Drupal.behaviors.print = function(context) {window.print();window.close();}>

Send Out the Work

May 01, 2012

Send Out the Work

By Ciara Curtin

Keith Robison says he'd like to have his own sequencer. But there are a number of reasons why he does not. His company, Warp Drive Bio, is small and housed in a small space. There is not much room for a sequencer — even one of the newer desktop models — let alone the capacity to dedicate someone to running the machine.

There are also scientific reasons. Warp Drive is interested in developing natural products as pharmaceuticals. Researchers there are hunting through genomes to find regions encoding natural products that could be used as drugs. Many of the regions they are interested in are mobile elements, and can be large and highly repetitive. "That's a challenge from a *de novo* assembly standpoint," Robison says, especially for short reads.

Further, the sequencing work they do comes in bursts. "We'll have a whole lot of things we want to sequence and then it's a drought for a while, and then there is another burst of things," he says. "With having your own facility, you'd really want to level that out."

So Robison outsources his sequencing work, and he has used a number of service providers. "They have reasonable pricing, we had a good experience with them, and we keep using them," he says.

As new sequencing instruments come online that have a higher throughput at a lower cost, there are more and more sequencing service providers to choose from. "I think that the relatively low entry point for sequencing now with the MiSeq and Ion Torrent is creating more competition in the individual lab as well as academic sequencing cores; it's not all corporate," says Callum Bell, vice president for research and program lead at the National Center for Genome Resources in New Mexico.

At the same time, the high prices of the machines, among other factors, are enough to discourage all but the heaviest of users from pulling sequencing instruments into individual labs.

It's that confluence of cheaper yet still expensive machines that allows for more choice in service providers — from local core labs to large genome centers to commercial enterprises — while still making sequencing impractical for many -researchers to do themselves.

There are a number of services that providers offer, and while what researchers look for in service providers may vary, there are three common qualities they tend to weigh: cost, turnaround time, and data quality.

"Those are the three fairly straightforward considerations that we have with any partner: scalability, which informs turnaround time and cost, and, of course, quality," says Knome CEO Jorge Conde. Knome outsources its sequencing work to a number of different providers.

[pagebreak]

Send it out

There are several reasons why researchers or companies choose to send out their sequencing work, from ranging from a lack of local capacity to cost control. "In many instances, it's cost-competitive and even cost-effective to send the samples out and have them done for you," says Children's Hospital of Philadelphia's Hakon Hakonarson.

However, Indiana University's Milan Radovich cautions researchers not to rely on price alone. "Sometimes people would just use price as the only metric to determine who they are going to use as a service. ... I think that the point to get across is that there is more to service than just price and all of those things need to be considered," he says.

Merck turned to BGI as a partner for its sequencing work, rather than building infrastructure there, and has been happy with the results, says Tom Fare. "BGI has consistently delivered high quality data for a range of projects and sequencing applications," he says. Merck and BGI are collaborating on a number of biomarker discovery and development projects.

Further, sequencing technology moves so quickly that it can be difficult and costly to keep up to date. "The machines that are being used to sequence are still very expensive, but also the technology is evolving very, very rapidly. The useful life of the current generations of machines and certainly the previous generations of machines was relatively short," Knome's Conde says. As Knome focuses on the interpretation of -genomes, the company decided that it didn't make sense to invest so much money on machines that could quickly become obsolete. "We made a conscious decision to say if we going to be an interpretation company, that's what we want to focus on. It probably isn't the best use of our resources to spend on building up [a sequencing] infrastructure," he adds.

Warp Drive's Robison adds that his company is also wary of spending money in a rapidly changing field. "Management is understandably leery to sink a lot of capital into an instrument now and having someone more or less dedicated to run it at least some of the time, and then six months later say, 'Oh, but now we want the new toy,'" he says.

The same logic holds for other users, Conde adds, especially for those who do not do a lot of sequencing. "In other words, if you are a lab and you are going to do relatively low-volume sequencing, it really doesn't make sense to go out and establish your own sequencing capability and capacity," he says.

There is, however, an inflection point. When people become heavier users of sequencing

capacity, Conde says, it then would then make sense for them to bring sequencing equipment in house.

When Hakonarson's group was just starting out in sequencing, it used both BGI and Complete Genomics' services. "At the time, we sort of were much more focused still on genotyping because we have one of the largest genotyping centers basically there is. We were just doing sequencing on a relatively small scale, and so it was just more convenient to send it out," he says. Now, Hakonarson adds, as they are ramping up the scale of their sequencing, they have entered into an alliance with BGI to bring that group's capabilities to their institution. BGI@CHOP opened last fall, and focuses on sequencing and bioinformatics studies of rare and common pediatric diseases.

"We are basically sequencing our samples utilizing BGI's workflow and informatics and technology, but the sequencers and the data are all based at CHOP," Hakonarson says.

