



Microsoft Expands Azure Software Ecosystem With New Informatics Partnerships

Oct 18, 2018 | [Uduak Grace Thomas](#)

NEW YORK (GenomeWeb) – Microsoft Genomics, part of Microsoft's healthcare unit, has announced several informatics partnerships aimed at providing a variety of analysis capabilities and solutions on the company's Azure platform and effectively extending its reach in the genomics market.

Specifically, Microsoft said recently that it is partnering with BizData, Eagle Genomics, Genoox, Gentera Biotechnology, L7 Informatics, Parabricks, Qiagen, and Veritas Genetics. The company says it now has more than 20 partners in the genomics space covering everything from sequencing prep through to variant interpretation. This includes [a deal with DNAnexus](#) to include its genome informatics and data management platform and services on Azure, which was announced in 2016.

As part of the current agreements, some partners have or are in the process of implementing their solutions on the Azure cloud. In some cases, these companies' pipelines will accept output from Microsoft's genomics analysis solution as input into their own products. Still other companies will work on implementing Microsoft's secondary genome analysis pipeline in their own platforms. Microsoft's analysis pipeline is an implementation of the Broad Institute's best practices pipeline and includes tools like the Burrows Wheeler Aligner and [the Broad's Genome Analysis Toolkit](#).

"The partnerships that we are forming are part of a multi-year journey to broaden the ecosystem of genomic partners" and extend the Azure platform to include a much wider variety of features and functionalities that support genomics activity, said GERALYN MILLER, director of Microsoft Genomics. The company also chose to partner with these particular companies in part because of customers' requests for specific solutions to be made available on Azure cloud.

"We acknowledge that we are never going to have all the pieces and parts to be able to deliver what's needed in the space of genomics," she said. "We choose to address that by partnering with others."

Veritas Genetics, for example, offers both software and access to sequencing services. On the informatics side, this includes the open-source bioinformatics platform Arvados, which the company obtained when it purchased [Harvard Medical School spinout Curoverse](#) last year.

Microsoft is an ideal partner for Veritas because of the company's footprint in the marketplace, said Marc Rubenfield, Veritas' director of business development, research services. However, "we can bring a lot to the table for them in [areas] they are not quite as familiar with." For example, Veritas is one of the few companies that Microsoft is partnering with that has a lab for generating sequence data, he said. In addition, the company offers tools to process sequence data including performing pharmacogenomic analysis and identifying actionable genes and traits of interest, among other capabilities.

The [Arvados platform already runs on the Azure cloud](#) and Veritas is now working on integrating Microsoft's analysis pipeline into the Arvados infrastructure, Rubenfield said, offering current Arvados users another option for their secondary analysis needs. They are implementing the pipeline using the Common Workflow Language so that anyone currently running an Arvados instance will be able to simply pick up the Microsoft pipeline and run it on their data, he said.

For its part, Genoox offers software applications for translating complex genetic information into actionable clinical information. Its platform includes tools for aggregating and analyzing differing kinds of data including genotypic and phenotypic information and correlating it with publicly available literature. Customers can cobble together tools from the Genoox platform into custom analysis pipelines or use pre-configured pipelines to process data from raw sequence through to clinical results.

Current customers of Genoox's products use its pipelines to analyze data in the context of reproductive health, inherited disorders, and oncology. Recently, Sanford Imagenetics, the molecular diagnostic laboratory of Sanford Health, announced that it is using [a free classification tool from Genoox](#) that is designed to help geneticists better classify genomic variants in accordance with guidelines from the American College of Medical Genetics. The company also has a partnership with [Bionano Genomics](#) to build an informatics workflow for integrating and annotating structural variants with sequencing data.

The current agreement allows Genoox to include Microsoft's secondary genomic analysis pipeline as an option for its customers, according to Guy Harel, Genoox's director of business development. It simultaneously expands Genoox's existing suite of analysis tools as well as its potential customer base. Meanwhile, Microsoft's customer base benefits from access to more applications and pipelines including tools for variant interpretation, classification, and reporting in clinical contexts.

