

GENOTYPING TECHNOLOGY: SNP-ING OUT INFORMATION

Researchers from academia and industry continue to make new technologies and tools to learn more from single nucleotide polymorphisms (SNPs). As expected, some of the advances come from higher density microarrays. Nevertheless, ancillary technologies, such as microfluidics, deliver additional advances. Moreover, researchers are also exploring variations beyond SNPs.
by Mike May

Despite ongoing discussions of the data overload, researchers still want more, especially from SNPs. “We really need to have high throughput techniques,” says Lise Lotte Hansen, associate professor of human genetics at the **University of Aarhus** in Denmark. Beyond just gathering more genotyping data, however, Hansen wants data that include more content. In some cases, any content would help. As Hansen says, “You look through databases and there is no information on some SNPs.”

In fact, some of the current data turn out to be wrong. “Sometimes, the SNP is not really located where the database places it,” Hansen says. “Also, the frequencies for the various genotypes can simply be wrong.” This arises from either methodological problems or because genotype frequencies vary considerably between populations.

To get better information on SNPs, scientists need more data, such as that gathered by the International HapMap Project (www.hapmap.org). This organization focuses on making a catalog of the human haplotype map, or HapMap, which is basically the variation in human genotypes. To obtain a representative sampling of human variation, the HapMap project looks at SNPs from 270 people from four populations: European, Han Chinese, Japanese, and Nigerian.

It will take even more information on SNPs to correlate them with diseases. Scientists need to look through tens of thousands of SNPs to find the one or two disease-related ones. So researchers need faster and more-accurate approaches to genotyping.

Expanding the View

Although scientists need more SNP data, part of today’s challenge in genotyping ironically comes from simply having too much to consider. “There are 10 million or so SNPs in public databases,” says Carsten Rosenow, senior marketing manager of DNA analysis products at **llumina** in San Diego, California. “Researchers don’t have the array density—or the money—to look at all 10 million.” Still, companies want to give scientists more SNPs on a chip—but no one can jam every SNP on one microarray. Instead, scientists at Illumina look for the SNPs that make what Rosenow calls “the best proxies for the others.”

Illumina’s tools depend on the company’s BeadChip design, which relies on beads of silica covered with DNA probes. “With enhancements to the beads and the slides,” says Rosenow, “we were able to make even higher density products.” Specifically, Illumina recently released its Infinium High-Density Human1M-Duo and Human610-Quad BeadChips, which run two [continued >](#)



“We really need to have high throughput techniques. You look through databases and there is no information on some SNPs.”

Look for these Upcoming Articles

RNAi — June 6

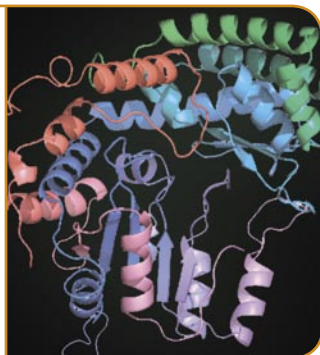
Proteomics 2 — August 1

Genomics 2 — September 19

Inclusion of companies in this article does not indicate endorsement by either AAAS or Science, nor is it meant to imply that their products or services are superior to those of other companies.

Genomics

“It is very clear that not only SNPs, but the combination of all known types of variants need to be discovered and typed—fast and cost efficiently.”



and four samples per chip, respectively. “These include up to 2.4 million markers on each chip,” says Rosenow. About 2.3 million of the markers are for typing SNPs and copy-number variation. The approximately 100,000 other markers improve the analysis of copy-number variation, because they increase the uniformity of probe distribution, reducing the number of large gaps without any probe coverage in the genome.

Rosenow sees these products used in many ways, such as studying the causes of complex diseases. When using Illumina chips to look for SNP-related diseases, though, Rosenow says that it’s not likely that the SNP causing the disease is on the chip. “But once you know that a SNP on the chip shows an association with the disease,” he says, “then you know where to go on a hunt.”

Other companies also look for more flexible ways to study SNPs. For example, the GenomeLab SNPStream Genotyping System from **Beckman Coulter** in Fullerton, California, works with nucleic acid samples in 384-well plates. Each well can contain 12 or 48 SNPs. “We can analyze 4,000 to 3.2 million SNP genotypes per day,” says Rene M. Oda, senior scientist at Beckman Coulter. “It’s very scalable.”

Moreover, this instrument does not limit scientists to a list of SNPs. Instead, says Oda, “We give them the design tools, and they can design any type of SNP or insertion or deletion.” This includes nucleic acids from humans or other mammals as well as bacteria, plants, or viruses.

