DNA sequencing by hybridization – a megasequencing method and a diagnostic tool?

Andrei D. Mirzabekov

The length of DNA from different sources that has been sequenced has already exceeded 1×10^8 bases, and this figure is continuing to grow exponentially. Even so, the very large quantity of DNA that remains to be sequenced and the impetus generated by the worldwide Human Genome Project has increased the need for the development of megasequencing procedures. Sequencing by hybridization (SbH) shows promise as an approach for developing such a method. The technique involves hybridization of the DNA of unknown sequence with an enormous set of short oligonucleotides; identification and analysis of the overlapping set of oligomers that form perfect duplexes with the DNA of interest permits reconstruction of the target-DNA sequence. Preliminary experiments have already demonstrated the feasibility of incorporating this approach into large-scale sequencing projects, and processes have been developed both for manufacturing the sequencing microchips that incorporate the immobilized oligonucleotides, and for detecting hybridization of the target DNA to these microchips.

The goal of sequencing the human genome is only part of a much greater international collaborative effort to sequence the genomes of many different human individuals, as well as those of various animals, plants and microorganisms. The major bottleneck for genome analysis is the sequencing technology itself. The efficiency, cost and reliability of the two classical sequencing methods – those of Sanger and Maxam–Gilbert[1] – are insufficient to cope with the requirements of large-scale genomic sequencing. Therefore, a number of innovative adaptations, as well as modifications to automate sequencing, have been developed that significantly improve these two methods.

A radically new approach towards DNA sequencing is sequencing by hybridization (SbH), an approach that was suggested simultaneously by four groups: two in the UK[2,3], one in the former Yugoslavia[4] and one in Russia[5,6]. SbH is based on sequence-specific DNA hybridization to a large set of oligonucleotides of specified length. The unknown DNA sequence can be determined by identifying overlapping blocks of oligonucleotides that form perfect duplexes with the target DNA sequence. Figure 1 shows a model SbH experiment for a 17-base target DNA fragment (I) that was hybridized with eight complementary, overlapping 8-mers to form perfect duplexes[7]. The overlap of the adjacent 8-mers by seven or fewer bases enables reconstruction of the complementary DNA sequence. Hybridization of a similar DNA fragment (II), which contains a single base substitution of C for T (at position 8) showed a significant decrease in the hybridization efficiency of the six duplexes spanning this base and containing internal mismatches, but a rather smaller decrease in the stability for duplexes with terminal G–T mismatches. Hybridization of fragment II with the ninth, complementary, 8-mer identifies the base change at position 8. While this example illustrates the principle of SbH, 'real' sequencing of unknown target DNAs would, of course, necessitate hybridization with the complete set of all possible 8-mers (i.e. $4^8 = 65536$).

A. Mirzabekov is at the Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov Street 32, Moscow 117984, Russia.

Matrices of immobilized DNA or oligonucleotides

There are two alternative formats for SbH. In the first (Fig. 1b), the sequencing is performed by consecutive hybridization of different oligomers with the target-DNA samples immobilized on a filter matrix[4,8]. The technologies and robots required for matrix manufacture and hybridization procedures are well

