Review article

Mutational analysis using oligonucleotide microarrays

Joseph G Hacia, Francis S Collins

The development of inexpensive high throughput methods to identify individual DNA sequence differences is important to the future growth of medical genetics. This has become increasingly apparent as epidemiologists, pathologists, and clinical geneticists focus more attention on the molecular basis of complex multifactorial diseases. Such undertakings will rely upon genetic maps based upon newly discovered, common, single nucleotide polymorphisms. Furthermore, candidate gene approaches used in identifying disease associated genes necessitate screening large sequence blocks for changes tracking with the disease state. Even after such genes are isolated, large scale mutational analyses will often be needed for risk assessment studies to define the likely medical consequences of carrying a mutated gene.

This review concentrates on the use of oligonucleotide arrays for hybridisation based comparative sequence analysis. Technological advances within the past decade have made it possible to apply this technology to many different aspects of medical genetics. These applications range from the detection and scoring of single nucleotide polymorphisms to mutational analysis of large genes. Although we discuss published scientific reports, unpublished work from the private sector^{1,2} could also significantly affect the future of this technology. (*J Med Genet* 1999;36:730-736)

Keywords: mutational analysis; oligonucleotide microarrays; DNA chips

DNA chip fabrication

The generic term "DNA chips" refers to miniaturised arrays of nucleic acid segments anchored on glass supports no larger than a microscope slide. Larger segments, usually greater than 100 bp, have been primarily used to measure the expression levels of thousands of RNA species within cells or tissues.^{3,4} They are often called cDNA microarrays in reference to the cDNA nucleic acid components. Shorter segments, in the range of 8-25 nt, can also be anchored on solid supports; these are often called oligonucleotide microarrays or high density oligonucleotide arrays. Oligonucleotide microarrays manufactured by Affymetrix (Santa Clara, CA) have been designated Gene Chips. Such microarrays may be used in RNA

expression level analysis^{5,6} in addition to mutational and comparative sequence analysis.

Oligonucleotide arrays can be manufactured using two distinct approaches. In one, oligonucleotide solutions are deposited onto glass surfaces using either simple spotting techniques^{7,8} or microfabricated ink jet pumps.⁹ The other approach involves synthesising oligonucleotides on the chip surface. Commercially available oligonucleotide arrays (comprising over 250 000 25mer oligonucleotides) supplied by Affymetrix are manufactured using photolithographic techniques developed in the semiconductor industry for computer microchip manufacture and modified oligonucleotide synthesis chemistry.^{6,10} High precision delivery of chemical reagents using microfluidic channel¹¹ and ink jet pump⁹ based technologies can also allow conventional oligonucleotide synthesis protocols to be adapted for this purpose.

Target preparation, hybridisation, and detection

Depending on the types of DNA chips used or the specific application, the terms "probe" and "target" can either refer to the arrayed nucleic acids or the nucleic acid sample in solution. In this review we will refer to the oligonucleotides on the surface of DNA chips as "probes" and the nucleic acid in solution as "target".

Target preparation begins with PCR amplification reactions using genomic DNA, RNA, or cloned templates. Random fragmentation of PCR products promotes hybridisation to DNA chips by producing targets with accessible single stranded segments.¹² This can also be accomplished using asymmetrical PCR reactions.¹³ Similarly, single stranded RNA can be generated through in vitro transcription reactions using PCR products with terminal RNA polymerase promoter sequences as templates.¹⁴ Single stranded targets should be randomly fragmented before analysis to decrease inter- and intramolecular structures which inhibit array hybridisation.^{15,16} Although hybridisation is typically diffusion controlled, it can be accelerated using spatially defined electric fields which increase target concentration near the array surface.¹⁷⁻²⁰

Target hybridisation is typically detected using fluorescent based approaches. Targets are usually internally labelled with fluorescent dyes or haptens at their 5' or 3' ends.^{13,14,21}

Genetics and
Molecular Biology
Branch, National
Human Genome
Research Institute,
Building 49/3A14,
National Institutes of
Health, Bethesda,
Maryland 20892, USA
J G Hacia
F S Collins

Correspondence to:
Dr Collins.

