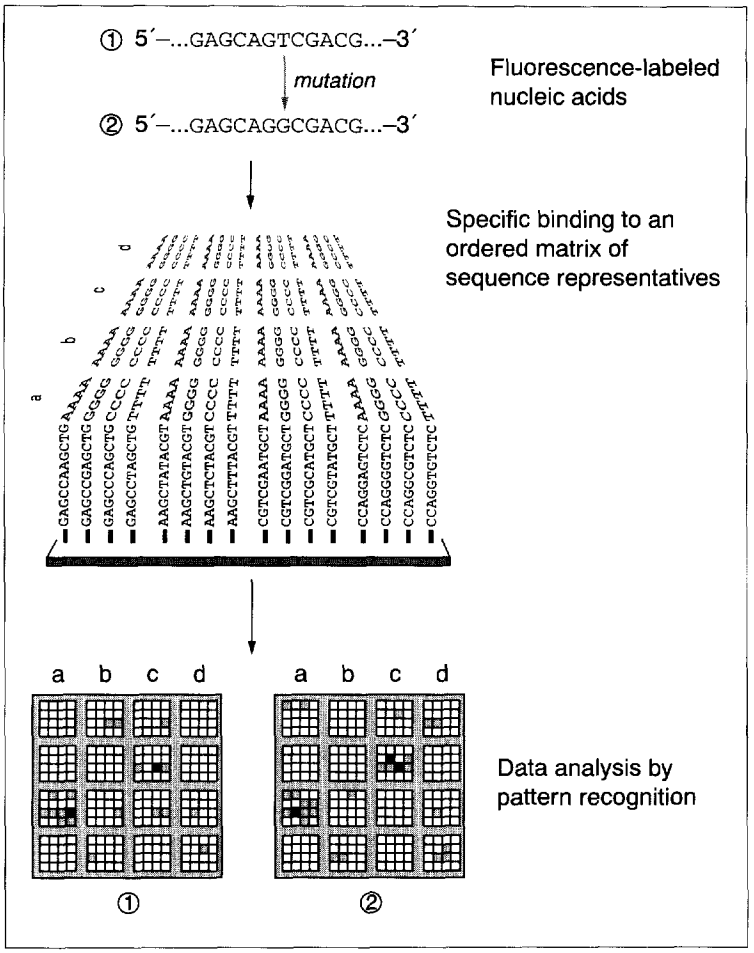

Figure 1

Basic principle of oligonucleotide-chip technology. Nucleic acids are exposed to an ordered grid of array-bound oligomer molecules. At each position, a distinct sequence is attached to the support medium, with the differences as small as a single base for the detection of polymorphisms, as displayed here. After hybridisation of the probe to chip molecules with a complementary sequence, analysis of the binding position is performed by pattern recognition. Evaluation of signal intensities is essential for the identification of the occurrence of mismatch hybridisation (modified from Ref. 6).

Even the current champion, photolithographic chip synthesis, has flaws, however, including relatively low yield obtained on the solid support[13]. This is at least one reason why a large degree of redundancy is required in actual analyses; although not necessarily an obstacle for some applications, this remains a problem for quantitative measurements. Work is ongoing to improve the efficiency of synthesis towards that achieved on standard support media[14,15].

Hybridisation

After production of the oligomer chip, the successful performance of target hybridisation is the most important issue of chip-based analyses. Many effects are still not completely understood, including the influence of the dangling ends of a bound molecule on the stability of the duplex[16] or, even more importantly, how much of a probe is really accessible to hybridisation; the relatively small amount of published data indicates that the latter issue, especially, is not yet resolved[17]. It is, however, possible that there is more data on this subject in the private domain that will not be published in the foreseeable future because it could give its owner a competitive edge. Fragmentation of the probe would be one way around the problem of intrastrand secondary and tertiary structure, which prevents binding to the arrayed oligomers. However, this would simultaneously cause another problem concerned with the labeling: usually, a base-specific fluorophor label is incorporated in the probe, but after fragmentation, there could be different amounts of label contained in the various, very-short pieces generated, due to, for example, enzymatic preferences. This could lead to a bias when comparative analyses are performed.