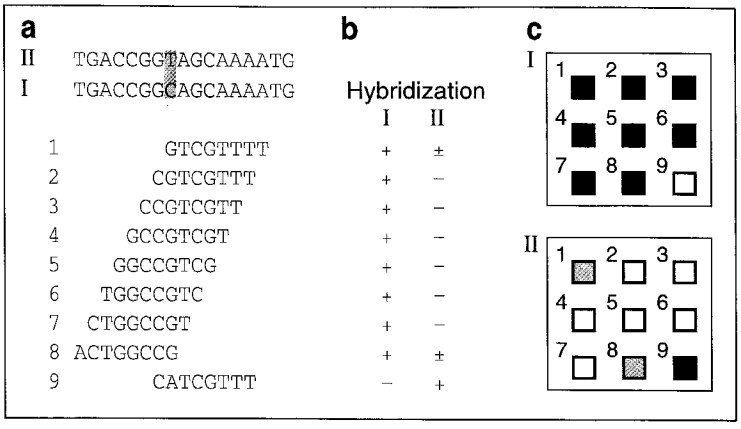

Figure 1

Sequencing by hybridization. (**a**) Two 17-mers, differing from each other by a single base change of C for T in position 8 (highlighted), were hybridized to nine 8-mers to make up perfect duplexes or duplexes with single base-pair mismatches. (**b**) Hybridization of labeled 8-mers with immobilized 17-mers in solution. (**c**) 8-mers were immobilized on a two-dimensional matrix and hybridized with labeled 17-mers. Hybridization is marked as positive for perfect duplexes (+, black), negative for duplexes with internal and some terminal mismatches (–, white) or intermediate for duplexes with G–T terminal mismatches (+–, striped) in (**b**) and (**c**), respectively.

developed and are already widely used for dot-blot hybridization[9].

Sequencing by hybridization using an oligonucleotide matrix (SHOM) (Fig. 1c) is being developed as an alternative approach[2,5,6]. Short oligonucleotides cannot bind sufficiently tightly to a filter (such as Hybond™); therefore they are immobilized as a defined two-dimensional array by covalent binding of their activated 3' or 5' terminal linkers to either a glass[2] or gel support[7]. Using a gel support enables larger numbers of oligomers to be immobilized than either a Hybond™ filter or glass[7,10]. In our laboratory, we have developed methods for producing a sequencing micromatrix comprising gel elements containing immobilized oligomers with dimensions as small as 30×30 μm (Refs 7,10; G. M. Ershov *et al.*, unpublished). A sequencing microchip made in this way and incorporating a complete set of 65536 8-mers would have dimensions of ~3×3 cm. The size of the chip element can be reduced even further by making additional improvements to this approach, or by using other methods such as printing technology, which has a resolution as high as 10 μm (Ref. 11).

Independent synthesis of the thousands of different oligomers needed for SbH is rather expensive. A method for parallel synthesis of oligomers on a glass surface as a specified two-dimensional array has therefore been developed[2,11,12]. If the coupling of every subsequent nucleotide to a precursor is carried out simultaneously for the whole array, then the synthesis of the complete 8-mer matrix requires only 4×8 = 32 rounds of coupling. However, since it is virtually impossible to purify and test the matrix-synthesized oligomers, it is essential that the efficiency and reliability of every coupling step is as high as possible.

A matrix incorporating 256 purine 8-mers has been manufactured by solid-phase oligonucleotide synthesis[11]. Spectacular opportunities are available through the use of addressable photo-activated chemistry in the parallel synthesis of oligonucleotides directly on a glass surface[13]. This technology may provide the means for effective industrial manufacture of highly complex sequencing microchips. If large numbers of small-dimension oligonucleotide matrices are to be produced industrially, and quality control for every immobilized oligomer is essential, then using robots to apply presynthesized oligomers to the matrix (Ref. 7; G. M. Ershov *et al.*, unpublished) has some advantages. This method is also applicable to the immobilization of a much wider variety of oligonucleotide analogs, as well as other compounds, such as proteins, antibodies, antigens and low-molecular-weight ligands, on matrices. Such microchips could find application in many other areas besides large-scale sequencing; for example, in diagnostics or molecular screening.

Choosing between immobilizing the target DNA or using oligomers on matrices in SbH needs to consider whether or not the number of DNA probes exceeds that of the oligonucleotides required for hybridization. SHOM has the advantage for routine sequencing of DNA, provided that the sequencing microchips can be produced economically on a commercial scale.

Detection of hybridization

Quantitative measurement of hybridization is usually performed using radioactively or fluorescently labeled probes. The use of a Phosphor-Imager (Molecular Dynamics) and image-analysis software simplifies the quantitative analysis of radioactivity[8,9,11]. Radiolabeling gives high-sensitivity detection, although it suffers from rather low resolution (~300 μm), and is also less convenient for measuring the kinetics of molecular interactions.

