

The Digital PCR Revolution

“Like finding a needle in a haystack” is an overused expression, but when it comes to some biological scavenger hunts, it fits. Researchers studying rare variant biomarkers often find themselves on the lookout for faint genetic signals against an overwhelming background, sometimes as little as a single positive in 100,000 negatives or more. Such a situation cries out for polymerase chain reaction (PCR), a technique uniquely capable of capturing the proverbial needle. But standard PCR won’t do—it is a qualitative technique—and neither will quantitative real-time PCR (qPCR), which often lacks the necessary accuracy and sensitivity. These days, there’s a third and increasingly popular option: digital PCR (dPCR). By discretizing those 100,000 molecules in a large number of individual reactions, dPCR makes the rare positive surprisingly easy to find. **By Jeffrey M. Perkel**

Digital PCR essentially lifts weak signals out of the noise, and researchers today are using that strategy for everything from detecting copy number variations and circulating tumor DNA to HIV viral load monitoring.

Kenneth Kinzler, co-director of the **Ludwig Center at Johns Hopkins University**, who (with colleague Bert Vogelstein) first coined the term “digital PCR,” says he and Vogelstein developed the approach to better identify rare cancer mutations. “The thought that occurred to us was that the most sensitive [detection] you can get is by looking at single molecules,” Kinzler explains. “When you start with single molecules, [reactions] are either 100% mutant or 100% wild type, which makes the distinction of a tumor from wild type much easier.”

Of course, seeing even very rare events using just standard PCR should be possible. After all, PCR excels at plucking needles from molecular haystacks. In theory, given the right primers and cycling conditions, the reaction can copy a single piece of template DNA into millions upon millions of daughter strands, enough to clone, sequence, or detect on a gel. But because the process isn’t quantitative, researchers cannot infer the DNA content of the starting sample from the number of molecules present at the end of the reaction.

qPCR addresses that problem by quantifying the reaction as it runs. By charting fluorescence intensity over time, researchers can compare, for instance, the relative expression of a given gene from sample to sample. But absolute quantification by qPCR isn’t straightforward. It generally requires standard curves to convert abundance into absolute concentrations, and those concentrations can sometimes vary day to day and across labs. qPCR also struggles with detecting subtle copy-number

changes—distinguishing six and seven copies, for instance—and limited sensitivity.

The resulting uncertainty can complicate, for instance, data sharing—a critical limitation when setting up multi-institutional trials, says Muneesh Tewari, associate professor at the **University of Michigan** and formerly associate member of the **Fred Hutchinson Cancer Research Center (FHCRC)**, who uses dPCR for microRNA biomarker discovery. “Real-time PCR just doesn’t have the day-to-day precision, or sometimes even the precision within the same day that we would like,” he says.

dPCR circumvents these issues using a “divide-and-conquer” strategy. A mixture of molecules is discretized or compartmentalized into a large number of reaction chambers—the more the better—such that each chamber has on average either one target nucleic acid or none. Following PCR amplification of each compartment, Poisson statistics can be used to convert the count of positive signals into an absolute number.

Kinzler likens the process to Sanger sequencing. Sequencing a mixture of 100 template molecules, one of which is mutant, in a single tube would yield an electropherogram in which the dominant signal at the mutated position would be the wild-type base. “You probably wouldn’t even be able to tell there was a little hump underneath [the peak] for the alternative base read, because it represents too small a percentage of the molecules.” But diluting the molecules and distributing them in multiple reactions would produce 99 wild-type signals and one clear mutant—a digital, and absolutely quantitative, result.

Digital PCR essentially lifts weak signals out of the noise, and researchers today are using that strategy for everything from detecting copy number variations and circulating tumor DNA to HIV viral load monitoring and fetal aneuploidy testing. Clinical applications, in particular, seem most likely to benefit from the technology, assuming it can be made simple enough for clinical laboratory workflows.

Upcoming Features

Microscopy—May 2

Big Data—June 13

Digital Lab Management—July 25





THE ORIGINAL DIGITAL PCR

Published in 1999, Kinzler and Vogelstein's original "digital PCR" approach was a tedious, manual affair. Looking to detect and quantify K-RAS mutations in stool samples from colorectal cancer patients, the pair diluted and aliquoted genomic DNA into each well of a 384-well microtiter plate, such that the entire plate represented a single sample. They then amplified the gene region that encompassed the mutational hotspots they were looking for.

Once the reactions came to completion—dPCR generally is an endpoint PCR assay, though it doesn't have to be—they hybridized two fluorescent probes, a control that should always hybridize and a second that only binds to wild-type sequences, and read the results. Just over 100 wells had genomic DNA, of which four appeared to be mutants.

The results, Kinzler says, indicated dPCR was "a very robust, reliable, and accurate" method for rare mutation detection. But it also was labor-intensive and difficult to scale. "It was often impossible to convince a grad student to do it," he says.