Target	
3' CCTTAACGATCGATGGATCGATCGATCTA 5'	
Sub A	5' AATTGCTAGCTAACTAGCTAGCTAG 3'
Sub C	AATTGCTAGCTACCTAGCTAGCTAG
Sub G	AATTGCTAGCTAGCTAGCTAGCTAG
Sub T	AATTGCTAGCTATCTAGCTAGCTAG
Del G	AATTGCTAGCTACTAGCTAGCTAGA
Ins A	AATTGCTAGCTAACTAGCTAGCTA
Ins C	AATTGCTAGCTACCTAGCTAGCTA
Ins G	AATTGCTAGCTAGCTAGCTAGCTA
Ins T	AATTGCTAGCTATCTAGCTAGCTA
Perfect Match	AATTGCTAGCTACCTAGCTAGCTAG

Figure 1 Typical probes used for gain of signal sequence analysis. An arbitrary target sequence given in red was chosen to show the nucleotide composition of various classes of arrayed probes. Substitution (sub) probes interrogate for all possible single nucleotide substitution sequence variations. Deletion (del) probes interrogate for all possible nucleotide deletion lengths of a given size. Insertion (ins) probes interrogate for inserted sequences (single nucleotide insertions are the most feasible to represent completely on the array). The inserted base is shown in green. Perfect match probes are fully complementary to wild type sequence. These are shown in blue along with the perfect match probe present within the substitution probe series.

Fluorescent signals are produced by exciting the array surface with a laser and detected using commercially available confocal microscopes equipped with either photon multiplier tubes or CCD cameras.²²

Universal combinatorial oligonucleotide arrays

In the 1980s, proposals were made to determine unknown DNA sequences by hybridising nucleic acid targets to an arrayed library of all possible oligonucleotides of a given length, usually 8mers (over 65 536 different species).^{23–26} Sophisticated signal processing algorithms would be used to analyse hybridisa-

Figure 2 Global target hybridisation to an ATM oligonucleotide microarray. Magnified false coloured image (1.2 cm × 1.2 cm edge length) of fluorescently labelled ATM antisense target to an oligonucleotide array interrogating the complete coding sequence of the ATM gene. Areas in white and yellow depict strong hybridisation signal while areas in blue and black show regions with weak hybridisation signal.

tion patterns and elucidate the sequence of virtually any DNA fragment. Significant technical issues involving complex hybridisation patterns and repetitive sequence elements still need to be addressed before these proposals are implemented on a practical basis.

Oligonucleotide microarrays consisting of all 262 144 possible 9-mer oligonucleotide probes have been used to screen for DNA sequence changes.¹² To circumvent problems of analysing unknown sequences (de novo sequence analysis), PCR products with a known reference sequence (resequencing analysis) were examined. Comparisons between test and reference sample hybridisation patterns help accentuate signals stemming from sequence differences. This approach was initially tested in a study analysing two 24mer homopyrimidine oligonucleotide targets containing single nucleotide differences using arrays of all 256 possible homopurine octamers.²⁷ The 9mer probes were present as the overhanging ends of short duplexes to allow an enzymatic ligation strategy which enhanced hybridisation specificity. High base calling accuracy was found for a 500 bp target (99% of the bases correctly called); however, it decreased for larger targets (89 and 74% of the bases called correctly in 2.5 and 5.4 kb targets respectively). It remains to be seen how effective this approach is towards detecting heterozygous sequence changes since only one such case was evaluated.

The format of this approach can be inverted with complex DNA targets now being placed on a surface and then hybridised to oligonucleotides in solution. This utility of approach was tested in a study where *TP53* exons 5–8 were amplified from 12 samples, spotted onto nylon filters, and then individually probed with 8192 non-complementary radiolabelled 7-mer oligonucleotides.²⁸ All 13 distinct homozygous or heterozygous sequence changes were detected. However, since thousands of separate hybridisation reactions are needed for each sample, this approach may only be suitable for repetitive assays in large clinical diagnostic laboratories.

