Acta Biochimica Polonica

Vol. 48 No. 3/2001

615–622

QUARTERLY

Minireview

An introduction to DNA chips: principles, technology, applications and analysis✪

Magdalena Gabig [1, 2] ✉ and Grzegorz Węgrzyn[3]

[1]*Department of Biotechnology, Royal Institute of Technology, Stockholm, Sweden;* [2]*Laboratory of Molecular Biology (affiliated with the University of Gdańsk), Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Gdańsk, Poland;* [3]*Department of Molecular Biology, University of Gdańsk, Gdańsk, Poland*

Received: 3 April, 2001; revised: 4 September, 2001; accepted: 5 September, 2001

Key words: oligonucleotide array, DNA array, gene expression, hybridisation

This review describes the recently developed GeneChip technology that provides efficient access to genetic information using miniaturised, high-density arrays of DNA or oligonucleotide probes. Such microarrays are powerful tools to study the molecular basis of interactions on a scale that would be impossible using conventional analysis. The recent development of the microarray technology has greatly accelerated the investigation of gene regulation. Arrays are mostly used to identify which genes are turned on or off in a cell or tissue, and also to evaluate the extent of a gene's expression under various conditions. Indeed, this technology has been successfully applied to investigate simultaneous expression of many thousands of genes and to the detection of mutations or polymorphisms, as well as for their mapping and sequencing.

MICROARRAY STRUCTURE AND USE

There are different names for the microarrays, like DNA/RNA Chips, BioChips or GeneChips. The array can be defined as an ordered collection of microspots, each spot containing a single defined species of a nucleic acid. The microarray technique is based on hybridisation of nucleic acids. In this technique, sequence complementarity leads to the hy-

✪This work was supported by the European Commission (contract No. QLK3-CT-1999-00533) and the State Committee for Scientific Research (KBN, Poland) grant No. 127/E-335/SPUB-M-5PR-UE/DZ 177/2000). G.W. acknowledges a financial support from the Foundation for Polish Science (subsidium 14/2000).

✉To whom correspondence should be addressed: Department of Biotechnology, Royal Institute of Technology (KTH), S-10691 Stockholm, Sweden; tel: (46 8) 5537 8310; fax: (46 8) 5537 8323; e-mail: magda.gabig@biotech.kth.se

bridisation between two single-stranded nucleic acid molecules, one of which is immobilised on a matrix [1].

There exist two variants of the chips: cDNA microarrays and oligonucleotide arrays [2–4]. Although both the DNA and oligonucleotide chips can be used to analyse patterns of gene expression, fundamental differences exist between these methods [5]. Two commonly used types of chips differ in the size of the arrayed nucleic acids. In cDNA microarrays, relatively long DNA molecules are immobilised by high-speed robots on a solid surface such as membranes, glass or silicon chips [2]. Sample DNAs are amplified by the polymerase chain reaction (PCR) and usually are longer than 100 nt. This type of arrays is used mostly for large-scale screening and expression studies. The oligonucleotide arrays are fabricated either by *in situ* light-directed chemical synthesis or by conventional synthesis followed by immobilisation on a glass substrate [6, 7]. Those with short nucleic acids (oligonucleotides up to 25 nt) are useful for the detection of mutations and expression monitoring, gene discovery and mapping. In the procedure of genomic analysis, both types of microarrays are exposed to a labelled sample, hybridised, and complementary sequences are determined (Figs. 1, 2, 3).

ARRAY PRODUCTION

A chip or microchip, in computer technology, is a small piece of semiconducting material containing an electronic circuit. Such chips are generally small, usually less than 5 cm per side. Their small size helps to make modern computers fast, compact and relatively inexpensive.

A comparable phenomenon is observed in molecular biology. The miniaturisation of certain tools is suitable for the construction of a smart and portable device – the spotted array system, which offers the pharmaceutical, biotechnology and agriculture industries more efficient and economical solutions.