In 2003, Kinzler and Vogelstein described an improved protocol, which they called "BEAMing" for its reliance on "beads, emulsion, amplification, and magnetics." In place of manual sample discretizing in a microtiter plate, BEAMing turns the individual droplets of a water-in-oil emulsion into reaction vessels. Using an emulsion PCR process that's now used widely in next generation DNA sequence library preparation, BEAMing distributes a mixture of template, primers, PCR reagents, and magnetic beads into droplets, again under conditions such that most droplets will contain either zero or one template. Following PCR, during which the amplified product is coupled to the bead via a biotin-streptavidin linkage, the emulsion is broken and the beads read by hybridization with detection oligonucleotides and flow cytometry.

Instead of manual separation on a plate, BEAMing "can do the equivalent of 100,000 wells," says Kinzler, who still uses the technology and (with Vogelstein and other colleagues) cofounded **Inostics** to commercialize it. (Inostics subsequently was acquired by Sysmex, and today Kinzler holds no equity in the company.)

Droplet-based strategies have also been commercialized by **Bio-Rad Laboratories** and **RainDance Technologies**. Instead of magnetic beads, both companies discretize (or, as RainDance President and Chief Executive Officer S. Roopom Banerjee puts it, "dropletize") the PCR mixture into either 20,000 nanoliter-sized (Bio-Rad's QX200) or 10 million picoliter-sized (RainDance RainDrop) droplets. Following PCR amplification, the reaction results are read by flowing the droplets past a fluorescence excitation source and detector, like a flow cytometer without the cells.

"Think of it as like a digital camera," says Banerjee, who says RainDance's RainDrop system was designed with single-button, "Apple-like simplicity. You have a complicated picture with lots of colors, lots of diversity. What we're doing is basically pixelizing that picture down to minuscule-sized pixels and then we literally read out every single pixel in order to say exactly what's in that biological picture." In RainDance's case, it takes four hours to read 80 million such "pixels," the output of eight parallel dPCR reactions.

POWERFUL POISSON

Because dPCR is an endpoint assay, the resulting concentration could in theory be thrown off if the target molecules didn't discretize perfectly—that is, some compartments ended up with more than one target. That's where Poisson statistics come in. According to George Karlin-Neumann, director of scientific affairs at Bio-Rad Laboratories' Digital Biology Center, Poisson

statistics describe a random distribution. As long the compartments aren't saturated, users can back-calculate how many molecules they started with, even if some wells actually receive more than one molecule (an event that, in dPCR, reads as a single count). "You tell me how many compartments or how many droplets you have, [and] what fraction of those are negative. That will tell me essentially what concentration I started with that would give me that ratio of positive compartments to negative compartments," he says.

Reginald Beer, medical diagnostics initiative leader at the **Lawrence Livermore National Laboratory**, who as a graduate student published the first report of dPCR in monodisperse (i.e., identically sized) droplets, says the advantage of the droplet-based approach lies in its scalability: It's relatively simple to increase the number of reaction vessels, which in turn increases data quality. "Poisson accuracy greatly improves as you scale or increase the number of your reactor vessels," he explains.

Published reports establish the efficacy of both Bio-Rad's and RainDance's dPCR platforms as well as the diversity of applications amenable to dPCR. RainDance's platform, for instance, has been used to detect mutant gene transcripts in cerebrospinal fluid from glioma patients and KRAS mutations in sera from colorectal cancer patients.

Jason Bielas, associate member of the FHCRC, used Bio-Rad's system to quantify tumor-infiltrating lymphocytes in ovarian cancer, while Keith Jerome, professor and head of the Virology Division in the Department of Laboratory Medicine at the **University of Washington** and associate member of the FHCRC, used it to develop a way to distinguish active HHV-6 viremia from latent, genome-integrated virus. (The former must be treated if it occurs during the course of bone marrow transplantation, but not the latter; the two are distinguished by the ratio of viral sequence copies to genome equivalents in blood, which in the case of genome integration should be 1.0.)

Bielas says his team was able to use their own method, termed QuanTILfy, to reproducibly resolve as little as one T lymphocyte in 10,000 cancer cells—enough to detect an association between T-cell tumor infiltration and patient survival. That method is considerably more quantitative, he explains, than the current standard of immunohistochemistry. "In **continued**>

The technique clearly offers considerable advantages for a growing number of specific (and often clinical) applications.

Featured Participants

Bio-Rad Laboratories
www.bio-rad.com

Fluidigm
www.fluidigm.com

**Fred Hutchinson Cancer
Research Center**
www.fhcrc.org

**Lawrence Livermore
National Laboratory**
www.llnl.gov

Life Technologies
www.lifetechnologies.com

**Ludwig Center at Johns
Hopkins**
www.ludwigcancerresearch.
org/location/johns-hopkins-
center

RainDance Technologies
www.raindancetech.com

Rutgers University
www.rutgers.edu

**Stanford Genome Technology
Center**
med.stanford.edu/sgtc/

Sysmex Inostics
inostics.com

University of Washington
www.washington.edu

essence, you're counting T-cell genomes."