Customised oligonucleotide arrays

Almost all of the recent applications of DNA chips for mutation detection have used oligonucleotide microarrays designed to evaluate a specific DNA sequence. Although this approach is less flexible than using universal microarrays, customised microarray based assays show high levels of sensitivity and specificity. Data from these experiments may be interpreted using two complementary analytical schemes relying upon “gain” and “loss” of hybridisation signal.²⁹

Gain of signal analysis and customised microarray design

“Gain” of hybridisation signal analysis compares signals from probes complementary to mutant and wild type sequences (figs 1 and 2). Relative to their wild type counterparts, mutant targets should have increased affinity towards a corresponding mutation specific probe (fig 3). This results in a “gain” of

Figure 3 Specific target hybridisation patterns to ATM oligonucleotide microarrays. Magnified digitised grey scale images, 50 μ feature size, showing hybridisation patterns of fluorescently labelled ATM reference and test antisense targets to oligonucleotide microarrays. Both panels show the region of the array interrogating nucleotides 7317-7337. Nucleotide identities are given under each respective column with the position 7327 underlined. The upper panel depicts the hybridisation pattern of a reference sample with a 7327 C/C genotype. The lower panel shows the hybridisation patterns of a sample with a 7327 C/T genotype, corresponding to a R2443X nonsense mutation.

hybridisation signal to this probe. However, only mutations with a corresponding complementary probe represented in the array can be detected. To interrogate both strands of an N bp long sequence for all single nucleotide substitutions, the array should consist of 8N probes (4N probes per strand). This total is derived from the fact that four probes complementary to each of the four possible target sequences for each nucleotide position on a given strand are represented in the array. When analysing both strands of an N bp sequence, 2N probes are needed to scan for all possible deletions of a given length while $2(4^X)N$ probes are needed to scan for all X nt long insertions.

Figure 4 Perfect match probes used in loss of signal sequence analysis. In theory, each target nucleotide hybridises to a set of n overlapping n mer perfect match probes in the oligonucleotide microarray. An arbitrary sequence is given to show the sequence composition of perfect match probes used in loss of signal analysis. In this example, hybridisation to 15 overlapping 15mer probes (shown in red) are affected by changes in a single target nucleotide. Probes shown in blue would be unaffected.

For example, 80 000 probes can be used to scan for all single nucleotide substitutions in a 10 kb target. However, 100 000 more probes are needed to scan for all 1-5 bp deletions and 27 280 000 more probes are needed to scan for all 1-5 bp insertions. Thus it is impractical to screen large targets for insertions beyond a single base pair using this approach.

In some cases, hybridisation to a mutation specific probe can be used to identify the nature of a sequence change; however, this is not always a correct or unambiguous assignment³⁰ and in general a putative mutation needs to be confirmed by an independent method. Misidentifying a sequence change could have substantial impact on assessing its functional relevance. Nevertheless, mutation identity assignment is robust for common variants having known hybridisation patterns.

Loss of hybridisation signal analysis

Sequence variations between test and reference samples may also be detected by quantitating relative losses of target hybridisation signals to oligonucleotides perfectly matched to the reference sequence (perfect match probes). In this "loss" of hybridisation signal approach, homozygous sequence change should cause a complete loss of signal to perfect match probes interrogating the sequence tract surrounding the change (fig 4). Ideally, heterozygous targets would produce a 50% loss of signal intensity relative to wild type target for perfect match probes interrogating the sequence change as well as diminished hybridisation signal to flanking probes.

Loss of signal analysis accommodates hybridisation based screens for virtually any sequence variation. Arrays interrogating both target strands N bp in length for all possible sequence changes minimally consist of 2N overlapping probes. For example, 11 000 oligonucleotides would be needed to screen the 5.5 kb *BRCA1* gene for all possible sequence variations. Furthermore, random sources of error are minimised since multiple probes (N probes N nucleotides in length) are involved in detecting a sequence variation. The sensitivity and specificity of this approach is improved in "two colour" assays where known reference and unknown test targets labelled with different fluorophores are cohybridised and directly compared.³⁰⁻³² By plotting the ratio of perfect match probe signal intensities from test and reference samples, peaks with distinct width and height properties which indicate the presence of sequence variations are shown (fig 5). Since the exact nucleotide sequence of the variation cannot be determined using loss of signal analysis, the region surrounding the proposed sequence variation must be dideoxy sequenced to identify the change.