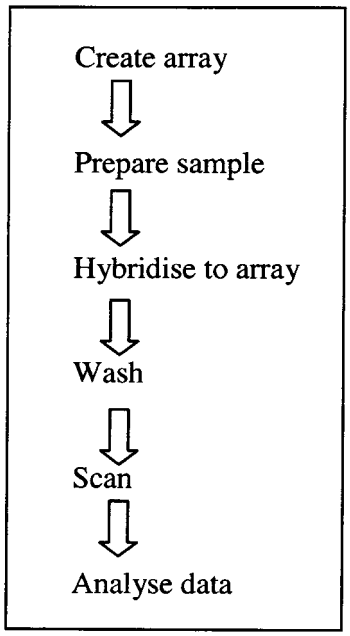

Figure 1. How to perform an array experiment.

Chips are built of semiconducting materials. A semiconductor is a material that is neither a good conductor of electricity nor a good insulator. The most common semiconductor materials used in making chips are the elements: silicon and germanium. However, nearly all chips are made from silicon. Production of microarrays begins with the selection of probes to be printed on the array. In most cases, these are chosen directly from databases including GeneBank [8] and UniGene [9]. There are three different methods for creating the microarray: spotting long DNA fragments, spotting prefabricated oligonucleotides, and *in situ* (on-chip) synthesis of oligonucleotides (Table 1).

For prokaryotes, probes are usually generated by amplifying genomic DNA with gene-specific primers. Spotting PCR products (of approximately 0.6–2.4 kb) representing specific genes onto a matrix produces DNA arrays. These PCR products are generated using chromosomal DNA as a template, and subsequently purified by precipitation or gel-filtration, or both. For both, glass and membrane,

each array dot is generated by depositing a few nanoliters of purified PCR product, usually at 100–500 μg/ml [10]. The printing is carried out by a robot (arrayer) that spots a sample of each gene product onto a number of matrices in a serial operation. The membranes commonly used are commercially available nitrocellulose and charged nylon that are employed in standard blotting assays (Southern blot, colony and plaque blot, dot and slot blot). The disadvantages of this method are that the genetic material is non-covalently attached which may result in its loss from the support, and that only a small amount of the DNA is available for hybridisation [11]. Glass-based arrays are most often made on microscope slides. They are coated with poly-lysine, amino silanes or amino-reactive silanes [4], which enhance both the hydrophobicity of the slide and the adherence of the deposited DNA. In most cases, DNA is cross-linked to the matrix by ultraviolet irradiation. After fixation, residual amines on the slide surface are reacted with succinic anhydride to reduce the positive charge at the surface. As the final step, the deposited DNA is split single-stranded by heat or alkali.