Of course, some people will always want to do sequencing themselves, no matter what. "Much like a guy who wants to go out in his garage and work on a '69 Camaro every day just because he wants to be an expert in '69 Camaro, those are the people that probably should not outsource. They should probably go buy a sequencer and dig into it," says Jon Armstrong, chief operating officer of Cofactor Genomics. "You have to be willing to accept the fact that probably the first five to 10 libraries that you make may fail. You have to be ready to accept that some of the early runs you do, if they turn out looking bad, you may not be able to troubleshoot."

[pagebreak]

Expertise

What many sequencing service providers say they offer their clients is expertise — sequencing is what they do. "What we found is that because we have done a thousand projects since our inception, that through those projects and running those through the whole system, we're experts and knowledgeable about how you design a sequencing project," Armstrong says. Cofactor houses an ABI SOLiD instrument, as well as machines from Illumina, Roche 454, and Ion Torrent.

Cornell University has a large web of core facilities that offer a number of services. The genomics core there has first-generation capillary electrophoresis machines as well as Roche 454 GS FLX, Illumina GAIIx, and HiSeq 2000s. "Everything that you think you can do with these instruments, we're doing, with large diversity of interest," says George Grills, director of operations for Cornell's life science cores. The draw for many researchers to use the Core Laboratories Center is that its cores offer more than sequencing and can integrate bioinformatics, proteomics, and more into a project, he adds.

"We have been sequencing essentially anything that can be sequenced in terms of samples, in terms of applications, but also we have a relatively large scale or portfolio of organisms," says Vladimir Benes, the head of the genomics core lab at the European Molecular Biology Laboratory. His core has a number of Illumina instruments, including HiSeq2000s, as well as an Ion PGM.

Other, larger sequencing providers including BGI, and more local resources like the National Center for Genome Resources in New Mexico and the Australian Genome Research Facility,

offer a number of services as well. BGI has 137 Illumina HiSeq 2000 instruments, 27 Applied Bio-systems SOLiD machines, and a number of Roche GS FLX Platform and Ion Torrent machines.

"One thing that differentiates BGI from just a regular service provider is its publications and research capabilities," says Joyce Peng, marketing director for BGI Americas. "It's the capability to do good science that attracts a lot of people to work with us."

NCGR and AGRF similarly offer a number of sequencing platforms — NCGR has a PacBio machine, in addition to Illumina machines, and AGRF has five HiSeqs, a 454, four Ion Torrents, and two MiSeqs. Both organizations serve academic researchers who focus on a range of fields, from biomedicine to agriculture and commercial enterprises. "Currently, next-gen sequencing comprises about 27 percent of our business, and that's been growing quite strongly year on year. The HiSeq platform, which is probably the main next-generation platform we use, comprises about 19 percent of our business," says Kirby Siemering, AGRF's business development manager.

Other enterprises like Complete -Genomics have sharper concentrations. "We focus exclusively on sequencing whole human genomes. We don't do any other organism and we don't do partial genomes or gene sets or exomes or anything like that," says Complete Genomics CEO Clifford Reid. Some of his firm's projects include cancer genome sequencing efforts including a study of neuroblastoma and one of hepatitis B virus integration into liver cancer, among others.

Part of the expertise that service providers say they offer is good quality control, which comes, in part, from the volume of projects they process and run. "Our wet lab people make hundreds of libraries a week, and we are running machines multiple, multiple, multiple times a month. … We are able to catch problems very early on in the process," Cofactor's Armstrong says.

"We can run thousands and thousands of whole human genomes and we can get better at it each time we do it," Reid adds.

Quality often trumps cost considerations. "I know that without a quality delivery I could offer almost for free and people wouldn't come," adds -EMBL's Benes. "The combination of competence and skills and going further and the guidance that we give them ... that makes people come to us."

And quality is naturally of high importance for customers. "When you are in the genome interpretation business it's important that the data is correct," Knome's Conde says. "In other words, the quality of that data is correct; that's paramount, otherwise we're interpreting the wrong thing."

As a customer himself, Warp Drive's Robison adds that "when the data shows up, I want high-quality data. I mean I can understand sometimes I've submitted lousy samples and I get lousy data back, but if I submit good samples, I want to see good data back. And most times you do."

Outsourcing can also spare customers the headaches of things going wrong. "For customers that have purchased instruments or reagents, if something goes wrong, the problem is theirs and [can] cost them a lot of reagents and instrument time." Reid says. "Our customers are completely insulated from that cost."

[pagebreak]

A matter of scale

Due to the number of machines and the number of projects coming through their pipelines, sequencing service providers also offer an economy- of scale, which is then passed on to their customers.