Screening for known and unknown sequence variants

In non-genetically isolated populations without a history of severe bottlenecks, many highly penetrant disease genes have complex mutation spectra. Such a situation is found for the hereditary breast and ovarian cancer gene

Figure 5 Mutation detection using two colour loss of signal analysis. Fluorescein labelled reference and biotinylated test targets were cohybridised to an oligonucleotide microarray designed to interrogate the coding region of the *ATM* gene for all possible sequence changes.³² To compensate for consistent differences in reference and test target hybridisation efficiencies, the ratio of reference to test signal at each wild type position was normalised relative to ratios derived from 10 separate cohybridisation experiments. The averaged sense and antisense strand ratios are given along with the identity of each exon listed below the appropriate data. Panels A, B, and C represent data derived from exons 4–24, 25–46, and 47–65 respectively. The peak encompassing the mutated 5932 G/T position (a nonsense mutation) is present in both sense and antisense strand data sets.

BRCA1, for example, where over 400 distinct mutations have been identified.³³ Because of this diversity, thorough comparative sequencing and mutational analysis involves screening for all possible changes in the homozygous and heterozygous states.

CFTR ASSAY

In an early example, oligonucleotide microarrays were designed to screen for all possible single nucleotide substitutions in the 95 nt long *CFTR* exon 11 coding region as well as 37 known mutations.¹³ In a blinded study, 10 genomic DNA samples were successfully genotyped by comparing hybridisation signals from test and wild type reference samples at mutation specific probes. This was the first published large scale application of the gain of hybridisation signal approach.

HIV-1 PROTEASE AND REVERSE TRANSCRIPTASE ASSAY

Oligonucleotide arrays were used to screen the entire 297 bp HIV-1 protease (*pr*) gene coding sequence for all possible single nucleotide substitutions using the gain of hybridisation signal approach.¹⁴ When screening 114 samples for HIV-1 *pr* sequence changes, there was a 98.26% agreement between oligonucleotide microarray based analysis and dideoxy sequencing. DNA chip analysis is especially well suited for this system for two reasons. First, base substitutions will represent the majority of changes since functional full length protein is needed for viral survival. On average, base substitutions are easier to detect than small inser-

tions and deletions using hybridisation analysis. Secondly, HIV-1 isolates probably represent clonal populations which makes the assay similar to screening for homozygous changes in genomic DNA sequences. Homozygous changes are much easier to detect than heterozygous changes since wild type allele hybridisation signals which may mask the hybridisation signatures of the mutant allele are completely absent.

MYCOBACTERIUM SPECIES IDENTIFICATION

An oligonucleotide microarray designed to analyse a 705 bp segment of the *Mycobacterium tuberculosis* *rpoB* gene accurately detected rifampin resistance associated with mutations of 44 clinical isolates of *M. tuberculosis*.³⁴ The nucleotide sequence diversity in 121 *Mycobacterium* isolates (comprised of 10 species) was examined both by dideoxy sequencing and oligonucleotide array based analysis. Species identification could be obtained with equivalent accuracy using either dideoxy sequencing or oligonucleotide microarray based hybridisation analysis. The same array could be used to identify non-tuberculous *Mycobacteria* species.

MITOCHONDRIAL GENOME ASSAY

A pair of oligonucleotide microarrays consisting of over 135 000 probes was used to interrogate the entire 16.6 kb human mitochondrial genome from 10 samples.³¹ Using gain of hybridisation signal analysis, 99% of the genome could be read correctly. The remaining 1% of the genome would have to be checked using a complementary technology such as

dideoxy sequencing analysis. While analysing a 2.5 kb sequence tract from 12 samples, 179/180 polymorphisms were detected using gain and loss of signal analysis although confirmatory dideoxy sequencing was recommended in some cases. Although this was the first reported study analysing large targets for all possible variants, it did not assay for heterozygous sequence changes.

BRCA1 ASSAY

Gain and loss of signal analysis were directly evaluated when scanning for heterozygous sequence variations in the 3.43 kb *BRCA1* exon 11 sequence.³⁰ A two tiered algorithm for mutational analysis based upon both forms of analysis allowed 14/15 heterozygous mutations scattered throughout the exon to be detected. Single nucleotide substitutions generally produced more robust gain and loss of hybridisation signal signatures than small insertions and deletions. Data from both target strands had to be assessed since several sequence changes were more readily detected on one strand than the other. The loss of signal assay showed increased sensitivity and specificity relative to the gain of signal assay. This results from the fact that the gain of signal assay is relatively insensitive towards detecting larger deletions and insertions owing to cross hybridisation of wild type target to mutation specific probes.