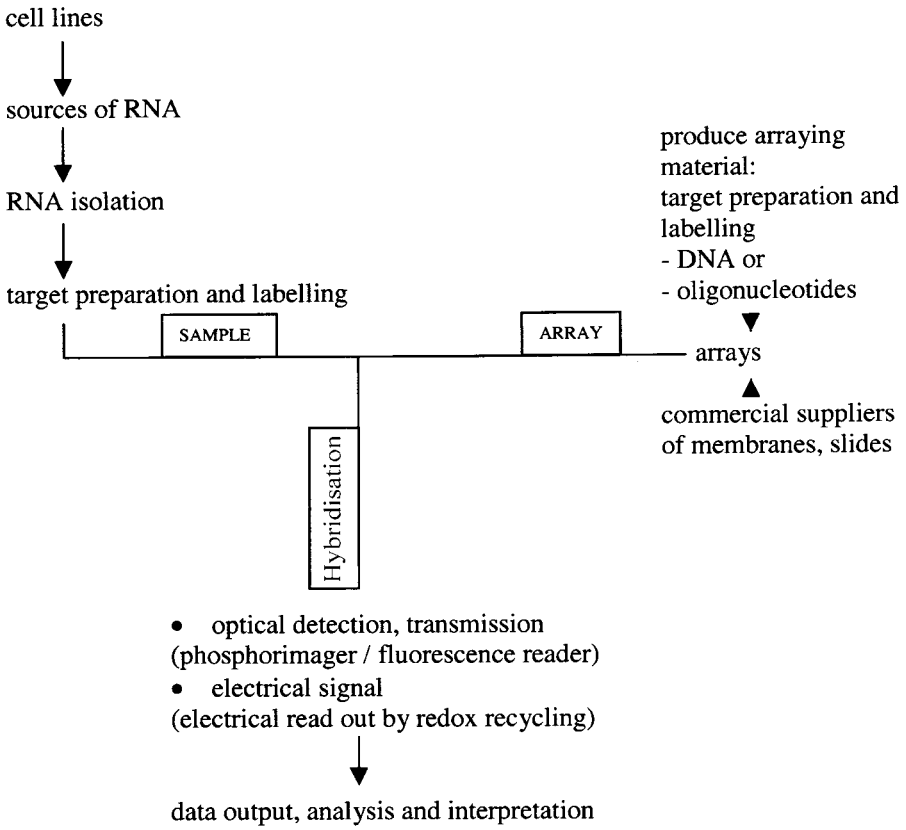

Figure 2. Components required for RNA-expression analysis using GeneChip-microarrays.

Oligonucleotide chips are produced by adapting semiconductor photolithography to synthesise oligonucleotide probes *in situ* on glass or membrane substrate. These chips are designed and produced on the basis of sequence information alone, without the need for any clones, PCR products, DNA and so on. Probe arrays are manufactured by light-directed chemical synthesis, which combines solid-phase chemical synthesis with photolithographic fabrication techniques employed in the semiconductor industry [12, 13]. Using a series of photolithographic masks to define chip exposure sites, followed by specific chemical synthesis steps, the process constructs high-density arrays of oligonucleotides, with each probe in a predefined position in the ar-

ray. Multiple probe arrays are synthesised simultaneously on a large glass wafer. This parallel process enhances reproducibility and helps achieve economies of scale [1].

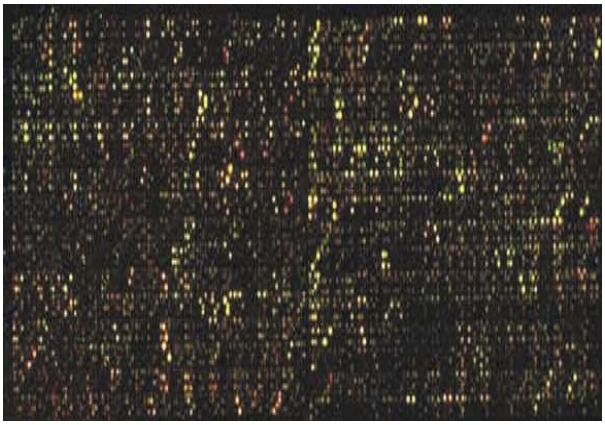

Figure 3. Example of the GeneChip-microarray, ready for analysis.

Oligonucleotides cannot be coupled directly to the surface by silanol groups of silicate glass or to most plastics. It is necessary to create the surface with a group from which the growth of the oligonucleotide chain can be initiated. Such spacers also help to overcome steric interference, consisting in the fact that the ends of the probes closest to the surface are less accessible than the ends further away [1, 14]. Oligonucleotides on long spacers extend away from their neighbours and from the surface, and thus they allow more efficient interaction with the target. The spacer's length has a marked effect on hybridisation yield. It has been shown that the optimal spacer length gives up to 150-fold increase in the yield of hybridisation [14].

The effects of different charged groups in the spacer were also examined, and it was shown that both positively and negatively charged groups in the spacer diminish the yield of hybridisation [14]. Additionally, the base composition and sequence of the oligonucleotides have also a large effect on duplex formation. The effect is of course due to the lower stability of A:T than G:C pairs. According to this, the oligonucleotides of the same length have correspondingly different T_m values. Adding an A:T base pair increases T_m by roughly 2°C, compared with 4°C for a G:C pair [1].

The technology for spotting arrays (DNA chips) is undoubtedly simpler than that for *in situ* fabrication. Simultaneous production of many arrays with the same set of probes makes the deposition more economical than *in situ* synthesis. Moreover, deposition is also a method of choice for long sequences, which are available as PCR products.

