How Capillary Electrophoresis Sequenced the Human Genome**

Norman J. Dovichi* and Jianzhong Zhang

1. Introduction to DNA Sequencing

Now that the sequence of the human genome is essentially complete, it is appropriate to review how we have come to this point in the history of science. In 1953, Watson and Crick reported the structure of DNA, which confirmed the central role of DNA for storage of genetic information.[1] A few years later, the genetic foundation that describes the translation of DNA to proteins was fully known.[2]

The relationship between the genetic and protein sequences was a powerful motivation for the development of methods to determine DNA sequences. Not only would DNA sequences provide insight into mutation and evolution, but they could also provide a shortcut to the determination of protein sequences. Until 1976, the determination of DNA sequences involved laborious, classical chemical methods. Sanger et al.'s "plus and minus" method was the first step in sequencing chemistry, which allowed the determination of the 5375 base sequence of the bacteriophage ΦX174 in 1977.[3] Two much more elegant sequencing methods were reported in 1977, the dideoxy chain-terminating method of Sanger and co-workers (Sanger's method) and Maxim and Gilbert's chemical degradation method.[4, 5] Sanger and Gilbert shared the 1980 Nobel Prize in Chemistry for their development of analytical methods for DNA sequencing.

Biologists pursued DNA sequencing with great enthusiasm, and a cottage industry developed for the manual determination of the sequences of viruses, genes, and genetic markers. GenBank was started in 1981 as a public depository for DNA sequence information.[6, 7] Figure 1 presents the number of bases recorded in GenBank since its inception, along with several important milestones in sequencing history. From 1984 to 1999, the rate at which data were deposited into GenBank increased exponentially, with a doubling period of roughly two years.

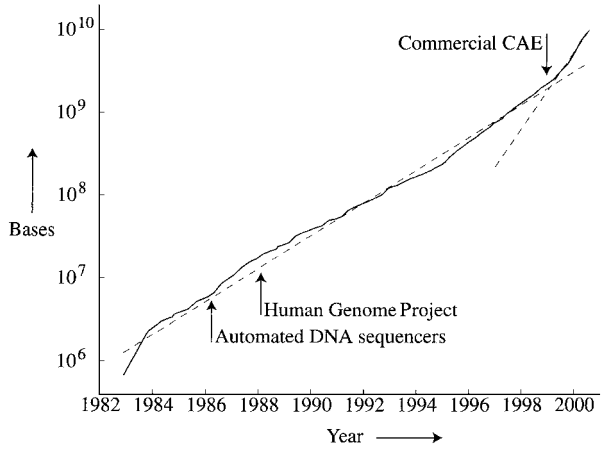

Figure 1. Number of bases of DNA sequences deposited in GenBank. The dashed lines are least-square fits to the logarithm of the linear data; the data increased exponentially with a two-year doubling rate until 1999. Since then, GenBank has increased in size with a six-month doubling rate.

Instruments for automated capillary-array electrophoresis (CAE) became available in 1999, which precipitated an explosive increase in the rate at which DNA sequences were reported in GenBank, reaching a doubling rate of six months during the year 2000. These instruments were used in large factories with robotic sample handling and huge bioinformatics capabilities. The industrialization of DNA sequencing has had a breathtaking impact on biology. More DNA sequences were deposited in GenBank over the past two weeks than reported during the whole of 1995. When the GenBank data are combined with sequences held in commercial databases, it is likely that the amount of DNA sequences known at the end of 2000 will be ten times larger than that known at the start of the year.

The development of capillary-array DNA sequencers and their installation in sequencing factories has changed biology forever. In this Essay, we trace the history of the sequencer that has made this revolution possible.