A confocal laser fluorescence microscope equipped with a sensitive detector, such as a photomultiplier tube (PMT), provides high resolution and sensitivity for detecting hybridization of fluorescently labeled probes. This arrangement can be used to scan hybridization on microchips manufactured by photolithography in a matter of minutes[14]. Less sensitive, but much faster (down to 50 msec) quantitative parallel measurements of dissociation curves for microchip-hybridized DNA are performed using a fluorescence microscope fitted with a charge-coupled device (CCD)-camera and a computer. A diagram of a SHOM sequencer is shown in Fig. 2 (Ref. 7; G. M. Ershov *et al.*, unpublished) and prototype equipment has been constructed and tested in our laboratory.

Hybridization with sequencing matrices

In early experiments, synthetic oligomers, 11–19 bases in length, were hybridized with filter-immobilized DNA. These experiments demonstrated that single, internal base-pair mismatches can be identified easily due to the resulting destabilization of the duplexes[15,16]. Further preliminary studies also indicated the feasibility of applying SbH to the detection of sequence differences.

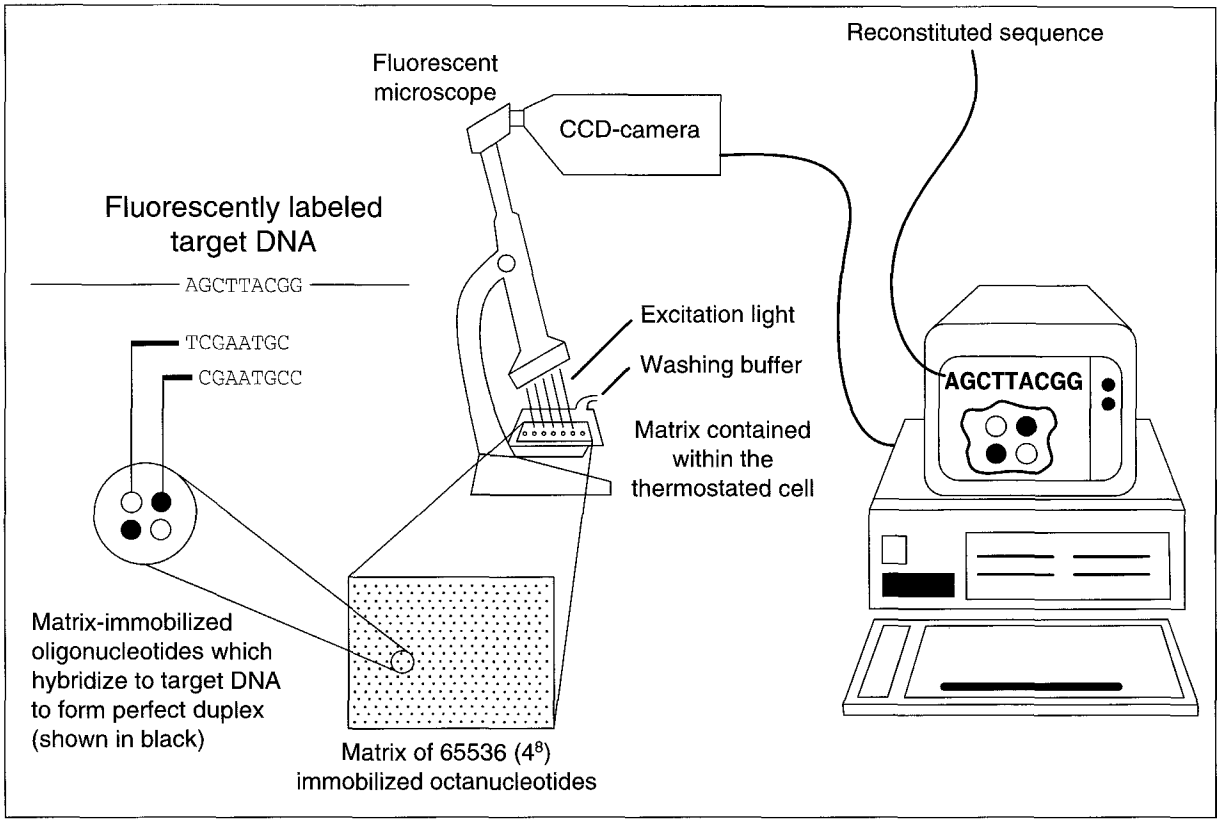

Figure 2
A schematic representation of a sequenator for sequencing by hybridization with an oligonucleotide matrix (SHOM).