Table 1. Methods for creating a microarray

Method	Features and/or applications
I) spotting long DNA fragments:	
– fragment based DNA printing (Stanford University)	• array of spotted PCR products • gene expression analysis
II) array of prefabricated oligonucleotides:	
– gel pads (Motorola)	• oligonucleotides are attached to patches of activated polyacrylamide
–- microelectrodes (Nanogen)	• controlled electric fields for immobilisation
III) *in situ* synthesis of oligonucleotides:	
–- photolitography (Affymetrix)	• light-directed oligonucleotide-synthesis on chip • adapted from semiconductor industry
–- inkjet technology (Agilent)	• oligonucleotides are synthesised drop-by-drop • adapted from the technique used in ink-jet printers

Comparing the two types of oligonucleotide arrays, arrays of prefabricated oligonucleotides and *in situ* (on-chip) synthesis of oligonucleotides, the latter has some advantages over deposition of pre-synthesised oligonucleotides. It is not profitable to make large arrays using pre-synthesised oligonucleotides attached to the surface. On the other hand, it is difficult to assess the quality of the oligonucleotides made on a surface. Therefore, this technique could be used for a quality control, but it is not available for most biological laboratories. In contrast, the pre-synthesised oligonucleotides can be assessed before they are attached to the surface.

There is still some confusion in the nomenclature of the target and probe. It is commonly accepted that target determines the labelled material (DNA or RNA), and immobilised DNA or oligonucleotide is the probe. But sometimes probe is also defined as a piece of labelled DNA or RNA which is used in a hybridisation assay.

The targets for arrays are usually labelled representations of cellular RNA or DNA pools. It has been found that the short targets can interact more efficienty with oligonucleotide or DNA arrays than large ones. Ideally, target and probe should have the same length [5]. RNA as a target has a stable secondary structure, which can interfere with hybridisation. To reduce this effect, RNA can be fragmented, preferably to a size close to that of the probes on the arrays. Secondary structure is less of a problem with DNA targets and PCR products [1, 5].

There was also a patent battle over the GeneChip (Table 2). This is understandable because this technology has already been used to an incredibly diverse set of applications. DNA microarrays continue to gain popularity as a number of biotechnology companies aggressively pursue DNA chip technology enhancement and cost reduction. In recent years a number of different methods for the microarrays have been developed. Electric fields have been used to greatly accelerate the hybridisation of labelled target to immobilised sample oligonucleotides [15, 16]. The microarray was fabricated on a silicon chip one-centimetre square. The silicon substrate was thermally oxidised and then platinised to form a 1 mm × 1 mm array of 25 microelectrodes. The electrodes were covered with a

Table 2. The patent battle over the GeneChip

Year	Event
1987	Patent filed on Sequencing By Hybridisation (SBH) R. Drmanac, Belgrade Univ. → Argonne National Laboratory → HySeq
1988–1991	Several groups published reports on SBH E. Southern, Oxford University (Oxford Gene Technology) A.Mirzabekov, Engelhard Institute, Moscow → Agronne National Laboratory S. Fodor, Affymetrix W. Bains, Bath University
1989	European patent granted to Southern on "Oligonucleotide arrays as a testing platform"
1993	US patent on SBH granted to HySeq
1997–1998	HySeq accuses Affymetrix of patent infringement "We are not doing sequencing, but mutation detection"
1998	Courtcase between Southern and seven chip companies (Affymetrix, HySeq, Hoffman La Roche, Abbot etc.)
1998	US patent to Incyte (Synteni) on technology to print microarrays with density higher than 100 polynucleotides per cm^2
1998–1999	Affymetrix and Incyte (and others) accuse each other of patent infringements

permeation layer of streptavidin-agarose, to which biotinylated DNA sample was coupled under a positive potential. Such a use of electric field to increase the transport rate of negatively charged probe leads to a 10-fold increase in the hybridisation rate [17].