Probe concentration is another issue that strongly affects hybridisation; in fact, it is the one major reason why chips are needed for this sort of analysis. Because the amount of sample is limited in many potential applications, the relevant assay must be performed in small volumes, and thus on small surfaces, in order to obtain the probe concentrations necessary. A possibly very useful technique in this context is directing the probe flow by applying electric fields[18]. In addition, this technology allows for electronic control of the hybridisation stringency, thus accelerating the process.

Detection

Currently, there is a sort of monopoly in detection methodology. The use of a fluorescently labeled probe requires that detection is performed optically, but there are, nevertheless, different approaches to this that could prove significant. Rather than using a confocal system[19], thus blocking out all signal from outside a certain layer, excitation could be done by taking advantage of optical wave guides, for example, which cause excitation to take place only in the zone reached by the evanescent wave along the wave-guide material[20]. Thus, only label that is very close to the support, almost exclusively bound probe molecules, will produce a signal. However, other techniques for circumventing the complex and costly confocal system are possible. The most elegant would be direct detection of probe binding by electronic means. There is research being done in this direction[21], but the published results are, to date, not very convincing. Nevertheless, this procedure would really merge DNA-chip technology with electronics, and make a pocket-sized machine possible.

Applications

The range of applications for this technique is already wide and will probably extend still further. However, I should mention that, for some assays, the use of longer array molecules could be advantageous; one of these applications is transcriptional profiling (Fig. 2), a technique pioneered in the early 1990s on filters[22–24]. Total mRNA is isolated and hybridised to arrays that contain gene representatives, either as PCR products[25,26] or oligonucleotides[27]. Miniaturisation has led to the technique now being fit to deal with entire genomes at high levels of sensitivity.