In our laboratory, the temperature-dependent dissociation curves for duplexes were measured for four 17-mers, differing by only a single base-change, and hybridized with the nine gel-immobilized 8-mers (Fig. 1). Perfect duplexes formed with the octamers were easily distinguishable from duplexes containing internal mismatches. This enabled reconstruction of the sequence of one hybridized 17-mer and identification of the position of some base changes in others[7].

The sequences of two oligopyrimidine 24-mer target DNAs were reconstructed on the basis of their hybridization to a matrix of all 256 possible oligopurine 8-mers synthesized on a glass surface[11]. In another test experiment, three homologous, 116-base regions of unknown sequence from primate T-cell receptor β loci were unambiguously sequenced by hybridization of filter-immobilized DNA with a preselected set of 156 8-mers and 9-mers[8]. Generating a preselected set of probes requires some knowledge of the sequence, and probes that will potentially hybridize are chosen. The likelihood of any of 156 random 8-mers hybridizing to a 116-base target would, of course, be very small.

These, and other studies reveal some key features of SbH. Oligonucleotides as short as 8-, 7- and even 6-mers can be used for efficient hybridization, although the stability of the duplexes formed decreases as the length of the oligomer decreases[17]. Perfect duplexes formed upon hybridization of short oligomers with DNA can easily be distinguished from duplexes containing base-pair mismatches, due to the destabilizing effect of the mismatches. Terminal mismatches, in particular G–T and G–A, are more difficult to detect. More efficient discrimination of mismatches is achieved with shorter duplexes. However, their lower stability may cause a decrease in the hybridization signal relative to the background.

The ability to distinguish perfect duplexes from imperfect ones by hybridization may, depending on the conditions used, either be temperature-independent or may be less effective at higher temperatures; as a general rule, it is therefore recommended that screening be carried out at lower temperatures[17]. However, washing to remove imperfect duplexes may take hours at low temperature, although it can be carried out in minutes or seconds if washing to dissociate preformed duplexes is carried out at ~5°C below the 'melting', or apparent dissociation temperature of the perfect duplex.

When the hybridization conditions have been optimized for the particular DNA being sequenced, hybridization signals from mismatched duplexes are at least several times lower than those from perfect ones. However, longer target DNAs may form many mismatched duplexes with a particular oligonucleotide and, if there are many such false-positive signals from imperfect hybrids, they may exceed the signal from a perfect duplex. Thus, the tolerance of error in the signal is dependent on the length of the target DNA, i.e. one can be more confident of the sequence obtained for shorter target DNAs than for longer DNAs, for a given error rate. According to computer-simulation experiments, this limits the size of the target DNA that can be analysed for the presence of complementary

nucleotides to 500, 2000 and 8000 bases for 6-, 7- and 8-mer matrices, respectively[18].

SbH with an 8-mer matrix allows every base in the DNA to be sequenced eight times: each base will be hybridized with eight overlapping 8-mers within its immediate sequence context. This additional information enables up to 10% false-positive or false-negative hybridization signals to be tolerated by the system, without compromising sequence accuracy. These erroneous signals may include poor resolution of terminal mismatches, multiple internal mismatches, or faults in experimental design[8].

Significant differences in the stability of A–T and G–C base pairs cause variation in the optimal hybridization temperature, or the time required to dissociate A–T- and G–C-rich duplexes. SbH experimental procedures can be simplified greatly if the hybridization can be carried out under the same conditions for all oligomers. Several approaches have been used to attempt to equalize the temperature stability of duplexes with different base compositions. Hybridization of filter-immobilized DNA with 9-mers for less-stable A–T-rich oligomers and with 8-mers for more stable G–C-rich oligomers, partly compensates for the difference[8]. The addition of tetramethylammonium chloride (TMAC) to the hybridization solution increases the relative A–T base-pair stability and equalizes, to some extent, the efficiency of hybridization of glass-immobilized A–T and G–C-rich oligomers[19]. However, it seems to be ineffective in smoothing base differences for hybridization of filter-immobilized DNA with short oligomers[18]. Betaine can also be used to help eliminate the base-composition dependence of DNA melting[20].