Another approach to perform and improve gene expression analysis is dynamic DNA hybridisation (DDH) on a chip using paramagnetic beads [18]. The advantages of this method are the dynamic supplies of both DNA sample (target) and probe, and also use of the paramagnetic beads as a transportable solid support. It reduces hybridisation time and makes the reaction more efficient. The magnetic coated beads are loaded with labelled capture sample or the capture samples are coated onto activated magnetic beads. In such a microfabricated device, simultaneous analysis of many samples is possible. The microfluidic platform of the device is developed for automated analysis of nanoliter volumes. A pneumatic pumping apparatus transports probes and other reagents into the microfluidic device while hydrostatic pumping is used for the introduction of beads with samples. The DNA sample/bead complex is introduced into the device in which hybridisation takes place with a complementary probe.

At the present time, the paramagnetic beads are extensively used for the preparation, separation and detection of biological molecules such as DNA, RNA and proteins. Their efficiency, simplicity and low cost are very favourable.

DNA chip is simple in concept, but generating probes on a solid array surface requires considerable expertise and technical tricks. Key matters in creating DNA arrays include: fidelity, reproducibility, ease of synthesis, flexibility, shelf life, cost, hybridisation conditions, steric considerations. As mentioned before, the major advantage of this technique is its high flexibility, which allows creation of a chip with any necessary sequences. But this approach has disadvantages difficult to break through and limiting its usage. First of all it is time and resources consuming to synthesise the needed number of different oligonucleotides. Oligonucleotide arrays have longer shelf lives, compared with DNA chips, which may only be useable for a few weeks. Another even more perspective direction is to extend the capability of the chip so one can detect thousands of genes simultaneously. Further, fluorescence technology, which is the most commonly used detection method for array readouts is reproducible, but is limited in sensitivity. That is why chemiluminescence, diode array detectors, direct electrical charge detection, and piezoelectric readout are all being developed as alternative detection methods. Accelerated hybridisation techniques are being developed by using electric field control.

Pointing out advantages and disadvantages of this technique shows us perspectives and directions of future researches. The innovation of DNA chips opens up a window into the complex world of biology and provides a good tool for the coming post genomic project research.

APPLICATIONS

The GeneChip technology may be employed in diagnostics (mutation detection), gene discovery, gene expression and mapping [15, 19]. It is used to measure expression levels of genes in bacteria, plant, yeast, animal and human samples [20–22].

At the present time, the main large-scale application of microarrays is comparative expression analysis [23, 24]. The microarray technology provides the possibility to analyse the expression profiles for thousands of genes in parallel. Another application is the analysis of DNA variation on a genome-wide scale. Both of these applications have many common requirements.

By hybridisation with labelled mRNA, cDNA, arrayed PCR products or oligonucleotides on a substrate have been successfully used for monitoring transcript levels [2], sin-

gle nucleotide polymorphism (SNP) [19], or genomic variations between different strains [25]. One of the most significant applications of this technique is, as mentioned above, gene expression profiling on the whole genomic scale [20, 21]. For example, the expression levels of the genes in the *Saccharomyces cerevisiae* genome have been successfully determined with both the DNA and oligonucleotide microarray technology [7]. This technique has also been used to investigate physiological changes in human cells [22]. DNA microarray technology was applied to detect differential transcription profiles of a subset of the *Escherichia coli* genome [20].

The microarray technology is a powerful yet economical tool for characterising gene expression regulation and will prove to be useful for strain improvement and bioprocess development. It may prove to be useful for strain development, process diagnosis, and process monitoring in bioreactors.

Information obtained from DNA chip analysis may enable researchers to determine the impact of a drug on a cell or group of cells, and consequently to determine the drug's efficacy or toxicity [26]. Knowledge of gene expression profiles can also help researchers to identify new drug targets [26].

The BioChip opens a new world of diagnostics based on genetics. This technology may be adequate to answer many medical questions. For example, gene expression profiles can be used for classification of tumours and for prognosis.

The technology finds increasing application in fundamental and applied research. The major feature of this technique is that it allows one to perform a simultaneous analysis of a great number of DNA sequences. We can study genome as a whole. This new direction is named "genomics". As the Human Genome Project nears completion a new era of genomic science is beginning.

The GeneChip technology is a new technique that undoubtedly will substantially increase the speed of molecular biology research. This paper gives a survey of gene-microarray technology and its use in gene expression studies. Perhaps more other applications of this technology will be possible in the near future.

We thank Joanna Potrykus and Katarzyna Potrykus for critical reading of the manuscript.