The SNP analysis performed by this instrument can be used in many ways. For example, **DNAPrint Genomics** in Sarasota, Florida, uses the GenomeLab SNPStream to provide law enforcement agencies with ancestry information about suspects and victims, based on samples taken from crime scenes. In addition, **MMI Genomics** in Davis, California, markets a SNP-based kit to consumers that reports the genetic makeup of a dog from among 38 breeds, providing insight into behavioral characteristics and predisposition to genetic-based diseases that afflict some breeds.

Moreover, scientists are searching for signs of human health risks in other features of chromosomes. “Somatic mutations, for example, cannot be seen through genotyping,” says James Wang, president of **TrimGen** in Sparks, Maryland. His company, though, uses a form of real-time polymerase chain reaction, called eQ-PCR, which uses a proprietary probe design and chemistry that the company says enhances the detection of the fluorescent signal. According to Wang, eQ-PCR can find somatic mutations at levels of just 1 percent. In other words, TrimGen’s technology could find a mutation composed

of one base in 100. “For cancers, that could mean earlier detection,” Wang says. “Our eQ-PCR is sensitive and accurate. So you could use this to predict risks based on somatic mutations.”

This technology can also be used to detect a virus in a sample. Using just 1 milliliter of blood, TrimGen’s eQ-PCR can detect as few as 10 copies of a virus. “Most technologies can only get down to 50–100 copies,” says Wang.

TrimGen already puts this technology to many uses. For example, the company is running a leukemia project. “In Europe,” says Wang, “physicians already use genetic testing to select the treatment for leukemia and to predict relapse.” Consequently, TrimGen developed leukemia panels for its eQ-PCR. “This tells you which patients are likely to relapse fast and which are likely to be stable,” says Wang, “and it has already been proven that there is a gene related to this.” The eQ-PCR test results, Wang says, will provide supportive genetic evidence and valuable information to transform traditional medicine into personalized medicine.

Controlling the Flow

Making SNP technology work depends in many ways on getting the right sample materials to the right detectors. **Fluidigm** in South San Francisco meets this challenge with integrated fluidic circuits (IFCs). “It’s a chip that incorporates a network of elements: inlets, chambers, channels, and valves,” explains Michael Y. Lucero, Fluidigm’s executive vice president of sales and marketing. “It’s like an electrical circuit, except that instead of channeling electrons, IFCs manipulate fluids. The IFCs include the logic that controls the flow with 10,000 or more valves.”

Fluidigm recently introduced its BioMark system, which uses IFCs for various applications, including real-time PCR and genotyping analysis. “With IFCs,” says Lucero, “we can load many samples and test them against many different SNP assays in the same run.” The current version can run 48 samples against 48 assays—for a total of 2,304 reactions in one run. In the near future, says Lucero, the company will increase the throughput to 96 x 96. “That will be 9,216 parallel PCR reactions for genotyping,” he says.

The BioMark system runs its 2,304 PCR reactions in about 75 minutes, which Lucero says is the same time as is required for running PCR using conventional 384-well plates. Moreover, the system lowers running costs because IFCs require far less sample, reagents, and pipetting steps to set up reactions, according to Lucero. “These factors also make IFCs ideal for analyses of many single-cell samples as each cell can be efficiently assayed for 48 genes at once,” he says.

Science Applications International Corporation (SAIC)—a contractor for the US National Cancer Institute in Frederick, Maryland—recently installed a BioMark system at its core genotyping facility (CGF) to use genotyping to track down cancer susceptibility genes. With the BioMark, scientists at the CGF will be able to use TaqMan assays that they’ve already developed, including more than 5,000 SNP assays to look for potentially cancer-related genotypes.

The TaqMan assays, which were developed by Applied Biosystems in Foster City, California, are used by many scientists for PCR-based research. The TaqMan assays continue to expand to new instruments,

including being used on the OpenArray Platform from **BioTrove** in Woburn, Massachusetts. “Researchers using this integrated platform can experience an end-to-end genotyping workflow of less than five hours, enabling them to analyze thousands of samples in days, in contrast to weeks on alternative genotyping platforms,” says Phoebe A. White, senior director genotyping applications at **Applied Biosystems**. She adds, “The platform offers a flexible range of assay configurations, allowing researchers to interrogate as few as 16 and as many as 3,072 SNPs within one array across hundreds to thousands of samples.” Moreover, a customer can pick the SNPs from Applied Biosystems’ 4.5 million TaqMan SNP Genotyping Assays, Custom TaqMan SNP Genotyping Assays, or TaqMan DME Genotyping Assays.