The same oligonucleotide microarrays were also used to analyse *BRCA1* exon 11 orthologues from great apes, Old and New World monkeys, prosimians, and other mammals.³⁵ These were all approximately 3.4 kb in length and ranged from 98.2% to 83.5% nucleotide identity relative to human. Retrospective guidelines for identifying high fidelity hybridisation based sequence calls were formulated based upon dideoxy sequencing analysis. Prospective application of these rules yielded base calling with at least 98.8% accuracy over sequence tracts shown to have approximately 99% identity relative to human. A second tier confirmatory DNA chip based strategy was proposed that could allow the complete sequence of the chimpanzee, gorilla, and orangutan orthologues to be deduced solely through hybridisation based methodologies. Furthermore, DNA chip based analysis of less highly conserved orthologues could identify conserved nucleotide tracts and can provide information for primer design.

Oligonucleotide microarrays of over 96 000 oligonucleotides were more recently designed to screen the entire 5.53 kb coding region of the human *BRCA1* gene for all possible sequence changes in the homozygous and heterozygous states.³⁶ Preliminary studies investigated the thermodynamic properties of array hybridisation. Fluorescent hybridisation signals from RNA targets containing the four natural bases to over 5592 different fully complementary 25mer oligonucleotide probes on the chip varied over two orders of magnitude. To examine the thermodynamic contribution of RNA/DNA base pairs to this variability, modified nucleoside 5'-triphosphates were incorporated into *BRCA1* targets. Targets con-

taining 5-methyluridine displayed promising localised enhancements in hybridisation signal, especially in pyrimidine rich target tracts, while maintaining single nucleotide mismatch hybridisation specificities comparable to those of unmodified targets.

ATM ASSAY

High oligonucleotide density arrays were used to screen for all possible heterozygous germline mutations in the 9.17 kb coding region of the *ATM* gene.³² A strategy for rapidly developing multiplex amplification protocols in DNA chip based hybridisation analysis was devised and implemented in preparing target for the 62 *ATM* coding exons. In a blinded study, 17 of 18 distinct heterozygous and eight of eight distinct homozygous sequence variants in the assayed region were accurately detected along with five false positive calls while scanning over 200 kb in 22 genomic DNA samples. Of eight heterozygous sequence changes found in more than one sample, six were detected in all cases. Five previously unreported sequence changes, not found by other mutational scanning methodologies on these same samples, were detected that led to either amino acid changes or premature truncation of the ATM protein.

SINGLE NUCLEOTIDE POLYMORPHISM SCREENS

Oligonucleotide arrays have been used in large scale identification and genotyping of single nucleotide polymorphisms (SNPs) in the human genome.²¹ One hundred and forty nine chip designs, each containing 150-300 000 oligonucleotides, were used in screening 2.3 Mb of human sequence for SNPs. A total of 3241 candidate SNPs were found through gain and loss of signal analysis. A separate chip containing a simple tiling scheme for scoring individual SNP alleles was designed to genotype 500 markers simultaneously. Since single nucleotide substitutions are generally easier to detect than insertions and deletions, and SNP screens need not be exhaustive, this is a robust application for microarray technology.

DNA chip assays designed to evaluate 500 human SNPs were applied to chimpanzee and gorilla genomic DNA samples to determine the distant history of these human polymorphic sites.³⁷ Hybridisation based analysis allowed 214 ancestral alleles (the sequence found in the last common ancestor of humans and chimpanzees) to be assigned. It was proposed that information about the ancestral states of SNP sites could increase their utility in linkage disequilibrium studies³⁸⁻⁴⁰ since recently derived SNP alleles might be associated with greater shared DNA segment lengths.