2. Sanger's DNA Sequencing

Most researchers used Maxim and Gilbert's sequencing method during the early 1980s, primarily because reagents

[*] Prof. N. J. Dovichi,[+] Dr. J. Zhang
Department of Chemistry
University of Alberta
Edmonton, Alberta T6G 2G2 (Canada)
Fax: (+1) 780-492-8231
E-mail: norm.dovichi@ualberta.ca

[+] After January 1, 2001:
Department of Chemistry
University of Washington
Seattle, WA 98195-1700 (USA)
Fax: (+1) 206-685-8665
E-mail: dovichi@chem.washington.edu

[**] This Essay is based on a lecture given at the Analytica 2000 conference in Munich (Germany) on the occasion of the Heinrich-Emanuel-Merck Prize presentation.

  1433-7851/00/3924-4463 $ 17.50+.50/0

were not widely available for Sanger's method. However, the commercialization of dideoxynucleotides and the development of automated DNA synthesizers made Sanger's sequencing method practical and virtually all DNA sequencing relies on this method today.

In the vocabulary of DNA sequencing, the piece of DNA with the unknown sequence is called the template. To sequence a single-stranded template using Sanger's method, an approximately 20-base primer is synthesized which is complementary to a specific region of the template. The primer is mixed with the template, where it hybridizes forming a short region of double-stranded DNA that is recognized as an initiation site for chain extension by a DNA polymerase. This enzyme extends the primer sequence, adding new nucleotide triphosphates to the 3′ end of the primer and synthesizing a DNA fragment that is complementary to the template.

The key to Sanger's sequencing reaction is the addition of a small amount of dideoxyadenosine triphosphate to the reaction mixture. This reagent has a triphosphate group and it can be incorporated into the growing strand. However, the reagent is missing a hydroxyl group at the 3′ position of the sugar; no nucleotides can be added to the dideoxynucleotide and the growing strand is terminated upon its incorporation. On completion of the reaction, a population of products is generated. All products start at the primer, are complementary to the template sequence, and terminate upon incorporation of a dideoxyadenosine (ddA), but they are of different lengths depending upon the position where the ddA was incorporated. These reaction products are called a sequencing ladder. Similar sequencing ladders are generated from the template with dideoxycytidine, dideoxyguanosine, and dideoxythymidine. This synthesis remains essentially unchanged from Sanger's original work, although improved enzymes are used today.

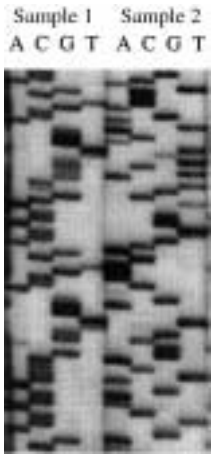

Figure 2. Autoradiogram generated from a DNA sequencing gel. Two samples were subjected to the Sanger sequencing reaction. Each sample generated four sequencing ladders that were separated on adjacent lanes of the polyacrylamide gel slab.

In Sanger's original method, the sequencing ladders were separated, based on their size, in four adjacent lanes of a polyacrylamide gel. Radioactive phosphorus was used in the reaction mixtures so the end base of the sequence was radiolabeled. The fragments were detected by autoradiography, where film was placed over the gel and the radioactive decay of the labeled DNA created dark bands on the film (Figure 2). These autoradiograms were inspected by a skilled technician who would determine the DNA sequence based on the pattern of bands and manually enter the data into a computer. Roughly 12 hours were required for the electrophoresis, another 12 hours for the autoradiography, and many hours to read the sequence, which was an error-prone and mind-numbing exercise.

Figure 2 shows the image from a small portion of a sequencing gel. Two samples are analyzed in adjacent lanes in this figure; the original gel analyzed 12 samples. Roughly 50 bases worth of sequence are shown in the figure; the entire gel would generate a few hundred bases of sequence for each sample. The smallest fragments migrated fastest and appear at the bottom of the gel. The sequence of sample 1 begins with CGACCACCACCCCGAGGTT.