Hanlee Ji, senior associate director of the **Stanford Genome Technology Center**, also uses Bio-Rad, both to validate genetic aberrations from genome-sequencing studies and more recently to quality control next-gen libraries prior to sequencing. For his group, the advantage of dPCR is largely simplicity: He can get an undergrad in the lab up to speed on dPCR in weeks, whereas qPCR can take considerably longer.

"One of my tests of technology is: If I take an undergrad who has reasonably good bench skills and sit them in front of an instrument, can they generate robust results from the positive and negative controls that they're asked to do within a matter of weeks? If so, that says a lot regarding how well the system works."

DIGITAL PCR-BY-SEQUENCING

In 2011, Kinzler and Vogelstein described yet another approach to dPCR for rare allele detection, this time based on sequencing. Called Safe-Sequencing System (Safe-SeqS), the strategy entails tagging individual template molecules with unique identifiers (or barcodes), which are then amplified and read out by next generation DNA sequencing.

According to Nickolas Papadopoulos, director of translational genetics at the Ludwig Center at Johns Hopkins, professor of oncology at Johns Hopkins Medical Institutes, and coauthor of the study: Safe-SeqS enables researchers to extend the capacity of dPCR to many more loci—up to 30 or more—than BEAMing can reasonably handle.

"Massively parallel sequencing represents a particularly powerful form of dPCR in that hundreds of millions of template molecules can be analyzed one by one," explain Papadopoulos and his coauthors in their 2011 study. "It has the advantage over conventional dPCR methods in that multiple bases can be queried sequentially and easily in an automated fashion." But NGS also has an unacceptably high inherent error rate caused by mistakes in amplification, sequencing, and detection. By tagging each starting molecule with a unique identifier, Safe-SeqS enables researchers to differentiate true mutations from procedural artifacts.

In the original study, the authors assessed mutations in the *CTN-NB1* gene in 100,000 normal human cells. Raw Illumina sequencing reads produced an error rate 2.1×10^{-4} . Taking into account the Safe-SeqS

barcodes lowered that rate 24-fold, to 9×10^{-6} . When applied to mitochondrial DNA, Safe-SeqS dropped the observed mutation rate 15-fold.

This past year, the team extended Safe-SeqS to search for multiplexed signatures of ovarian, endometrial, and cervical cancer in the scraped cervical material that comes from routine Pap testing, a first step in developing an early detection system for gynecologic tumors. In this case, the team used the assay to assess mutations in 46 gene regions from 12 cancer-associated genes.

Papadopoulos, who discussed the method at a recent digital biology conference in San Diego, says Safe-SeqS may be particularly useful for early tumor detection from plasma. Overall, his team could detect mutations in more than 80% of tumor plasma samples. But some tumor types worked better than others. Particularly difficult, he says, were malignancies in brain. "Currently, we can detect mutations in the plasma from only 10% of the patients with these tumor types," he says.

ARRAY-BASED STRATEGIES

For Andrzej Pietrzykowski, assistant professor of animal sciences at **Rutgers University**, who studies the molecular and genetic bases of alcoholism, dPCR's precision is particularly beneficial. Many of the changes he sees are less than twofold and thus difficult to discern by qPCR. That difference is readily detected digitally, he says.

But Pietrzykowski also cites another advantage. He has identified a particular microRNA that seems to be associated with alcoholism, and can quantify that molecule using qPCR. But qPCR assays for microRNA precursors are harder to pull off, as identifying reliable controls for standard curve calibration can be problematic. "You need to ... have that housekeeping gene double or triple-checked that it doesn't change due to conditions." That's a non-issue with dPCR, though, as the technique doesn't require a standard curve.

Pietrzykowski was one of five Innovation Grant recipients of **Life Technologies'** Digital PCR Applications Grant program, which in late 2013 provided him with the company's QuantStudio 3-D chip reader, chip loader, and thermo cycler. The QuantStudio 3D Digital PCR system is like a souped-up version of Kinzler's original dPCR implementation. Researchers use a tool akin to a windshield "squeegee" to discretize samples onto 20,000 wells etched into the surface of a silicon wafer. **Fluidigm**, too, uses this physical array strategy; its qdPCR 37K "integrated fluidic circuit" leverages the company's microfluidics expertise to distribute 48 samples into 770 chambers each (though a single sample can be spread over multiple sets of chambers to increase accuracy, up to 37,000).

The technique clearly offers considerable advantages for a growing number of specific (and often clinical) applications. But don't discard your real-time thermocycler just yet. For most researchers dPCR represents a complement to qPCR, not a replacement. Indeed, for many applications, says Beer, qPCR is "by far sufficient"—plus, the technology is far more mature and the assays established. "qPCR certainly still has a significant advantage from a throughput perspective," says Iain Russell, senior product manager for dPCR at Life Technologies (recently purchased by Thermo Fisher Scientific), "and to be quite honest, for a large number of applications out there it meets all the needs of the customer."

Says Pietrzykowski, "It really depends on what kind of question you're asking."

Jeffrey M. Perkel is a freelance science writer based in Pocatello, Idaho.

DOI: 10.1126/science.opms.p1400084