OLIGONUCLEOTIDE ARRAY BASED MINISEQUENCING ASSAYS

Minisequencing assays are another oligonucleotide microarray based approach towards scanning for all possible sequence variations.⁴¹⁻⁴³ In this strategy, oligonucleotides are typically tethered to the surface via a 5' linkage leaving an exposed 3'-OH group or placed into a microtitre dish format.^{44 45} Labelled dideoxyribonucleoside triphosphates

are used in primer extension reactions with the annealed target and oligonucleotide probes respectively serving as template and primer. A mixture of all four dideoxyribonucleoside triphosphates, each labelled with a different dye, can be used in the extension reactions.⁴⁵ The identity of the added dideoxyribonucleotide is used to assign the identity or identities of the target nucleotide adjacent to the 3' end of each probe. Similar to the loss of hybridisation signal approach, minisequencing arrays designed to interrogate both N bp target strands for all possible sequence changes minimally consist of 2N overlapping probes. Heterozygous single nucleotide substitutions will produce two signals corresponding to the identity of the two alleles. Although the exact nature of insertions and deletions cannot always be determined, their end points can be elucidated. Thus far the most complex system analysed by this approach involved scanning a 33 base region of the *TP53* gene for all possible sequence changes.⁴⁵

Current strengths and limitations

Oligonucleotide array based sequence analysis is especially well suited for several applications. One is the detection of commonly occurring sequence changes for which specific hybridisation patterns are known. Another is in the detection of unknown homozygous base changes as was previously described for HIV-1¹⁴ and mitochondrial genome³¹ analysis. This can also be useful in screening for fixed base changes among different species³⁵ and in the analysis of hemizygous X and Y chromosome sequences. Microarray based assays for heterozygous sequence changes are most useful when a 5–10% false negative error rate is acceptable. Such studies could include non-exhaustive SNP screens or epidemiological studies where genotype information is not shared with participants.

The nucleotide sequence composition of the DNA segment being analysed is crucial in defining the sensitivity and specificity of DNA chip assays. Repetitive sequence elements and highly structured nucleic acids can greatly decrease the sensitivity of hybridisation based analysis. For example, DNA chip technology would be particularly ill suited for analysing genes commonly having triplet repeat based mutations⁴⁶ which represent an extreme form of repetitive sequence elements.

DNA chip assays are also vulnerable to problems involving target preparation. Like other PCR based assays, DNA chip assays are subject to errors associated with primer binding site mutations which allow only one allele to be evaluated. As with other methodologies, analysing rare viral species or tumour DNA samples where sequence changes are present in a trace population presents substantial problems. In the latter analysis, advances in target preparation, such as laser capture microdissection,⁴⁷ may be used to increase assay sensitivity.

Even an infallible mutation detection technology with perfect analytical validity can yield unclear results owing to an incomplete under-

standing of the structure and function of the gene being analysed. For example, single nucleotide substitutions causing amino acid changes may represent benign variants or missense mutations which deleteriously alter protein function. Such changes do not lend themselves to unambiguous interpretation without the assistance of robust functional assays.

Conclusions

Although certain technical challenges must still be overcome before hybridisation based sequence analysis becomes a common tool for molecular geneticists, it has significant potential to alter the way resequencing analysis is currently performed. Nevertheless, a balanced view is needed to assess its current strengths and limitations. It is unrealistic to expect any single technology will detect all possible sequence changes in complex DNA samples. However, with anticipated incremental technological improvements in assay conditions^{36–48–53} and array manufacture,^{54–55} DNA chip based assays will probably become commonly used tools for resequencing and mutational analysis.

We especially thank Larry Brody, Keith Edgemon, Aeryn Mayer, and Bryan Sun from NIH for helpful suggestions and discussion.

