DNA analysis is becoming more automated

More than 20 years have passed since GATC Biotech AG launched its first DNA sequencing device. Based on voltage, the device would have taken 1875 years to sequence three billion base pairs. Today's ultramodern sequencers can decipher the same number of base pairs in 10 days. GATC Biotech AG is already on the lookout for next-generation sequencing technologies in the form of real-time single-molecule sequencing. In an interview with BIOPRO, Thomas Pohl, CTO of GATC Biotech AG, explains why the company is continuously expanding its existing equipment with the latest global technologies.

Always on the go: Thomas Pohl, CTO of GATC Biotech AG, is constantly on the lookout for new sequencing technologies. © GATC Biotech

Mr. Pohl, your laboratories are equipped with ultramodern, top global sequencing technologies with a total sequencing capacity of more than 20 terabases per year. The systems' performance is continuously improving and the time-to-customer gets ever shorter. What has made this progress possible?

The main reasons are that most of the new sequencers are equipped with high-resolution cameras

for detection, the sequencing chemistry has changed and the digital capture of the signals has become more efficient. The enormous increase in performance and efficiency have come about thanks to a number of different developments in a broad range of areas.

Your equipment pool involves a broad range of technologies, including Sanger, pyrosequencing and sequencing by synthesis. Can you briefly explain the main differences between these technologies?

The main differences are related to the detection of the nucleotides that are incorporated into the DNA. The Sanger and Illumina sequencing methods involve detection of fluorescence-labelled molecules with lasers. The Sanger method uses capillary electrophoresis and Illumina uses a flow cell. Pyrosequencing relies on the detection of pyrophosphate release on nucleotide incorporation. A secondary enzyme reaction generates visible light, which is detected by a camera and analysed with a special software. Sequencing by synthesis involves the detection of nucleotides while they are being incorporated into the DNA template. Another method is 'sequencing by ligation' in which oligonucleotides hybridise to a target DNA sequence.

One of your most modern technologies is the Illumina HiSeq2000 sequencer, which has a capacity of 200 gigabytes per run. What exactly does this mean?

This means that the HiSeq2000 sequencer can sequence two human genomes in a single run. The human genome consists of more than three gigabases. We can sequence two human genomes at 30-fold coverage with the HiSeq2000 sequencer. The 30-fold coverage is necessary for the subsequent high-quality analysis of, for example, the exchange of nucleotides in a genome.

GATC Biotech is constantly expanding its pool of equipment with new devices and technologies. GATC staff also undergo intensive training to stay up-to-date with the latest technologies. © GATC Biotech

Can you explain how you decide whether a large project requires the use of a Roche GS FLX, a Genome Analyzer II or an Illumina HiSeq2000?

For the de novo sequencing of larger genomes, i.e. genomes that are larger than those of bacteria, the GS FLX technology is the method of choice because it enables a reliable bioinformatic analysis of the data obtained. The Illumina technology is mainly used in re-sequencing projects when a sequence is already available to be compared with the new data, which might contain mutations. In simple terms, the choice of sequencing technology depends on whether a new genome is being sequenced for the first time, or whether a sequence is being compared with an already existing sequence.

You are due to implement a new generation of technology platforms, the "single molecule sequencing" technology, into your lab routines next year. How does this technology work and what advantages does it have?

With these platforms, sequencing and detection are done in real-time. PacBio RS also uses fluorescence-labelled nucleotides. However, these are bound to the phosphate groups rather than to the nucleotides. This phosphate group is removed, leaving behind a native DNA strand. All this happens in so-called zero-mode waveguides (ZMW), a kind of hole in a metal foil on a glass plate. Each ZMW has a polymerase enzyme attached to the bottom, which is where the sequencing reaction takes place. This enables us to observe DNA synthesis molecule by molecule, as it happens in real time. The great advantage is that PacBio does not require the routine PCR amplification that many second-generation sequencers need. In addition, PacBio also offers strobe sequencing which generates pulses of sequence data from a molecule separated by arbitrary stretches of unsequenced DNA during which sequencing is put on hold. Strobe sequencing thus

allows researchers to sequence multiple portions of a fragment and renders the creation of libraries containing differently sized inserts unnecessary. In addition, we are also testing other technologies that will enter the market at some point. These include Ion Torrent's PostLight sequencing technology, which measures the pH change caused by the release of a positively-charged hydrogen ion when nucleotides rather than fluorescence-labelled nucleotides are added to a DNA strand. Another major advantage is of course the low cost of the system and its short run times.

Why do you choose to expand your equipment with new technologies? Is it because projects are getting more and more tricky and client requirements more demanding? Or is it more a matter of saving resources?

Sequencing hardware is developing very fast indeed. Research institutions, to name just one example, are not able to purchase state-of-the-art devices whenever a new one is launched, for financial reasons if nothing else. As the leading sequencing services provider in Europe, our objective is to offer universities access to ultramodern technologies shortly after market introduction. Of course, we also have to follow the general trend, i.e. higher throughput and lower sequencing costs. We are mainly interested in devices that help us reduce the spend on chemicals. We change our existing sequencers with new sequencer generations every two to three years and invest around 400,000 to 500,000 euros per device. It goes without saying that, as a service provider, we need to make enough money during the same two to three years in order to be able to purchase this equipment.

The ultramodern Illumina HiSeq2000 can sequence around 200 gigabytes, representing two human genomes, per run. © GATC Biotech

Do new and more effective technologies require more intensive staff training, as the equipment becomes more complex?

Each new development brings with it its own particular challenges. Modern technologies increasingly rely on information technology. This means that we require greater computing capacities as well as bioinformatics, mathematics and statistics professionals in our team. Our personnel structure has changed a great deal. Before, we were able to analyse DNA sequences with standard notebooks; now we require computer clusters.

Can you tell me anything about any future innovations GATA Biotech has in mind?

There are still a number of microorganisms and infectious pathogens in hospitals that are not yet sequenced and we want to work with our research partners to close this gap. We are already visiting companies that are working on technologies that will be launched in about two to three years' time. There are around five to ten companies worldwide working on the development of new sequencing methods.

Thanks to ultramodern hardware you are already able to sequence eukaryotic genomes in a high-throughput manner. Can you tell me more about these projects?

We are part of the International Cancer Genome Consortium (ICGC) whose goal is to sequence tens of thousands human genomes with the aim of building up comprehensive information about tumour types that are of clinical and societal importance around the globe. We will have sequenced more than 100 human genomes by the end of 2010. In addition, we also focus on the de novo sequencing of big eukaryotes, i.e. of eukaryotes whose genome has not previously been sequenced. This involves crops as well as organisms that have properties of great scientific interest.

Contact:

Elke Decker
Director Strategic Marketing & Corporate Communication
GATC Biotech AG
Jakob-Stadler-Platz 7
D-78467 Konstanz
Phone: (0) 7531-816030
Fax: (0) 7531-816081
E-mail: e.decker(at)gatc-biotech.com

Article

18-Oct-2010
Michael Statnik
BioLAGO

The article is part of the following dossiers

New machines for the life sciences