2.1. Automated DNA Sequencing

In 1986, Smith, Hood, and co-workers reported an important improvement in Sanger's method whereby the radioactive labels, autoradiographic detection, and manual data interpretation were replaced with fluorescent labels, laser-induced fluorescence detection, and computer-based data analysis.[8] In their method, a primer was synthesized, split into four batches, and each batch was labeled with one of four different fluorescent dyes. Each dye-labeled primer was used in a sequencing reaction with one of the dideoxynucleotides. These reaction products were pooled and analyzed in a single lane of a sequencing gel. A four-color laser-induced fluorescence detector monitored the DNA fragments as they migrated to the bottom of the gel; the fluorescence signature was used to identify the terminal nucleotide. This technology was commercialized by Applied Biosystems in 1987.

Figure 3 presents raw data generated by this sequencing protocol; the fluorescence from each DNA fragment generates a peak with a particular spectral signature. The spectrum from each dye is rather broad and generates a signal in several spectral channels. A simple matrix inversion procedure is used to convert the spectral data into dye-space data, and this dye information is then used to determine the identity of each sequencing fragment.

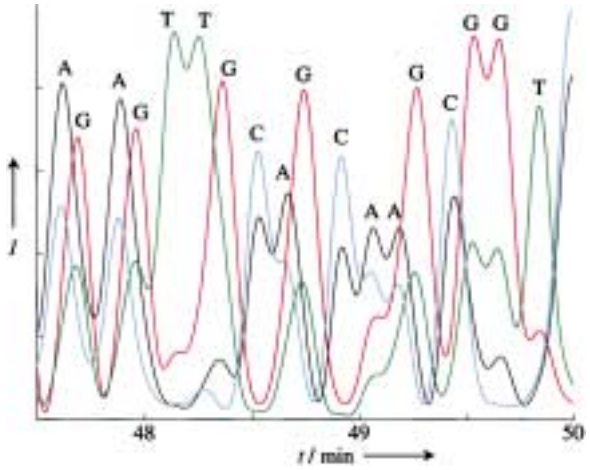

Figure 3. Four-color DNA sequencing data for bases 376–393 of M13mp18. Four different fluorescent labels were used to tag the primer, which was then was used in the Sanger sequencing reaction. After the four sequencing ladders were synthesized, the reaction products were pooled and separated in a single electrophoresis capillary. A laser-induced fluorescence detector monitored the fluorescence signal in four spectral channels, which are plotted as the four-color traces. I = fluorescence.

  1433-7851/00/3924-4464 $ 17.50+.50/0

Prober et al. at DuPont developed an automated sequencer that used more elegant chemistry based on a set of fluorescently labeled dideoxynucleotides;[9] these labeled dideoxynucleotides were combined and used in a single sequencing reaction, which simplifies the synthesis of the DNA sequencing ladders. The fragments were separated in a single lane of sequencing gel and identified by the fluorescence signature of the dye-labeled dideoxynucleotide. DuPont commercialized this technology for a brief period but ultimately sold the license to Applied Biosystems.

Ansorge et al. developed a simpler system based on a single fluorescent label and separation of fragments in four lanes of an automated sequencer. This technology was commercialized by Pharmacia.[10]

The introduction of these automated DNA sequencers was an important event in biotechnology. A group of visionary biologists realized that this new technology would allow the completion of very large DNA sequencing efforts. The Human Genome Project was initiated in 1988,[11] two years after Smith, Hood et al.'s report of an automated DNA sequencer. The audacious goal of the genome project was to determine the primary structure of a set of molecules with a combined molecular weight of 2×10^{12} Da and to complete this task by 2005 with the commercial, automated sequencers. The National Institutes of Health of the United States of America has provided the lion's share of public funding for the Human Genome Project, which has grown into an international effort of global proportions. The Wellcome Foundation in Great Britain has been notable in their support of the Human Genome Project. Since 1999, Celera has become an important private-sector company active in very large-scale sequencing efforts.