In addition to the costs of the machines, setting them up and bringing in reagents can be expensive. -Sequencing on a small scale means that there are fewer samples to spread that cost around on. "If you go out and you buy, for example, an Ion Torrent machine or a MiSeq, it may cost you $400,000 or $500,000 to get that machine set up, on top of the machine cost," Cofactor's Armstrong says. "That means that the first run that you do on that machine costs $500,000." With the volume his company sees, it can offer lower average total cost, he adds.

The more projects that are processed, the lower the costs can get. "We have economies of scale by consolidating projects across dozen to hundreds of customers than any one customer can achieve themselves," adds Complete Genomics' Reid.

Scale often translates to faster turnaround. Quick turnaround is what customers like Warp Drive's Robison are looking for, though the speed often depends on the machine and the scope of the project.

"We have the capacity to do about a thousand genomes a month," Reid says. "A 500 or 1,000 genome project, we can turn it around in our standard turnaround time rather than queue up all those samples."

For the HiSeq at AGRF, Siemering says that for a project of moderate size, they have an average turnaround time of six weeks. EMBL's Benes notes that they, too, have an approximate six-week turnaround time. Similarly, Cofactor's Armstrong gives a range of four to eight weeks, again depending on the project. "The reason that there's that big variability is mainly because it owes itself to the difference between a single-end one by 50 run and a paired-end run two by 100 run, where a two by 100 run and base calling may take two weeks and a single end may only take a week," Armstrong says.

Further, he adds that his company offers a number of different timing options — at different prices — and that four to six weeks is for their 'rush,' while their standard pipeline takes between six and 10 weeks, and what Armstrong calls their 'super saver' pipeline is reserved for projects that can fit into the workflow as the opportunity arises.

However, delays do happen, and sequencing isn't always very fast. "Even today despite all of the advances in sequencing technology, it still takes quite a while to get data from a sequencing project," Knome's Conde says.

In 2011, Complete Genomics had a bit of a lag in its turnaround time. The company sent out fewer genomes in the last quarter of 2011 than it expected, due to a delay in a lab expansion. "It went up to about 70 days and it will probably be up around that this quarter," Reid says, adding that one of the company's initiatives aims to reduce turnaround time. "We anticipate that by the end of this year, our turnaround time will be between 30 and 45 days and we expect that by the end of next year to be working on turnaround time of approximately two weeks."

Clinical needs

Researchers working on clinical projects often have the added concern of making sure that their results conform to laboratory testing regulations, like the Clinical Laboratory Improvement Amendments certification in the US. That need may drive where a sequencing project is sent.

"Another thing to consider is whether or not your project is actually a medical genomics project that requires a CLIA-certified clinical diagnostics environment for data generation," Cornell's Grills says. "Setting that up de novo at most institutions is a challenge. There are a lot of institutes that have already done that, or are on the road to do that, or are seriously considering doing that."

[pagebreak]

Indeed, he says that Cornell has two working groups — one at its medical center and the other at its veterinary school — looking into setting up medical genomics capabilities there so that researchers could do CLIA-certified diagnostics using next--generation sequencing. Cornell isn't the only place looking to receive CLIA certification, he adds. Harvard and the Mayo Clinic are also working on it, and other core labs are interested in garnering certification as well.

"There are many institutions which are still in the planning stage, and if you need that type of environment to meet regulatory compliance for data generation, then sometimes commercial services might be your only option," Grills adds.

The new BGI@CHOP center is also working toward CLIA certification, Hakonarson notes. At the time this article went to press, he estimated that certification was "about six weeks away." There are a number of programs there that would benefit from CLIA certification. One project is examining genes that are mutated in patients with attention-deficient hyperactivity disorder, and CHOP researchers think they have a drug that could reverse the effects of those mutations. "[We] now need sequencing to find patients to screen and test for the compound," Hakonarson says.

Several companies are also going through the certification process, though some already are certified.

Indeed, insurance company UnitedHealth Group issued a working paper in March that estimated that the US spent about $5 billion on genetic testing and molecular diagnostics, an amount it said would likely grow to between $15 billion and $25 billion by 2021. "The big market here in the long run is really the clinical market for genetic diseases," Reid says. By comparison, he adds that the sequencing research market is at $1 billion to $3 billion a year, and is expected to grow to $2 billion to $4 billion by the end of the decade. "It's pretty clear where the interest lies from a business and commercial point of view in sequencing," he says, adding that Complete Genomics is seeking CLIA certification for its entire facility so that "every genome that we produce will come through a CLIA-certified lab."

The company is already working with clinic-oriented customers. In 2011, it partnered with Inova Health System, a network of hospitals in northern Virginia, to sequence 1,500 genomes to try to determine whether there is a genetic component to premature birth. They are sequencing 500 babies — 250 of whom were normal births and 250 of whom were premature — and their parents. "The reason they are doing this is not for a basic research agenda. It is

for a translational research agenda, because premature births are extraordinarily expensive," Reid says, adding that "if there is a genetic component of premature birth and it can be identified before the event and can be mitigated in some way, then [there will be an] economic benefit for major hospitals."