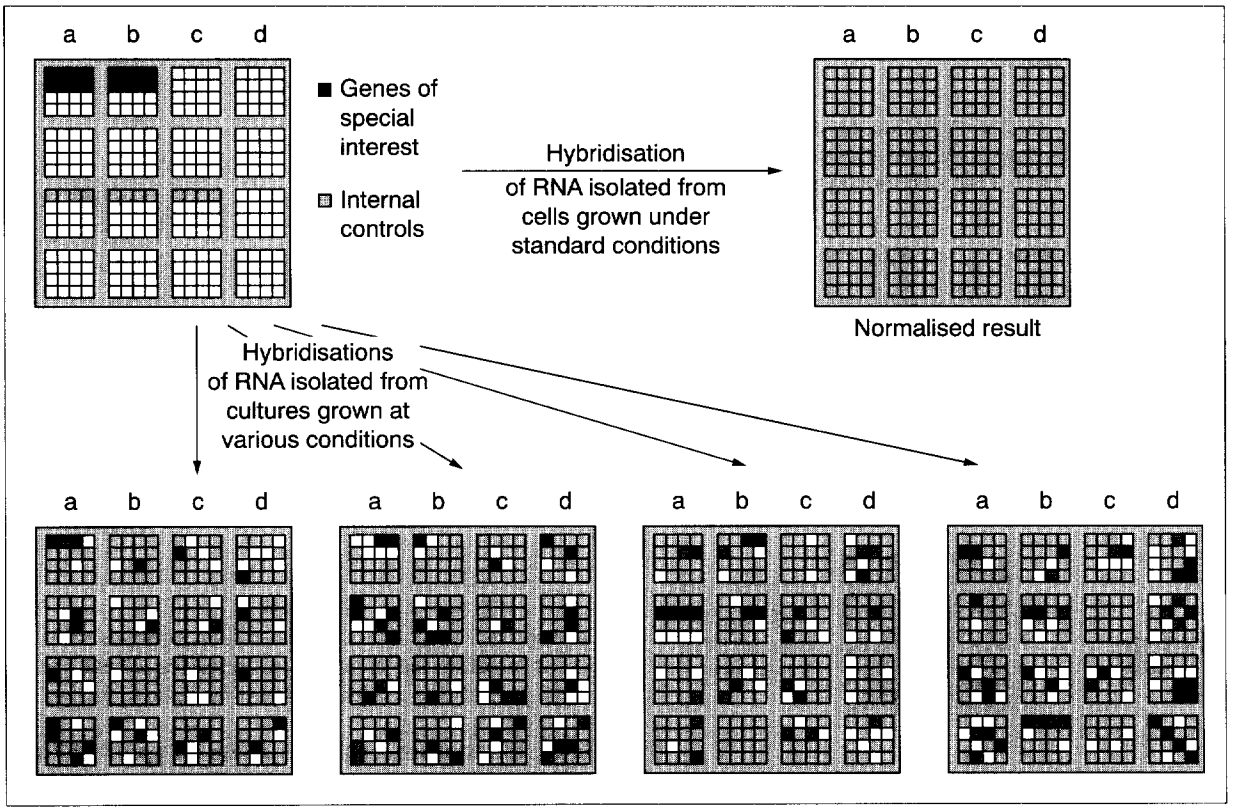

Figure 2

Parallel analysis of gene transcription. Total mRNA is isolated from cells and hybridised to a grid of molecules that represents (all) the genes of an organism. By analysing the binding positions and signal intensities, the transcription rate of each individual gene can be determined and compared with similar data generated with RNA from cells grown under different conditions. Using two dyes, two preparations could be mixed prior to hybridisation for immediate comparison.

Also currently at the forefront is the identification of mutations and polymorphisms[28–30]. By hybridising a sequence to a set of oligonucleotides containing all possible sequence deviations (Fig. 1), such analyses could be performed very quickly and also eventually lead to another sort of genetic map. While the physical mapping of individual cosmid or bacterial-artificial-chromosome (BAC) clones[31] seems an unlikely application, because it loses out on the issue of parallelism (and thus efficiency) compared with other hybridisation-based techniques[32], determination of the sequence-tagged-site (STS) content of larger genomic regions, such as radiation hybrids, could be worthwhile. Once one-tube PCR amplifications with very many pooled primer pairs can be performed, the presence of STSs in a given region could be checked by applying the result of such a PCR reaction to a chip that contains oligomer representatives for all the individual tag sequences (D. Cox, pers. commun.).

Other applications of oligonucleotide arrays range from studies in cell biology to chemical matters and include quantitative phenotypic analysis in a population of different (yeast) cells[33] (Fig. 3) and even highly parallel, and potentially extremely cheap, synthesis of primer molecules for subsequent use in molecular biology[14].

The most ambitious aim is still SBH, and a chip with 65 536 octamers, for instance, would also be the ultimate tool for many of the analyses listed above. The technique will probably never be used for *de novo* sequencing, however, because the read length is too limited in this sort of analysis; certain octamer sequences will occur more than once, thus creating many branching points within a given DNA fragment. However, the main sequencing task will only start after the human and other genomes have been completely sequenced. Then, individual patients, or distinct descendants of bacteria or other organisms, will be analysed in comparison with the sequence 'standard' that is in the database; for this application, the read-length of SBH is more than sufficient[34]. The last major technical obstacle was the need for equalising the duplex stabilities of all the different sequences. As there are ways to solve or get around this complication[35,36], no basic problem remains that could prevent SBH from becoming more than just a clever idea.