3. High-Throughput DNA Sequencing

Conventional DNA separations are performed using glass plates that are coated with an approximately 200 μm thick layer of cross-linked polyacrylamide gel. The preparation of these gels is tedious and manually intensive. The plates are separated, scraped to clean off old polymer, and reassembled with appropriate spacers. An acrylamide monomer mixture is prepared and poured between the plates to polymerize; this mixture contains an aqueous buffer and 8 M urea. Once polymerization is complete, the slab of gel is placed in the electrophoresis apparatus, sample is pipetted into wells formed at the top of the gel, electrodes are attached, and the DNA sequencing fragments are separated. Between 16 and 96 samples are separated simultaneously on a single gel.

The thickness of the slab of gel limits the speed with which DNA can be separated. The poor heat dissipation of the thick gel limits the electric field used for DNA separation. At high fields, the temperature rise due to Joule heating causes the sequencing gel to distort and ultimately boil, which destroys the separation. The modest electric field used in conventional slab-gel electrophoresis results in relatively tedious separation of the sequencing fragments.

3.1. Capillary Electrophoresis for DNA Sequencing

The foundation for capillary electrophoresis was laid by Hjerten in the 1960s,[12] Everaerts and co-workers in the late 1970s,[13] and Jorgenson and Lukacs in the early 1980s.[14, 15] Jorgenson and Lukacs' work was particularly important because they treated electrophoresis as an instrumental technique and introduced capillary electrophoresis to the analytical chemistry community.

Capillary electrophoresis is performed with capillaries that are typically 50 μm in inner diameter, 150 to 350 μm in outer diameter, and about 30 cm in length. The capillaries are made from high-purity fused silica which is similar to that used in the manufacture of fiber optics.

Instrumentation for capillary electrophoresis is quite simple (Figure 4). The capillary holds a sieving medium, which allows separation of the DNA fragments based on their size. The sample is injected into the capillary by dipping the capillary and an electrode into a sample solution and briefly applying electric current through the capillary so that DNA fragments migrate onto the tip of the capillary. Once injection is complete, the sample is replaced with running buffer and electric field is reapplied to drive the DNA fragments through the capillary. A laser-induced fluorescence detector near the end of the capillary records the fluorescence signal in four different spectral channels to resolve the fluorescence signature from the four dyes.

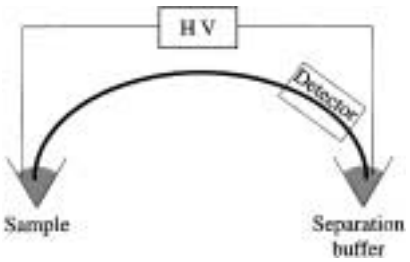

Figure 4. Single-capillary electrophoresis instrument. A fused-silica capillary is used for the separation. One end of the capillary is dipped into the sample or buffer reservoir, while the other end passes through a detector before being placed in a buffer-filled reservoir. High voltage (HV) is applied through platinum electrodes. The high voltage end of the capillary is held in a safety interlock equipped chamber.

The small inner diameter of the capillary reduces Joule heating to negligible levels and allows the use of extremely high electric fields for very rapid separations of DNA sequencing fragments.[16] However, the phenomenon of biased reptation with orientation is observed at high electric field and degrades the separation of longer sequencing fragments.[17] The electric fields employed in capillary electrophoresis represent a two-fold improvement over the best slab-gel instruments.

Despite the ability to use high electric field for electrophoresis, the most important property of capillaries is that they are highly flexible and easily incorporated into an automated instrument. It is the ease of automation, rather than fundamental thermal properties, that has resulted in the successful development of advanced sequencing instruments.

  1433-7851/00/3924-4465 $ 17.50+.50/0

The technician time necessary to operate a slab-gel sequencer was the major bottleneck in large-scale DNA sequencing. Now, a single technician can service dozens of instruments, by just loading microtiter plates containing the sequencing fragments in each machine and occasionally replenishing reagents.