Each in its place

While there is increased competition among sequencing service providers, different providers may emphasize different aspects of their technology, service, or expertise. There are certain projects that would not make sense for one provider to do, but might be a good fit for another.

"Cofactor has put a lot of emphasis on doing *de novo* assemblies, and then there are only a few shops that have PacBio instruments," Warp Drive's Robison says. "If you want to do optical mapping with OpGen, that's a very rare beast out there. As the newer instruments like the Ion Proton, I've already had one vendor claim that they are definitely going to have one. MiSeq, we're actually working with a local core lab with one." He adds that he doesn't feel "wedded to any one provider."

Grills adds the many companies occupy certain niches. "Companies like Complete Genomics, which offer excellent services, are really restricted to human-based sequencing. That's outside the interest of a lot of the investigators at our Cornell campus," he says.

Grills also draws on his past experience setting up a genome center at Harvard University. "Of course I had to juggle a lot in terms of making sure that what we were offering was appropriate and unique. What I found was, yes, there are overlaps and areas of competition, but there are certain types of projects which are unique to the scope and scale of, at that point, a mid-size genome center, or in my current environment, are appropriate to academic core facility," he says.

There are projects being done at Cornell that are beyond the scope of what Grills says the core there could do, practically speaking. "There's hundreds of thousands of samples and those are sent to some of the larger genome centers, particularly BGI or the Broad or WashU, and those are the scope and scale where it wouldn't make sense to do it in an academic core environment," he says.

Indeed, EMBL's Benes adds that there are some projects that do not make sense to do at his core. Projects that are not very high throughput, like CLIP-seq, GRO-seq, and Hi-C, can tie up his machines for too long.

"There is no one-size-fits-all," Reid adds. "There is a role and a place for in-house sequencing, and I think it is in these small, diverse projects where it doesn't make a lot of sense to send them out. They can be done just as well internally as externally and having that flexibility is very valuable. And then there are the large standardized projects and [they] naturally fall into a centralized model because it can do the same thing over and over and over again."

[pagebreak]

A new model

Just how the sequencing field will play out is anybody's guess. New technologies could still shake it up. "Sequencing is obviously still in its infancy. There is a massive amount of talent

and efficiency that has come out of BGI. I think they will continue to be the leading institute in sequencing unless the technology totally changes, like the Oxford Nanopore or something. If that works, the landscape may change," Hakonarson says.

In the nearer term, there appears to be space for sequencing providers both big and small.

Both Complete Genomics' Reid and CHOP's Hakonarson say that the current set-up for clinical diagnostic labs could serve as a model for sequencing service providers. Hospitals do some tests internally and outsource others to references labs like Quest Diagnostics or the Laboratory Corporation of America. Typically, Reid says, hospitals will do their own simple tests, like blood tests, in house and send samples out when there is a need for more complex or esoteric tests. "There's a very comfortable division of the market between internal and external testing in the existing diagnostics market," he says. "I think the sequencing market is going right there."

"I think there will continue to be large international groups who have strategic interest in genomics ... but then I think there will be other organizations, probably nationally-based who are working with local researchers, working on local projects of -national interest in their jurisdictions," Australia's Siemering adds.

Then again, the field is moving very quickly. If cheap desktop sequencers come to the market, Siemering says the field could be democratized, to an extent.

Outsourcing makes sense when machines are expensive, adds Knome's Conde. But should they come down further in price — and he notes that they appear to be on course to do so — more researchers may bring them into their labs. "If the sequencer itself really becomes another desktop tool that generates data very quickly, I think it is more likely for small-end users to invest in that machinery and then essentially run the sequencing themselves in house," he says.

Even then, Conde adds, there will still be a role for larger facilities. He likens it to copy machines: Most offices have copiers for workers to do routine copying for a decent price. But if there's a big job to be done, they'll still head out to Kinko's.

"I think there will always be a place for the larger sequencing cores simply because people will just make their questions bigger and bigger to take advantage of the technology's capabilities," Siemering adds.

Ciara Curtin is the editorial director of Genome Technology as well as the GenomeWeb blogs, including the Daily Scan. She also blogs about molecular tools in the clinical lab at the Sample. E-mail her here and follow @GenomeTechMag, @DailyScan, and @SampleGW on Twitter.

Related Stories

- The Cancer Fight
  April 1, 2012 / Genome Technology
- Sixth Annual Young Investigators
  December 1, 2011 / Genome Technology

- Humans as Host
  August 26, 2011 / Genome Technology
- Personalized Perspectives
  May 1, 2011 / Genome Technology
- To Stop Cancer
  April 1, 2011 / Genome Technology

footer