Peptide nucleic acids

More recently, a new class of derivatives, the peptide nucleic acids (PNAs; Fig. 4), have been developed[37]. Exhibiting many features that are identical, or at least very similar, to the behaviour of DNA oligonucleotides, such as the formation of a duplex with nucleic acids of complementary sequence, their application to oligomer-chip sensors might be advantageous for several reasons: their duplex stability greatly

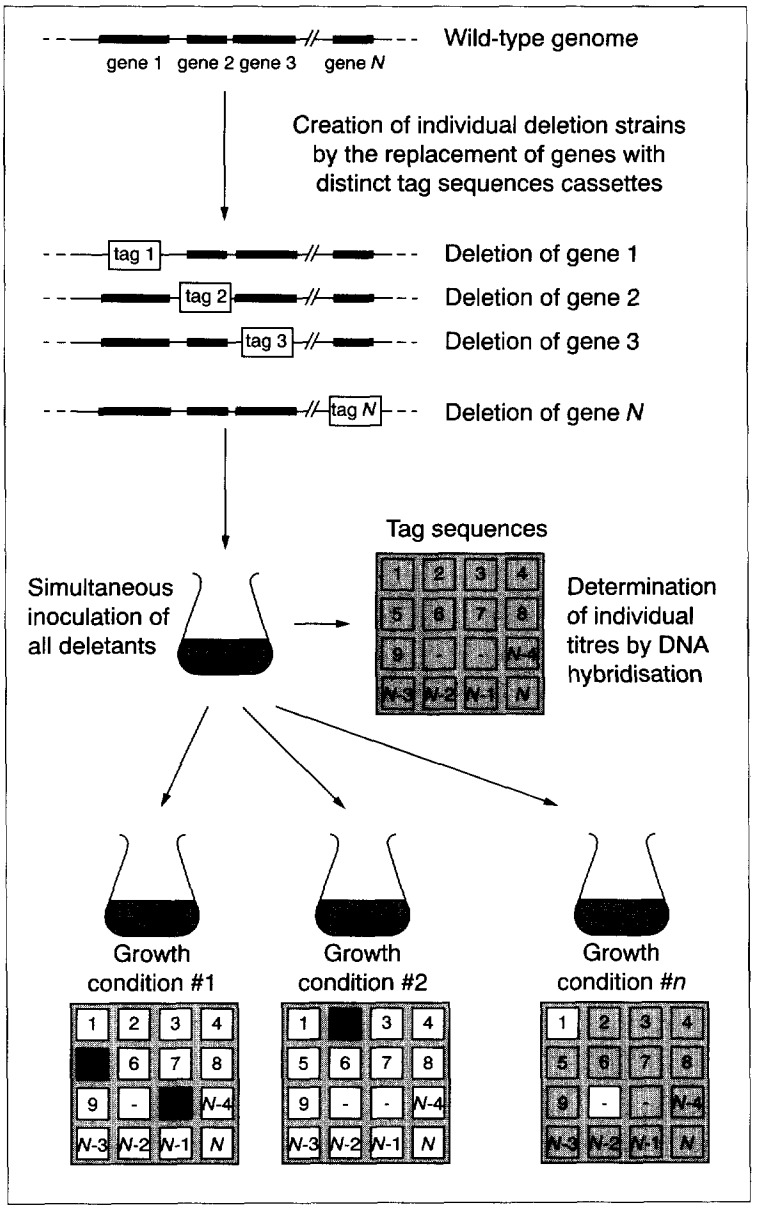

Figure 3

Phenotypic analysis in a population of different yeast cells. For functional analyses of gene activity, individual genes are replaced by DNA cassettes, each containing a distinct sequence of 20 to 40 bp. The deletants are all grown in one culture; at carefully designed growth conditions, individual deletions will be at an advantage or disadvantage, and this will lead to overgrowth or disappearance of the particular deletant strains. This is detected by hybridising DNA from the population of all cultured cells to a chip containing the tag sequences at specific locations.

exceeds that of equivalent DNA oligonucleotides; their binding to a complementary DNA strand is even more specific[38]; hybridisation takes place in a low-salt to possibly even no-salt buffer, and, because interstrand repulsion is not counteracted by ions, intramolecular double-strand formation of the DNA or RNA probe is disfavoured, making the entire probe sequence more accessible to the oligomers on the chip; and, less scientific but nonetheless important, the patent situation seems to be clear cut. Analyses with relatively small arrays of PNA oligomers has demonstrated the potential of these compounds as an alternative substrate for chip technology[39].