Beyond focusing TaqMan assays on cancer, IFCs can also tackle environmental questions. Alaska fisheries, for example, are using Fluidigm’s IFCs to genotype salmon populations in the ocean. “A panel of SNP assays,” says Lucero, “is used to determine which stream head the salmon will return to for spawning, allowing controlled harvesting.”

Targeting Mutations

Instead of screening an ocean of possibilities, scientists sometimes prefer to focus genotyping on specific chromosomal regions. According to Rob van Miltenburg, director, global marketing, **Roche Applied Science** in Penzberg, Germany, “With the development of high-resolution melting, or HRM, Roche is now offering a method that can be used to scan for mutations in selected gene regions.” Roche put this technology in its LightCycler 480 real-time PCR. Running this HRM method is easy and cost effective, according to van Miltenburg. “In addition,” he says, “mutation scanning by HRM can be performed in high throughput, only requiring approximately 90 minutes to analyze a 96-well or 384-well plate.”

The LightScanner made by **Idaho Technology** in Salt Lake City, Utah, uses a similar technology. “Using a new double-stranded DNA binding dye that is included in the PCR reaction, it can monitor the

thermal denaturation—the melting—of a sample,” explains Rachel Jones, marketing manager for life science at Idaho Technology. “It can detect very subtle differences in DNA melting behavior and identify sequence variants or genotype samples based on the unique melting profile.” It works with both 96- and 384-well plates.

The latest application using the LightScanner can genotype a small amplicon generated from primers placed immediately adjacent to a SNP. However, if SNPs of interest melt at very similar temperatures, they can be difficult to discriminate. “Our Master Mix,” says Jones, “includes normalizing, temperature-calibrator probes that allow more accurate differentiation of the melting temps of SNPs whose T_m differs by less than 0.5 degree Celsius.”

Moreover, melting a sample is nondestructive. As a result, a researcher could go back to a plate and collect samples for further processing. “Lots of customers identify novel SNPs by using sequencing as a second step,” says Jones, “but by using the LightScanner as a prescreen, they can reduce the amount of sequencing by more than 90 percent in most cases.” This technology currently is most widely used for gene scanning, such as looking at entire exons for rare or novel mutations. For products that are 400 bases in length or less, says Jones, the LightScanner’s sensitivity and specificity are 95–99 percent.

In addition to exploring specific regions or small amplicons from a chromosome, researchers want to look across wider areas and for a variety of differences in genotypes. “With the Roche/454 Genome Sequencer FLX System,” says Marcus Droege, director, global marketing, Roche Applied Science, “it is possible to discover and analyze the whole spectrum of different types of mutations, from a SNP to smaller insertions and deletions to large structural variations of the genome.” As a result, this next-generation sequencer can be used to determine both sequences and genotypes. Other sequencers can also be used this way. In the April 2007 *Journal of Clinical Microbiology*, for instance, Donald Murphy of the Institut National de Santé Publique du Québec, and his colleagues used Applied Biosystems’ Prism 3100 genetic analyzer to genotype variants of the hepatitis C virus.

Whether genotyping in selected regions or across the genome, scientists need a range of tools, as do companies. For example, Roche’s acquisition of **NimbleGen** (Madison, Wisconsin) adds to the company’s variety of analytical capabilities, according to Janti Masani, vice president marketing at Nimblegen. He says, “Now, Roche will also be able to offer an optimized workflow composed of enrichment and sequencing of targeted gene regions or the complete human exome.”

Beyond the growing range of applications of genotyping, even its definition continues to evolve. “It is very clear that not only SNPs, but the combination of all known types of variants need to be discovered and typed—fast and cost efficiently,” says Droege. “This will lead to a wider definition of ‘genotyping,’ which is still often used with respect to SNP analysis only.” As scientists study ever-more variations, the very meaning of genotyping will expand over the next few years

Mike May is a publishing consultant for science and technology.

DOI: 10.1126/science.opms.p0800025

Featured Participants

Applied Biosystems
www.appliedbiosystems.com

Beckman Coulter
www.beckmancoulter.com

BioTrove
www.biotrove.com

DNAPrint Genomics
www.dnaint.com

Fluidigm
www.fluidigm.com

Idaho Technology
www.idahotech.com

Illumina
www.illumina.com

MMI Genomics
www.mmigenomics.com

NimbleGen
www.nimblegen.com

Roche Applied Science
www.roche-applied-science.com

**Science Applications
International Corporation**
www.saic.com

TrimGen
www.trimgen.com

University of Aarhus
www.au.dk/en