The first capillary electrophoresis DNA sequencer that used flexible, fused-silica capillaries, which ultimately became the standard for DNA sequencing, was reported in 1990 by Swerdlow when he was a graduate student in Gesteland's group at the University of Utah.[18]

That year, Swerdlow moved to our laboratory at the University of Alberta as a visiting student. During his time in our lab, we published the first description of the use of a sheath-flow cuvette detector for DNA sequencing by capillary electrophoresis (Figure 5).[19] This cuvette, modified from a flow cytometer, proves to be an extraordinarily efficient fluorescence detector. The cuvette has an approximately 175 μm square flow chamber that passes through an approximately 1 mm square piece of quartz. The separation capillary fits snugly in the flow chamber. Buffer is pumped at a very low rate in the space between the capillary and the cuvette window, creating a stream that draws DNA sequencing fragments as a thin filament in the center of the flow chamber. Due to the excellent optical quality of the cuvette windows, light scatter is virtually eliminated as a source of background signal in fluorescence measurements. When used with carefully designed collection optics and a low power helium – neon excitation laser, the cuvette has achieved detection of single molecules of β-phycoerythrin after separation by capillary electrophoresis.[20]

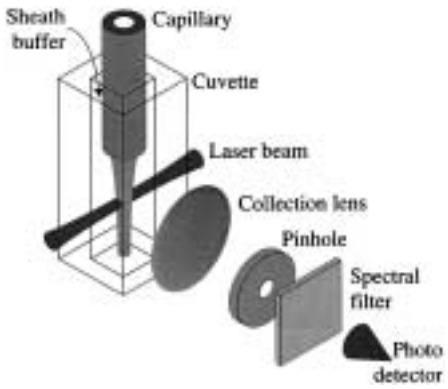

Figure 5. Single-capillary sheath-flow cuvette. A fused-silica capillary is placed inside of a square quartz flow chamber with high optical quality windows. Sheath fluid is pumped in the space between the capillary and the window, and this fluid draws the analyte into a thin stream in the center of the flow chamber. A laser beam is focused beneath the capillary tip on the sample stream. Fluorescence is collected by a lens, spectrally filtered, imaged onto an aperture, and detected with a photomultiplier tube.

In 1993, Karger and co-workers at Northeastern University realized that cross-linked polyacrylamide was an inappropriate matrix for DNA sequencing by capillary electrophoresis because the capillary must be discarded after each use. Instead, he reported the use of non-cross-linked polyacrylamide for the separation.[21] This matrix can be pumped from the capillary after each use and replaced with fresh material without having to replace the capillary itself.

The non-cross-linked material has relatively low viscosity, so that it may be pumped into the capillary under modest pressures. However, there was concern that the viscosity of the matrix would drop at higher temperatures and that electroosmosis would pump the matrix from the capillary. The use of high temperatures for separation is important because DNA denatures at elevated temperature. Intrastrand hydrogen bonding forms hairpin structures in the DNA that give rise to anomalous migration patterns during electrophoresis that degrade sequencing accuracy.

We studied the effect of temperature on the separation of DNA sequencing fragments in cross-linked and non-cross-linked polyacrylamide under high electric fields in capillary electrophoresis.[22, 23] In 1995, we reported the first separation of DNA sequencing fragments at 60 °C with non-cross-linked polymer capillary electrophoresis.[24] This temperature was sufficiently high to eliminate sequencing artifacts, thereby generating a high-quality sequence. We also demonstrated that the sequencing read length increased with temperature, which opened the door to practical, accurate long read length separation of DNA sequencing fragments.

3.2. Capillary Array Electrophoresis

While the performance of capillary electrophoresis is superior to that of conventional slab-gel electrophoresis, a single-capillary instrument does not offer significant advantages over a multi-lane slab-gel system. By separating many samples simultaneously, the slab-gel system produces much more data than is possible from a single-capillary instrument.