A specific feature of gel-immobilized oligomers was used in our laboratory to achieve equalization. The apparent dissociation temperature of duplexes rises by ~5°C upon increasing the concentration of gel-immobilized oligomers threefold[6,7]. Thus, a 'normalized' sequencing matrix was prepared by adjusting the concentration of the oligonucleotides, using a higher concentration of A–T- and a lower concentration of G–C-rich immobilized oligomers to equalize their apparent temperature stability[7]. A theoretical analysis of this effect has enabled the development of a method for predicting the normalized concentration of oligomers[21]. The theory assumes that diffusion of the dissociated DNA through the gel is retarded by encountering and reversibly associating many times with immobilized oligomers. This process is similar to that behavior observed in chromatographic processes and appears to sharpen the dissociation curves.

Single-stranded DNA forms hairpins, loops and other secondary and tertiary structure elements. These may interfere with the process of DNA hybridization with short oligomers[8,22]. Fragmentation of DNA into random small pieces of ~10–30 nucleotides in length that lack, or have reduced internal structure can be helpful. More ingenious approaches are based on increasing the stability of duplexes formed on sequencing chips over the stability of DNA internal structures[7].

Efficiency of SbH

Straightforward use of SbH is limited to sequences that lack three or more copies of repeats longer than $n-1$, where n is the size (in bases) of the oligonucleotides used in the hybridization. SbH cannot be applied to direct sequencing of simple-sequence DNA, such as $(G–T)_n$ repeats, polyA stretches, or satellite DNA. (Although SbH cannot be used for sequencing simple repeats, it can be used to determine what the repeat-unit length is in such simple sequences, provided that the repeat is less than the length of the oligomer probe; it can also be used to determine the number of repeats, provided that the hybridization is quantitative.) The sequence reconstruction becomes ambiguous if repeats are present, particularly if the repeats are interspersed among nonrepetitive sequences, as neither their order, nor the order of the nonrepetitive stretches of sequences can be determined. As is becoming increasingly obvious from genome projects, the DNA of many organisms is highly repetitive; human DNA is packed with repeats in exons, introns and noncoding regions. Analysis of such sequences by SbH may be facilitated by fragmentation of the target DNA into shorter pieces prior to hybridization experiments. The longer the DNA and the shorter the oligomers used, the more ambiguous the reconstruction[4,23,24].

The most straightforward way of improving the efficiency of SbH is to increase the matrix complexity. Sequencing a target DNA of completely unknown sequence neccssitates using the complete set of all possible oligomers of a particular length. Calculations show that using a set of all 7-mers ($4^7 = 16384$), or 8-mers (65536), or 9-mers (262144), target DNAs of ~ 100, 200 and 400 bases in length, respectively, may be sequenced. Thus, increasing oligomer size by just one base enables DNA of double the length to be sequenced[12,24], although the complexity of the matrix is increased fourfold.

The ambiguity that arises during sequence reconstruction, particularly of longer stretches of DNA, could be diminished if additional information, such as quantitative hybridization to determine the number of repeats, or measuring the total DNA length, is also taken into account. Measuring the distance between unique segments that flank repeats and simple-sequence DNA by gel analysis of restriction fragments or PCR products significantly decreases the ambiguity of reconstruction and, thus, widens the scope of SbH applications[25]. Interestingly, more information can be obtained using a complete set of 'gapped' oligomers of length $n+1$ or $n+2$, containing one or two undefined bases in their central positions, respectively, as compared with a set of the same size of n-length oligomers[3,24].