- 1 Marshall A, Hodgson J. DNA chips: an array of possibilities. *Nat Biotech* 1998;16:27–31.
- 2 Schena M, Heller RA, Theriault TP, Konrad K, Lachmeier E, Davis RW. Microarray: biotechnology's discovery platform for functional genomics. *Trends Biotech* 1998;16:301–6.
- 3 Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270:467–0.
- 4 Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. *Nat Genet Suppl* 1999;21:10–14.
- 5 Lockhart DJ, Dong H, Byrne MC, et al. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotech* 1996;14:1675–80.
- 6 Lipshutz RJ, Fodor SPA, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet Suppl* 1999;21:20–4.
- 7 Guschin D, Yershov G, Zaslavsky A. Manual manufacturing of oligonucleotide, DNA, and protein microchips. *Anal Biochem* 1997;250:203–11.
- 8 Proudnikov D, Timofeev E, Mirzabekov A. Immobilization of DNA in polyacrylamide gel for the manufacture of DNA and DNA oligonucleotide microchips. *Anal Biochem* 1998;259:34–41.
- 9 Blanchard AP, Kaiser RJ, Hood LE. High-density oligonucleotide arrays. *Biosens Bioelectron* 1996;11:687–90.
- 10 Fodor SPA, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D. Light-directed spatially addressable parallel chemical synthesis. *Science* 1991;251:767–73.
- 11 Matson RS, Rampal J, Pentoney SL, Anderson PD, Coassin P. Biopolymer synthesis on polypropylene supports: oligonucleotide arrays. *Anal Biochem* 1995;224:110–16.
- 12 Gunderson KL, Huang XC, Morris MS, Lipshutz RJ, Lockhart DJ, Chee MS. Mutation detection by ligation to complete *n*-mer DNA arrays. *Genome Res* 1998;8:1142–53.
- 13 Cronin MT, Fucini RV, Kim SM, Masino RS, Wespi RM, Miyada CG. Cystic fibrosis mutation detection by hybridization to light-generated DNA probe arrays. *Hum Mutat* 1996;7:244–55.
- 14 Kozal MJ, Shah N, Shen N, et al. Extensive polymorphisms observed in HIV-1 cladeB protease gene using high-density oligonucleotide arrays. *Nat Med* 1996;2:753–9.
- 15 Southern ES, Mir K, Shchepinov M. Molecular interactions on microarrays. *Nat Genet Suppl* 1999;21:5–9.
- 16 Milner N, Mir KU, Southern EM. Selecting effective antisense reagents on combinatorial oligonucleotide arrays. *Nat Biotech* 1997;15:537–41.
- 17 Edman CF, Raymond DE, Wu DJ, et al. Electric field directed nucleic acid hybridization on microchips. *Nucleic Acids Res* 1997;25:4907–14.
- 18 Sosnowski RG, Tu E, Butler WF, O'Connell JP, Heller MJ. Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control. *Proc Natl Acad Sci USA* 1997;94:1119–23.
- 19 Cheng J, Sheldon EL, Wu L, et al. Preparation and hybridization analysis of DNA/RNA from *E coli* on microfabricated bioelectronic chips. *Nat Biotechnol* 1998;16:541–6.