It is necessary to operate an array of capillaries to obtain throughput that is comparable with conventional electrophoresis systems. Zagurski and McCormick presented the first report of capillary array electrophoresis (CAE) for DNA sequencing in 1990.[25] That instrument's detector scanned across the capillaries, recording the fluorescence signal sequentially from each capillary. This general approach of scanning a detector across the capillary array has been duplicated by other groups. However, this scanning method suffers from limited duty cycle: while one capillary is interrogated by the detector, DNA is also migrating in the other capillaries so that most DNA is not detected. Limited duty cycle can be severe. For example, each capillary in a 96 capillary array is probed for 1 % of the time; 99 % of the DNA is lost without detection.[26]

The more successful capillary-array sequencers rely on continuous monitoring of each capillary. Two instruments are of particular interest, one from Kambara and co-workers at Hitachi and the other from this group.[27–32] Both instruments rely on a sheath-flow cuvette to simultaneously monitor fluorescence from a linear array of capillaries (Figure 6). Sheath fluid is pumped through the interstitial space between the capillaries and entrains the DNA sequencing fragments as discrete streams, with one stream per capillary. A laser beam is focused into the cuvette and skims beneath the capillary tips (Figure 4), thus exciting fluorescence from all of the capil-

  1433-7851/00/3924-4466 $ 17.50+.50/0

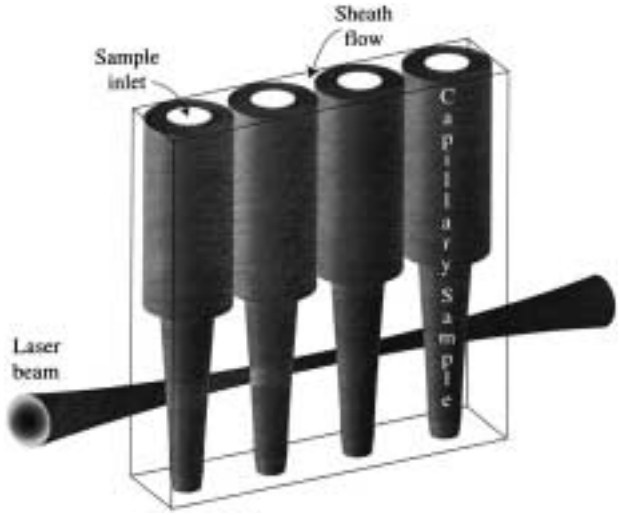

Figure 6. Capillary linear-array sheath-flow cuvette. A linear array of fused-silica capillaries is placed inside of a rectangular glass flow chamber. Sheath fluid draws the analyte into a thin streams in the center of the flow chamber, with one stream produced down-stream from each capillary. A laser beam is focused beneath the capillary tips on the sample streams. Fluorescence is collected by a lens, spectrally filtered, and detected with either an array of photodiodes or a CCD camera.

laries simultaneously. When observed from the front, the detector generates a set of fluorescent spots, with one spot beneath each capillary and with the spots separated by the capillary outer diameter. This detection cuvette is efficient in its use of laser light. As the sheath fluid is highly transparent, there is no disruption of the laser beam as it traverses the cuvette; one low-power laser beam is used to simultaneously illuminate samples migrating in all capillaries. An optical system images the fluorescence from each capillary onto a CCD camera or an array of photodiodes. In this way, fluorescence from each capillary is monitored continuously and the instrument's duty cycle approaches 100%.

The sheath-flow cuvette CAE instrument has been licensed to Applied Biosystems and is commercialized as the "Model 3700 DNA sequencer". The instrument operates with 96 capillaries for separation and has eight spare capillaries. A robot rapidly transfers samples from a 96-well microtiter plate to an injection block, where samples are loaded onto the capillaries. Sample loading, sieving-matrix replacement, electrophoresis, data collection, and base calling are all automated, so that the instrument can run for 24 hours without intervention.