Recently, we suggested a method for increasing the efficiency of SHOM by multistep, 'continuous-stacking' hybridizations[6,7,10]. It has been shown that the duplex formed upon hybridization of DNA with, for example, a matrix-immobilized octamer and a nonimmobilized pentamer is fairly stable, despite the

presence of a broken phosphodiester bond. The stability is provided by continuously stacked bases in the perfect 13 bp duplex. The stability of the pentamer association in this complex decreases significantly if any mismatch, gap or loop is present. Stable and perfect 13 bp duplexes can be assembled by additional rounds of DNA hybridization on a matrix of 8-mers, in the presence of a preselected mixture of labeled pentamers. This procedure imitates SHOM with a matrix of 13-mers. Computer simulations demonstrated that five additional rounds of continuous-stacking hybridization of DNA on an 8-mer matrix with computer-selected mixture of 300–500 different, non-immobilized 5-mers could be used for sequencing target DNAs up to 4000 bases in length, and a matrix of 9-mers with 6-mers will sequence target DNA of ~16 000 bases[26,27]. Continuous-stacking hybridization has also been used recently to assemble primers from several 6-mers stacked on a template for DNA sequencing polymerization[28].

A theoretical approach for predicting the temperature stability of any perfect duplex and those with different base mismatches would be helpful in choosing the best conditions for SbH, since experimental determination of the stability of all potential duplexes would involve enormous effort. At present, only semiempirical rules are available for predicting hybridization in solution[29,30] or with gel-immobilized oligonucleotides[21].

SbH applications

The ultimate goal of SbH is to provide a means of rapid, simple, inexpensive, reliable and automated DNA sequencing. Experience with the technique to date shows that it is effective for sequencing short target-DNA fragments with a preselected set of oligomers. However, validation of the method for large-scale sequencing of DNA fragments of unknown sequence necessitates the use of a very large set of oligomers (if not the complete set of all possible oligomers) of defined length. In addition, technical improvements in matrix preparation, hybridization procedures, hybridization measurements, reliable methods for discriminating between perfect and mismatched duplexes, and development of appropriate instrumentation and software are also required. The feasibility of each step has already been demonstrated – the considerable task that lies ahead is the optimization and integration of all the individual steps into one practical sequencing technology.

One factor contributing to the promise that SbH holds is the experimental simplicity of the hybridization step: it can be automated with comparatively inexpensive equipment and can be performed in minutes or even seconds. The computer technology currently available is sufficiently powerful to enable the development of automated systems for analysing hybridization signals, for discriminating between mismatches and perfect duplexes, and for the sequence reconstruction. The basic technologies required to manufacture sequencing chips and microchips already exist[7,11–14], although it is likely that it will take a few more years to develop and automate SbH and SHOM. However, both methods can already be used reliably to detect mutations. Mutations, including single base-changes, have been detected reliably using a microchip incorporating immobilized oligomers (Refs 2,7,11,14; G. M. Ershov *et al.*, unpublished), or by hybridization of these oligomers with immobilized DNA[15,16]. It should not be difficult to fabricate diagnostic microchips in the near future to identify any specified point mutation. Such a chip would contain about $4 \times N$ oligomers, where N is the length of the region of the gene where the mutation being screened for is located, and 4 corresponds to any of the four possible bases.

The use of a partial set of oligomers of different lengths for sequence verification, or for sequencing when the sequence is partially known, has been investigated in some detail[9,31]. This approach can be applied to sequence comparison, to searches for homologies, to the identification of particular sequence motifs, such as consensus sequences, promoters, enhancers, and to the selection of DNA for complete sequencing by determining its gross oligomer composition. A matrix of predetermined sequences could be used to fractionate DNA or to test for the expression of specific genes in different cell types or cells grown under different conditions. For this purpose, matrices incorporating immobilized expressed-sequence tags (ESTs) would be particularly valuable.

