TECHVIEW: SEQUENCING

Gels and Genomes

Jane Rogers

The Sixth International Conference on Automation in Mapping and DNA Sequencing (AMS) (held 1 to 3 September, 1999, at the Sanger Centre, the Wellcome Trust Genome Campus, Hinxton, UK; www.sanger.ac.uk/Info/Events/ams99) gave genome scientists, engineers, and biotechnologists a focused forum to discuss automation of the mining of genomic information. The need for automation became apparent when the feasibility of large-scale sequencing was demonstrated by the production of the first megabase of genomic sequence from *Caenorhabditis elegans*. This achievement prompted A. Watson, T. Hawkins, and R. Wilson to set up the inaugural AMS conference at the Sanger Centre in 1993.

The requirement for inexpensive, high-throughput, highly accurate DNA sequencing has risen dramatically since that time as the value of sequence information has become fully recognized by biologists. During the last 3 years, deposits of finished genomic sequence in the public databases [GenBank, EMBL (European Molecular Biology Laboratory), DDBJ (DNA Data Bank of Japan)] have risen from 270 Mb to 800 Mb of sequence, encompassing organisms from bacteria to man. Through advances in automation, the rate of sequence deposition for the human genome has trebled this year, as reviewed by F. Collins at the meeting.

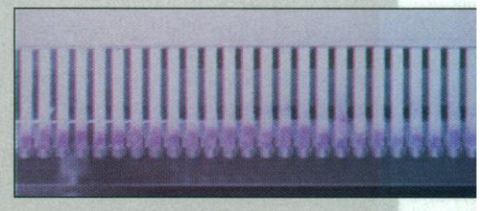
The technological aspects of automated sequencing are as topical now as in 1993, driven by the increased demand for de novo genomic sequencing and by the increasing numbers of resequencing applications that are emerging. For example, sequencing selected regions of the genome in a large number of individuals to identify polymorphisms, which should enable identification of genetic traits associated with human disease, is seen as an important derivative of the human genome sequencing project. Several large programs are now funded to discover single nucleotide polymorphisms (SNPs) in the human genome. The SNP Consortium (comprising 10 pharmaceutical companies, the Wellcome Trust, the Genome Sequencing Center, St. Louis, the Sanger Centre, the Whitehead Genome Center, and Cold Spring Harbor Laboratory) will make the first wave of SNPs publicly available before the end of 1999.

The author is at The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK. E-mail: jrh@sanger.ac.uk

New instrumentation for sequence generation continues to be developed, both in research institutes and by commercial organizations. The MegaBace (Amersham Pharmacia Biotech) and Perkin Elmer 3700 capillary sequencers now dominate human sequence production in the major genome centers, but it is important to remember that they have been in this position for less than a year. Improvements are still under development that will enhance their capacity (through reducing run times and increasing read lengths) and make them easier to use in a high-throughput environment (by reducing operator interaction times). Alternative capillary sequencers containing 96, 384, or 1024 capillaries are being developed in a number of laboratories but are not yet ready for use in a production environment. To compete with the performance of the capillary sequencers, slab-gel sequencing systems need to incorporate simultaneous running of at least 96 lanes/gel, automated loading, and sample retracking. The option of obtaining long read lengths is also desirable, as is flexibility with respect to sequencing chemistry. Two slab-gel systems that use different approaches to meet these demands were presented. The ARAKIS sequencer, developed with European Union funding by W. Ansorge at EMBL, now uses five-color fluorescence laser detection of multiplexed samples (labeled template-specifically with different dyes). Automated loading of up to 480 samples per gel is achieved using porous combs (see box), and runs can be completed twice daily. The BaseStation (MJ Research) is being developed to simultaneously run 192 samples (labeled with four-color fluorescent dyes) per gel, the samples being loaded directly from a 384-well plate. Using ultra-thin (75 μm) gels, 850 bases can be read in 4.5 hours.

Looking to the future, there is considerable optimism about miniaturization of sequencing platforms. One of the most attractive features of moving from high-throughput sequencers to smaller, faster DNA separation systems is the opportunity to reduce sequencing costs. D. Ehrlich from the Whitehead Institute reported on the development of a microdevice that is capable of reading 500 bases in 27 min in 11.5-cm channels filled with 3% linear polyacrylamide. ACLARA Biosciences has achieved separations of up to 700 bases in 25 min in 18-cm microfabricated channels in glass and plastic. Three groups presented sequencing reaction systems that are capable of handling volumes of only 500 nl. A. Marziali and co-workers (University of British Columbia and Stanford University) reported on the development of an instrument designed to handle 576 samples/run. The ACAPELLA-1K system, which handles sub-microliter reactions in glass capillaries, was developed at the University of Washington, Seattle, is capable of handling 5000 samples in 8 hours. Molecular Dynamics demonstrated the successful separation of 500-nl sequencing reactions on their MegaBace 1000 capillary sequencer and the interface of this technology with a prototype automated DNA analysis microchip (ADAM).

In addition to sequencing, genotyping applications were reported for all of the sequencing platforms. Competing with the sequencing platforms are fluorescence polarization assays, INVADER assays, and microsphere-based microsequencing flow cytometry assays. There will be keen competition in a field in which the throughput could be tens or hundreds of millions of assays per year per laboratory, and the aim of the investigator is reduction of the cost per reaction to 0.1 cents/sample. Clearly, data tracking, data management, and database development will also be key to the success of these projects.



CREDIT: HOLGER ERFLE

A simple technique for automating sample loading of sequencing gels (which involves cutting a comb template from a porous membrane) has been developed at EMBL [H. Erfle *et al.*, *Nucleic Acids Res.* **25**, 2229 (1997); www.mwgdna.com/products/pcr/robots/index.htm]. The comb is lowered onto sample wells that are machined into a plexiglass block at the same distance apart as the comb teeth. Each well is filled with sample (either manually or automatically), and as the teeth touch the wells the samples are drawn up into the teeth by capillary attraction (see photo). Alternatively, the samples can be robotically spotted onto the comb's teeth directly, and the comb can be stored until use. The comb is placed between the glass plates of the gel apparatus above the flat surface of the polymerized gel, and the samples are driven from the comb into the gel by electrophoresis.