It has been claimed that these CAE instruments represent a tenfold increase in sequencing throughput compared with automated slab-gel sequencers.[33] This claim is probably an underestimate. The fourfold increase in the sequencing rate over the past year (Figure 1) is due to the output of a couple of thousand automated sequencers that replaced tens of thousands of automated slab-gel electrophoresis instruments.

Conclusions

Advances in DNA sequencing include the development of large-scale shotgun sequencing,[34] highly automated DNA sequencers, and powerful data-processing algorithms.[35] This Essay focuses on the latest generation of highly automated DNA sequencers, which is based on CAE.

The Human Genome Project is the most ambitious and important effort in the history of biology. It has provided the complete genetic blueprint for human life, and will provide important insights into human health and development. The vast majority of this sequence was determined with high-throughput DNA sequencers based on CAE with a sheath-flow detector.

We are at the threshold of a new era in the biological sciences, which is based on genomic information that was unimaginable two decades ago. In 1980, the sequencing of a small gene could be the foundation of a complete PhD thesis. Due to the development of powerful analytical instrumentation, it is now possible to complete the sequence of a prokaryote in a few weeks and the sequence of a vertebrate in less than a year. This revolution in the biological sciences is possible only because of advances in analytical instrumentation for DNA sequencing.

Acknowledgements

It may be instructive to document the funding history for this work. Between 1989 and 1992, we submitted five proposals to local, national, and international funding agencies for the support of this project; all were rejected. Our first successful proposal for DNA sequencing was supported by the Department of Energy (USA) Human Genome Program and we held this funding from 1992–1998. Additional funding was received from SCIEX, a division of MDS, Inc. from 1995–1998. That funding was matched by the Natural Sciences and Engineering Research Council for the same period. Funding was also received from the Canadian Bacterial Diseases Network and the Canadian Genetic Diseases Network starting in 1994. The CBDN funding was recently terminated due to lack of commercial success. We gratefully acknowledge our sole source of funding for our current work in DNA analysis, the Canadian Genetic Diseases Network.

[1] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 964–967.
[2] "The Genetic Code": *Cold Spring Harbor Symp. Quant. Biol.* **1966**, *31*.
[3] F. Sanger, G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, C. A. Fiddes, C. A. Hutchison, P. M. Slocombe, M. Smith, *Nature* **1977**, 264, 687–695.
[4] F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. USA* **1997**, 74, 5463–5467.
[5] A. M. Maxam, W. Gilbert, *Proc. Natl. Acad. Sci. USA* **1997**, *74*, 560–564.
[6] D. A. Benson, M. S. Boguski, D. J. Lipman, J. Ostell, B. F. Ouellette, B. A. Rapp, D. L. Wheeler, *Nucleic Acids Res.* **1999**, *27*, 12–17.
[7] http://www.ncbi.nlm.nih.gov.
[8] L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connell, C. Heiner, S. B. H. Kent, L. E. Hood, *Nature* **1986**, *321*, 674–679.
[9] J. M. Prober, G. L. Trainor, R. J. Dam, F. W. Hobbs, C. W. Robertson, R. J. Zagursky, A. J. Cocuzza, M. A. Jensen, K. Baumeister, *Science* **1987**, *238*, 336–341.
[10] W. Ansorge, B. S. Sproat, J. Stegemann, C. Schwager, *J. Biochem. Biophys. Methods* **1986**, *13*, 315–317.
[11] F. Collins, D. Galas, *Science* **1993**, *262*, 43–46.