Even while still in the early stages of development, SbH could be particularly useful for verifying sequence data obtained by the classical gel-based methods. Since SbH is based on a different principle, it might detect errors that would be missed by repeated classical sequencing; such checking could diminish considerably the number of mistakes entered in databases. SHOM might prove to be the method of choice for comparative sequencing, i.e. sequencing very similar sequences, differing by a few base changes, but where only one of these sequences is known. This aspect of genome analysis will increase when the genes or the whole genome of a particular organism has been sequenced, and there is the need to determine the sequence of the corresponding genes or genome structures of different individuals or organisms.

Conclusions

Will SbH eventually be developed as a mega-sequencing method? It appears that all the crucial components of this approach have already been tested in several laboratories, and the results inspire reasonable optimism. However, the technology is still far from being fully developed, and only crude estimates of the efficiency of SbH are possible at present. Let us consider SHOM, extended by continuous-stacking hybridization, on a gel microchip that could be applied to sequencing DNA of ~4000 bases. The incubation of a prehybridized microchip at an appropriate temperature to eliminate imperfect duplexes can be

carried out in a matter of minutes for several rounds of hybridization when the stability of G–C- and A–T-rich duplexes is equalized. Measurement and analysis of the hybridization signals can be automated. Similarly, preparation of the DNA samples for sequencing and their prehybridization with microchips can be achieved robotically. When the automation is optimized, it should be feasible to sequence at the rate of several megabases per 24-hour period.

The cost of a microchip will incorporate the cost of manufacturing. By analogy with electronic microchip production, the per-unit cost should not be excessive for large-scale production. Where the matrices are manufactured with presynthesized oligomers, the cost of the oligomers must also be considered. The cost of synthesizing the oligonucleotides for an 8-mer gel-microchip (requiring ~100 fmol of immobilized oligomers for each of its 65536 elements) should amount to just a few US dollars. The challenge will be in establishing manufacturing procedures able to produce, reproducibly, matrices incorporating the oligonucleotides synthesized to the degree of accuracy required.

SbH does not work with DNA containing many internal repeats and simple-sequence elements, and it can therefore only be applied to the direct sequencing of simple genomes of organisms such as viruses, bacteria, yeast or for unique regions of the genomes of higher organisms. However, for more complex genomes, SbH may be used to identify repeats and determine sequence polymorphisms. These problems associated with sequencing complex genomes could be solved by the use of SbH in conjunction with the gel-based sequencing methods or measuring the length between several unique segments of DNA sequenced by SbH.

SbH can already be applied for diagnostic purposes, and it shows promise as an approach for developing a sufficiently fast and comparatively inexpensive method for DNA megasequencing.

Acknowledgements

The author would like to thank Yu. Lysov, V. Florentiev and I. Ivanov for their helpful comments, and E. Meier and E. Novikova for critical reading of the manuscript. This work was supported by grant numbers 562 and 558 from the Russian Human Genome Project; grant N DE-FG02-93ER1538 from the US Department of Energy; and was also sponsored by Affymax Research Institute, CA, USA.