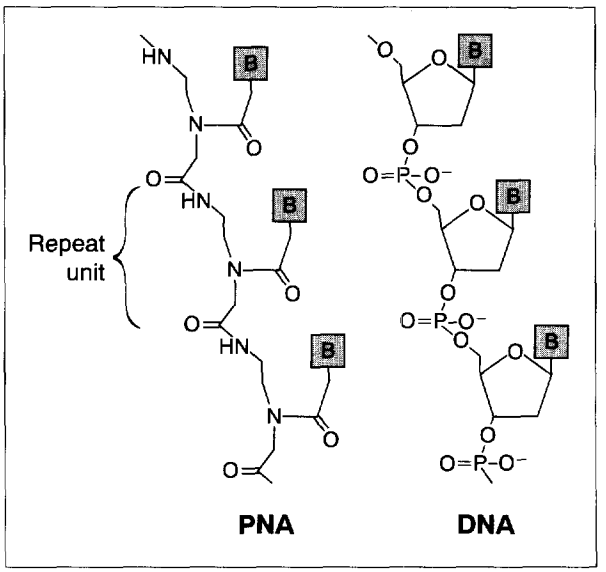

Figure 4

Comparison of the structures of PNA and DNA molecules. PNA consists of repeating units of *N*-(2-aminoethyl)-glycine linked by amino bonds. Bases (B) are attached by methylene–carbonyl linkages.

Scientific, commercial and medical applicability

There is no question that oligomer-chip technology will be put to productive and profitable use in science and many commercial applications, such as screening the molecular responses caused by the large numbers of compounds generated by pharmaceutical companies. Medical diagnosis in individual human beings is a different story, however, requiring the various analytical techniques to reach a standard sufficient to perform vital diagnoses reliably on individual patients. In practical terms of development and commercialisation, I wonder how regulative bodies such as the US Food and Drug Administration will set their accuracy requirements, given their different perspective. Considering the specificity required in practical uses of much simpler analytical tools such as primer oligonucleotides for PCR-based techniques, this is a crucial question, and a delicate one once it comes to commercial issues. Still, with the advances that have already been made being so enormous, the prospects for further improvement so substantial, and in view of related developments towards microfabricated structures for DNA analysis[40], I am convinced that a time will come when patients will give a drop of blood from their fingertip at a doctor's surgery, and shortly thereafter will be given a diagnosis based on chip-hybridisation data.

Acknowledgments

Work in the author's laboratory was funded by the Deutsche Forschungsgemeinschaft, the German Bundesforschungsministerium (BMBF) and the European Commission.