REFERENCES

1. Southern, E., Mir, K. & Shchepinov, M. (1999) Molecular interactions on microarrays. *Nat. Genet.* **21**, 5–9.
2. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.
3. Lockhart, D.J. *et al.* (1996) Expression monitoring by hybridisation to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14**, 1675–1680.
4. Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O. & Davis, R.O. (1996) Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10614–10619.
5. Lipshutz, R.J., Fodor, S.P.A., Gingeras, T.R. & Lockhart, D.J. (1999) High density synthetic oligonucleotide arrays. *Nat. Genet.* **21**, 20–24.
6. McGall, G.H. *et al.* (1997) The efficiency of light-directed synthesis of DNA arrays on glass substrates. *J. Am. Chem. Soc.* **119**, 5081–5090.
7. Wodicka, L., Dong, H., Mittmann, M., Ho, M.-H. & Lockhart, D.J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15**, 1–15.
8. Benson, D.A., Boguski, M.S., Lipman, D.J. & Ostell, J. (1997) GenBank. *Nucleic Acids Res.* **25**, 1–6.

9. Schuler, G.D. *et al.* (1996) A gene map of the human genome. *Science* **274**, 540–546.

10. Cheung, V.G., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R. & Childs, G. (1999) Making and reading microarrays. *Nat. Genet.* **21**, 15–19.

11. Gingeras, T.R., Kwoh, D.Y. & Davis, G.R. (1987) Hybridization properties of immobilized nucleic acids. *Nucleic Acids Res.* **15**, 5373–5390.

12. Fodor, S.P.A., Read, L.J., Pirrung, M.C., Stryer, L., Lu, A.M. & Solas, D. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science* **251**, 767–773.

13. McGall, G., Labadie, J., Brook, P., Wallraff, G., Nguyen, T. & Hinsberg, W. (1996) Ligh-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5022–5026.

14. Shchepinov, M.S., Case-Green, S.C. & Southern, E.M. (1997) Steric factors influencing hybridisation of nucleic acids to nucleotide arrays. *Nucleic Acids Res.* **25**, 1155–1161.

15. Sosnowski, R.G., Tu, E., Butler, W.F., O'Connell, J.P. & Heller, M.J. (1997) Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1119–1123.

16. Heller, R.N. & Tu, E. (1993) Active programmable electronic devices for molecular biomolecules. US Pat. 5605662.

17. McIntyre, P.E. (1996) Microfabrication technology for DNA sequencing. *Trends Biotechnol.* **14**, 69–73.

18. Hugh Fan, Z., Mangru, S., Granzow, R., Heaney, R., Ho, W., Dong, Q. & Kumar, R. (1999) Dynamic DNA hybridisation on a chip using paramagnetic beads. *Anal. Chem.* **71**, 4851–4859.

19. Wang, D.G. *et al.* (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280**, 1077–1082.

20. Richmond, C.S., Glasner, J.D., Mau, R., Jin, H. & Blattner, F.R. (1999) Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* **27**, 3821–3835.

21. Oh, M.K. & Liao, J.C. (2000) Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. *Biotechnol. Prog.* **16**, 278–286.

22. Iyer, V., DeRisi, J., Eisen, M., Ross, D., Spellman, P., Hudson, J., Jr. *et al.* (1997) Use of DNA microarrays to monitor differential gene expression in yeast and humans. *FASEB J.* **11**, 1126–1131.

23. Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J.M. (1999) Expression profiling using cDNA microarrays. *Nat. Genet.* **21**, 10–14.

24. Southern, E.M. (1996) DNA chip: analysing sequence by hybridisation to oligonucleotides on a large scale. *Trends Genet.* **12**, 110–115.

25. Behr, M.A., Wilson, M.A., Gill, W.P., Salamon, H., Schoolnik, G.K., Rane, S. & Small, P.M. (1999) Comparative genomics of BCG vaccines by whole genome DNA microarray. *Science* **284**, 1520–1523.

26. Debouck, C. & Goodfellow, P. (1999) DNA microarrays in drug discovery and development. *Nat. Genet.* **21**, 48–50.