References

1 Hunkapiller, T., Kaiser, R. J., Koop, B. F. and Hood, L. (1991) *Science* 254, 59–67
2 Southern, E. M. (1988) International Patent Application *PCT GB 89/00460*
3 Bains, W. and Smith, G. C. (1988) *J. Theoret. Biol.* 135, 303–307
4 Drmanac, R., Labat, I., Brukner, I. and Crkvenjakov, R. (1989) *Genomics* 4, 114–128
5 Lysov, Yu. P., Khorlin, A. A., Khrapko, K. R., Shick, V. V., Florentiev, V. L. and Mirzabekov, A. D. (1988) *Proc. USSR Acad. Sci.* 303, 1508–1511
6 Khrapko, K. R., Lysov, Yu. P., Khorlin, A. A., Shick, V. V., Florentiev, V. L. and Mirzabekov, A. D. (1989) *FEBS Lett.* 256, 118–122
7 Khrapko, K. R., Lysov, Yu. P., Khorlin, A. A., Ivanov, I. B., Yershov, G. M., Vasilenko, S. K., Florentiev, V. L. and Mirzabekov, A. D. (1991) *DNA Sequence* 1, 375–388
8 Drmanac, R., Drmanac, S., Strezoska, Z., Paunesku, T., Labat, I., Zeremski, M. T., Snoddy, M. J., Funkhouser, W. K., Koop, B., Hood, L. and Crkvenjakov, R. (1993) *Science* 260, 1649–1652
9 Meier-Eccert, S., Mayer, E., Ahmadi, A., Curtis, J. and Lehrach, H. (1993) *Nature* 361, 375–376
10 Lysov, Yu. P., Khrapko, K. R., Chernyi, A. A., Khorlin, A. A., Florentiev, V. L. and Mirzabekov, A. D. (1991) in *Electrophoreses, Supercomputing and the Human Genome* (Cantor, C. and Lim, H. A., eds), pp. 157–163, World Scientific
11 Southern, E. M., Maskos, U. and Elder, J. K. (1992) *Genomics* 13, 1008–1017
12 Maskos, U. and Southern, E. M. (1992) *Nucleic Acids Res.* 20, 1697–1684
13 Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. and Solas, D. (1991) *Science* 251, 767–773
14 Fodor, S. P. A., Pava, R. P., Xuang, X. C., Pease, A. C., Holmes, C. P. and Adams, C. L. (1993) *Nature* 364, 555–556
15 Wallace, R. B., Schaffer, J., Murphy, R. F., Bonner, J., Hirose, T. and Itakura, K. (1979) *Nucleic Acids Res.* 6, 3543–3557
16 Ikuta, S., Takagi, K., Bruce, R. and Itakura, K. (1987) *Nucleic Acids Res.* 15, 797–810
17 Drmanac, R., Strezoska, Z., Labat, I., Drmanac, S. and Crkvenjakov, R. (1990) *DNA Cell Biol.* 9, 527–534
18 Strezosca, Z., Paunescu, T., Radoslavljevic, D., Labat, I., Drmanac, R. and Crkvenjakov, R. (1991) *Proc. Natl Acad. Sci. USA* 88, 10089–10093
19 Maskos, U. and Southern, E. M. (1992) *Nucleic Acids Res.* 20, 1675–1678
20 Rees, W. A., Yagen, T. D., Korte, J. and von Hippel, P. H. (1993) *Biochemistry* 32, 137–144
21 Livshits, M. A., Florentiev, V. L. and Mirzabekov, A. D. *J. Biomol. Struc. Dynamics* (in press)
22 Maskos, U. and Southern, E. M. (1993) *Nucleic Acids Res.* 21, 2267–2270
23 Pevzner, P. (1989) *J. Biomol. Struct. Dynamics* 7, 63–73
24 Pevzner, P. A., Lysov, Yu. P., Khrapko, K. R., Belyavsky, A. V., Florentiev, V. L. and Mirzabekov, A. D. (1991) *J. Biomol. Struct. Dynamics* 9, 399–410
25 Lysov, Yu. P., Mironov, A. A., Gnuchev, F. N., Chernyi, A. A. and Mirzabekov, A. D. *Mol. Biol.* (in press) (in Russian)
26 Lysov, Yu. P., Chernyi, A. A., Balaeff, A. A., Gnuchev, F. N., Beattie, K. L. and Mirzabekov, A. D. *Mol. Biol.* (in press) (in Russian)
27 Lysov, Yu. P., Chernyi, A. A., Balaeff, A. A., Beattie, K. L. and Mirzabekov, A. D. (1993) *Mol. Biol.* 27, 1126–1138 (in Russian)
28 Kieleczawa, K., Dunn, J. J. and Studier, F. W. (1992) *Science* 258, 1787–1792
29 Breslauer, K., Frank, R., Blocker, H. and Marky, L. (1986) *Proc. Natl Acad. Sci. USA* 83, 3746–3750
30 Wetmur, J. G. and Davidson, N. (1986) *J. Mol. Biol.* 31, 349–370
31 Drmanac, R., Lennon, G., Drmanac, S., Crkvenjakov, R. and Lerach, H. (1991) in *Electrophoreses, Supercomputing and the Human Genome* (Cantor, C. and Lim, H. A., eds), pp. 60–74, World Scientific