References

1 Khrapko, K., Lysov, Y., Khorlyn, A., Shick, V., Florentiev, V. and Mirzabekov, A. (1989) *FEBS Lett.* 256, 118–122

2 Drmanac, R., Labat. I., Brukner, I. and Crkvenjakov, R. (1989) *Genomics* 4, 114–128
3 Bains, W. and Smith, G. (1988) *J. Theor. Biol.* 135, 303–307
4 Southern, E. M., Maskos, U. and Elder, J. K. (1992) *Genomics* 13, 1008–1017
5 Cantor, C. R., Mirzabekov, A. and Southern, E. (1992) *Genomics* 13, 1378–1383
6 Hoheisel, J. D., Maier, E., Meier-Ewert, S. and Lehrach, H. (1993) *Ann. Biol. Clin.* 50, 827–829
7 Maskos, U. and Southern, E. M. (1992) *Nucleic Acids Res.* 20, 1675–1679
8 Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. and Solas, D. (1991) *Science* 251, 767–773
9 Pease, A. C., Solas, D., Sullivan, E. J., Cronin, M. T., Holmes, C. P. and Fodor, S. P. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 5022–5026
10 Yershov, G. *et al.* (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 4913–4918
11 Lysov, Y. P., Chernyi, A. A., Balaeff, A. A., Beattie, K. L. and Mirzabekov, A. D. (1994) *J. Biomol. Struct. Dyn.* 11, 797–812
12 Southern, E. M. (1996) International Patent PTO WO 93/22480
13 Pirrung, M. C. and Bradley, J-C. (1995) *J. Org. Chem.* 60, 6270–6276
14 Weiler, J. and Hoheisel, J. D. (1996) *Anal. Biochem.* 243, 218–227
15 Matson, R. S., Rampal, J., Pentoney, S. L., Anderson, P. D. and Coassin, P. (1995) *Anal. Biochem.* 224, 110–116
16 Williams, J. C., Case-Green, S. C., Mir, K. U. and Southern, E. M. (1994) *Nucleic Acids Res.* 22, 1365–1367
17 Southern, E. M. *et al.* (1994) *Nucleic Acids Res.* 22, 1368–1373
18 Sosnowski, R. G., Tu, E., Butler, W. F., O'Connell, J. P. and Heller, M. J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 1119–1123
19 Lipshutz, R. J. *et al.* (1995) *Biotechniques* 19, 442–447
20 Stimpson, D. I. *et al.* (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 6379–6383
21 Hintsche, R., Paeschke, M., Uhlig, A. and Seitz, R. (1997) in *Frontiers in Biosensors: Fundamental Aspects* (Scheller, F. W., Schaubert, F. and Fedrowitz, J., eds), pp. 267–283, Birkhäuser Verlag
22 Augenlicht, L. H., Taylor, J., Anderson, L. and Lipkin, M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 3286–3289
23 Höög, C. (1991) *Nucleic Acids Res.* 19, 6123–6127
24 Zhang, Y., Coyne, M. Y., Will, S. G., Levenson, C. H. and Kawasaki, E. S. (1991) *Nucleic Acids Res.* 19, 3929–3933
25 Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995) *Science* 270, 467–470
26 DeRisi, J. *et al.* (1996) *Nat. Genet.* 14, 457–460
27 Lockhart, D. L. *et al.* (1996) *Nat. Biotechnol.* 14, 1675–1680
28 Maskos, U. and Southern, E. M. (1993) *Nucleic Acids Res.* 21, 2267–2268
29 Kozal, M. *et al.* (1996) *Nat. Med.* 2, 753–759
30 Chee, M. *et al.* (1996) *Science* 274, 610–614
31 Sapolsky, R. and Lipshutz, R. J. (1996) *Genomics* 33, 445–456
32 Hoheisel, J. D. *et al.* (1993) *Cell* 73, 109–120
33 Shoemaker, D. D., Lashkari, D. A., Morris, D., Mittmann, M. and Davies, R. W. (1996) *Nat. Genet.* 14, 450–456
34 Bains, W. (1993) *Genet. Anal. Tech. Appl.* 10, 84–94
35 Hoheisel, J. D. (1996) *Nucleic Acids Res.* 24, 430–432
36 Parinov, S. *et al.* (1996) *Nucleic Acids Res.* 24, 2998–3004
37 Nielsen, P. E., Egholm, M., Berg, R. H. and Burchardt, O. (1991) *Science* 254, 1497–1500
38 Nielsen, P. E. (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 167–183
39 Weiler, J., Gausepohl, H., Hauser, N., Jensen, O. N. and Hoheisel, J. D. (1997) *Nucleic Acids Res.* 25, 2792–2799
40 Burns, M. A. *et al.* (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 5556–5561