  1433-7851/00/3924-4467 $ 17.50+.50/0

[12] S. Hjerten, *Chromatogr. Rev.* **1967**, *9*, 122 – 219.
[13] F. E. P. Mikkers, F. M. Everaerts, T. P. E. Verheggen, *J. Chromatogr.* **1979**, *169*, 11 – 20.
[14] J. W. Jorgenson, K. D. Lukacs, *Science* **1981**, *222*, 266 – 272.
[15] J. W. Jorgenson, K. D. Lukacs, *Anal. Chem.* **1981**, *53*, 1298 – 1302.
[16] M. J. Rocheleau, N. J. Dovichi, *J. Microcolumn Sep.* **1992**, *4*, 449 – 453.
[17] J. L. Viovy, T. Duke, *Electrophoresis* **1993**, *14*, 322 – 329.
[18] H. Swerdlow, R. Gesteland, *Nucleic Acids Res.* **1990**, 18, 1415 – 141.
[19] H. Swerdlow, S. Wu, H. Harke, N. J. Dovichi, *J. Chromatogr.* **1990**, *516*, 61 – 67.
[20] D. Y. Chen, N. J. Dovichi, *Anal. Chem.* **1996**, *68*, 690 – 696.
[21] M. C. Ruiz-Martinez, J. Berka, A. Belenkii, F. Foret, A. W. Miller, B. L. Karger, *Anal. Chem*. **1993**, *65*, 2851 – 2858.
[22] H. Lu, E. Arriaga, D. Y. Chen, D. Figeys, N. J. Dovichi, *J. Chromatogr.* **1994**, *680*, 503 – 510.
[23] Y. Fang, J. Z. Zhang, J. Y. Hou, H. Lu, N. J. Dovichi, *Electrophoresis* **1996**, *17*, 1436 – 1442.
[24] J. Z. Zhang, Y. Fang, J. Y. Hou, H. J. Ren, R. Jiang, P. Roos, N. J. Dovichi, *Anal. Chem.* **1995**, *67*, 4589 – 4593.
[25] R. J. Zagursky, R. M. McCormick, *BioTechniques* **1990**, *9*, 74 – 79.
[26] X. C. Huang, M. A. Quesada, R. A. Mathies, *Anal. Chem.* **1992**, *64*, 2149 – 2154.
[27] N. J. Dovichi, J. Z. Zhang, US-A 5,439,578, **1995**.
[28] S. Takahashi, H. Kambara, US-A 5,529,679, **1996**.
[29] J. Z. Zhang, PhD Thesis, University of Alberta, **1994**.
[30] H. Kambara, S. Takahashi, *Nature* **1993**, *361*, 565 – 566.
[31] J. Z. Zhang, K. O. Voss, D. F. Shaw, K. P. Roos, D. F. Lewis, J. Yan, R. Jiang, H. Ren, J. Y. Hou, Y. Fang, X. Puyang, H. Ahmadzadeh, N. J. Dovichi, *Nucleic Acids Res.* **1999**, *27*, E36.
[32] H. J. Crabtree, S. J. Bay, D. F. Lewis, L. D. Coulson, G. Fitzpatrick, D. J. Harrison, S. L. Delinger, J. Z. Zhang, N. J. Dovichi, *Electrophoresis* **2000**, *21*, 1329 – 1335.
[33] R. Preston, *The New Yorker* **2000**, June 12, 66 – 83.
[34] J. C. Venter, M. D. Adams, G. G. Sutton, A. R. Kerlavage, H. O. Smith, M. Hunkapiller, *Science* **1998**, *280*, 1540 – 1542.
[35] E. W. Myers, G. G. Sutton, A. L. Delcher, I. M. Dew, D. P. Fasulo, M. J. Flanigan, S. A. Kravitz, C. M. Mobarry, K. H. J. Reinert, K. A. Remington, E. L. Anson, R. A. Bolanos, H. H. Chou, C. M. Jordan, A. L. Halpern, S. Lonardi, E. M. Beasley, R. C. Brandon, L. Chen, P. J. Dunn, Z. Lai, Y. Liang, D. R. Nusskern, M. Zhan, Q. Zhang, X. Zheng, G. M. Rubin, M. D. Adams, J. C. Venter, *Science* **2000**, *287*, 2196 – 2204.

  1433-7851/00/3924-4468 $ 17.50+.50/0