- 20 Gilles PN, Wu DJ, Foster CB, Dillon PJ, Chanock SJ. Single nucleotide polymorphic discrimination by an electronic dot blot assay on semiconductor microchips. *Nat Biotechnol* 1999;17:365-70.
- 21 Wang DG, Fan JB, Siao CJ, *et al.* Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 1998;280:1077-82.
- 22 Bowtell DD. Options available-from start to finish-for obtaining expression data by microarray. *Nat Genet Suppl* 1999;21:25-32.
- 23 Bains W, Smith G. A novel method for nucleic acid sequence determination. *J Theor Biol* 1988;135:303-7.
- 24 Lysov IU, Florent'ev VL, Khorlin AA, Khrapko KR, Shik VV. Determination of the nucleotide sequence of DNA using hybridization with oligonucleotides. A new method. *Dokl Akad Nauk SSSR Ser Biol* 1988;303:1508-11.
- 25 Khrapko KR, Lysov YP, Khorlyn AA, Shick VV, Florentiev VL, Mirzabekov AD. An oligonucleotide hybridization approach to DNA sequencing. *FEBS Lett* 1989;256:118-22.
- 26 Drmanac R, Labat I, Brukner I, Crkvenjakov R. Sequencing of megabase plus DNA by hybridization: theory of the method. *Genomics* 1989;4:114-28.
- 27 Southern EM, Maskos U, Elder JK. Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics* 1992;13:1008-17.
- 28 Drmanac S, Kita D, Labat I, *et al.* Accurate sequencing by hybridization for DNA diagnostics and individual genomics. *Nat Biotech* 1998;16:54-8.
- 29 Hacia JG. Resequencing and mutational analysis using oligonucleotide microarrays. *Nat Genet Suppl* 1999;21:42-7.
- 30 Hacia JG, Brody LC, Chee MS, Fodor SP, Collins FS. Detection of heterozygous mutations in *BRCA1* using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nat Genet* 1996;14:441-7.
- 31 Chee M, Yang R, Hubbell E, *et al.* Accessing genetic information with high-density DNA arrays. *Science* 1996;274:610-14.
- 32 Hacia JG, Sun B, Hunt N, *et al.* Strategies for mutational analysis of the large multiexon *ATM* gene using high-density oligonucleotide arrays. *Genome Res* 1998;8:1245-58.
- 33 A listing of *BRCA1* mutations can be found in the Breast Cancer Information Core Database located on the World Wide Web at http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/
- 34 Gingeras TR, Ghandour G, Wang E, *et al.* Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Res* 1998;8:435-48.
- 35 Hacia JG, Makalowski W, Edgemon K, *et al.* Evolutionary sequence comparisons using high-density oligonucleotide arrays. *Nat Genet* 1998;18:155-8.
- 36 Hacia JG, Woski SA, Fidanza J, *et al.* Enhanced high density oligonucleotide array-based sequence analysis using modified nucleoside triphosphates. *Nucleic Acids Res* 1998;26:4975-82.
- 37 Hacia JG, Fan JB, Ryder O, *et al.* Determination of ancestral alleles for human single nucleotide polymorphisms using high density oligonucleotide arrays. *Nat Genet* 1999;22:164-7.
- 38 Collins FS, Guyer MS, Chakravarti A. Variations on a theme: cataloging human DNA sequence variation. *Science* 1997;278:1580-81.
- 39 Kruglyak L. The use of a genetic map of biallelic markers in linkage studies. *Nat Genet* 1997;17:21-4.
- 40 Chakravarti A. Population genetics - making sense out of sequence. *Nat Genet Suppl* 1999;21:56-60.
- 41 Shumaker JM, Metspalu A, Caskey CT. Mutation detection by solid phase primer extension. *Hum Mutat* 1996;7:346-54.
- 42 Pastinen T, Kurg A, Metspalu A, Peltonen L, Syvanen AC. Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res* 1997;7:606-14.
- 43 Syvanen AC. From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum Mutat* 1999;13:1-10.
- 44 Nikiforov TT, Rendle RB, Goelet P, *et al.* Genetic bit analysis: a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res* 1994;22:4167-75.
- 45 Head SR, Rogers YH, Parikh K, *et al.* Nested genetic bit analysis (N-GBA) for mutation detection in the p53 tumor suppressor gene. *Nucleic Acids Res* 1997;25:5065-71.
- 46 Brice A. Unstable mutations and neurodegenerative disorders. *J Neurol* 1998;245:505-10.
- 47 Simone NL, Bonner RF, Gillespie JW, Emmert-Buck MR, Liotta LA. Laser-capture microdissection: opening the microscopic frontier to molecular analysis. *Trends Genet* 1998;14:272-6.
- 48 Broude NE, Sano T, Smith CL, Cantor CR. Enhanced DNA sequencing by hybridization. *Proc Natl Acad Sci USA* 1994;91:3072-6.
- 49 Hoheisel JD. Sequence-independent and linear variation of oligonucleotide DNA binding stabilities. *Nucleic Acids Res* 1996;24:430-2.
- 50 Hacia JG, Edgemon K, Sun B, Stern D, Fodor SP, Collins FS. Two color hybridization analysis using high density oligonucleotide arrays and energy transfer dyes. *Nucleic Acids Res* 1998;26:3865-6.
- 51 Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC, Ward DC. Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat Genet* 1998;19:225-32.
- 52 Nguyen HK, Bonfils E, Auffray P, *et al.* The stability of duplexes involving AT and/or G4EtC base pairs is not dependent on their AT/G4EtC ratio content. Implication for DNA sequencing by hybridization. *Nucleic Acids Res* 1998;26:4249-58.
- 53 Nguyen HK, Fournier O, Asseline U, Dupret D, Thuong NT. Smoothing of the thermal stability of DNA duplexes by using modified nucleosides and chaotropic agents. *Nucleic Acids Res* 1999;27:1492-8.
- 54 McGall G, Labadie J, Brock P, Wallraff G, Nguyen T, Hinsberg W. Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. *Proc Natl Acad Sci USA* 1996;93:13555-60.
- 55 Weiler J, Gausepohl H, Hauser N, Jensen ON, Hoheisel JD. Hybridisation based DNA screening on peptide nucleic acid (PNA) oligomer arrays. *Nucleic Acids Res* 1997;25:2792-99.