It "broadens our healthcare partner portfolio and brings us closer to actionable insights at the patient-provider level while addressing the heavy manual requirements of transforming disparate datatypes," Trent Norris, senior program manager at Microsoft Healthcare, said in a statement announcing the partnership. "This proposition is in line with our goal of providing AI and machine learning processing services for healthcare to handle a diverse set of data-rich workloads."

Meanwhile, the partnership with Eagle Genomics is emphasizing tools for analyzing microbiome data. The partnership will combine Eagle's expertise in curating, analyzing, and interpreting life sciences data with Microsoft's Azure cloud technology and artificial intelligence capabilities to address microbiome data analysis challenges. The platform will support applications in various contexts including healthcare, personal care, cosmetics, and food, the company said.

Parabricks' agreement with Microsoft will see its graphics processing unit-based sequence analysis software made available to customers on the Azure cloud. Earlier this week Parabricks announced that it had won a [Phase II Small Business Innovation Research grant](#) from the National Science Foundation worth roughly \$750,000 to further develop its secondary genomic analysis capabilities.

Parabricks' CEO and Cofounder Mehrzad Samadi said that the combined solution will offer yet another alternative for customers looking to shorten analysis times and lower the costs associated with analyzing large quantities of genomic data. Besides Microsoft, Parabricks' solution also runs on Amazon and Google clouds. The company also offers the solution as a standalone software. Moreover, he added, customers don't need a lot informatics know-how to run Parabricks' software on the cloud, and they'll benefit from the comprehensive support that Microsoft offers to Azure customers.

According to Vasu Rangadass, L7 Informatics' President and CEO, combining his company's [Enterprise Science Platform software](#) with Microsoft Azure's Genomics Service will give sequencing centers a complete software solution that covers sample management, laboratory information management system functionalities, and sequence data analysis capabilities. His company will also benefit from the broader sales distribution network that Microsoft can provide.

Under the terms of the agreement, L7 Informatics will be offering the Microsoft genomics pipeline as part of its existing ESP software product, Rangadass said in an email. The integrated solution will offer, among other benefits, tools for transferring data to and from the Azure cloud to on-premise ESP iterations, as well as Microsoft Genomics-specific configuration stubs to ESP.

As part of its deal with Microsoft, BizData will work with customers looking to implement their pipelines on the Azure cloud, said Nadav Rayman, BizData's director of sales and marketing. The analytics company, which began working in the genomics research market some three years ago, helps customers find more cost-effective solutions to their big data problems. In the context of this partnership, "we are helping customers onboard to Azure," he said. "So we are taking their existing pipelines and putting them on Azure and [helping] run them on the cloud in the most cost effective way."

It's an opportunity to showcase BizData's analytics capabilities as well as its expertise in executing pipelines in the cloud, Rayman added "The value-add that we provide is that we are not just narrowly focused on one step in the genomics pipeline. We understand analytics really well so we can help researchers with that."

Partnering with Microsoft will help boost Genera Biotechnology's profile in Turkey and the Middle East, enabling the company to secure more customers for its products, Erdi Keleş, a public relations specialist with Genera, said in an interview. The partnership also extends Microsoft's reach in the region, he added.

Genera, which was founded in 2013, provides next-generation sequencing and analysis and microarray services as well as primer design and synthesis services. It has also developed a Linux-based server that offers tools for bioinformatics analysis tasks. Customers can load their raw sequence data onto the server and receive a report that includes information on variants present as well as links to relevant public literature. Under the terms of the agreement with Microsoft, the company will implement its bioinformatics server on the Azure cloud, Keles said.

For its part, Qiagen is working on implementing software products that it inherited [with its purchase of OmicSoft](#) last year, Matt Newman, OmicSoft's vice president of business development for bioinformatics and the global product manager for the OmicSoft product line, said in an interview.

This includes the company's solutions for genomic sequence data analysis and visualization as well as its Land database portfolio which includes a repository of public RNA-Seq, microarray expression, and methylation datasets across hundreds of diseases and a database of oncology data among other resources. "We've integrated with Azure Batch so that the customers could run the OmicSoft pipelines directly on Azure," he said. "It also works with Azure files and so they can store their NGS data and that could be anything from the raw fastq files coming off the sequencer all the way through to the downstream variant call or expression data or other kinds of data."

According to Newman, Qiagen made the decision to partner with Microsoft in response to customers' requests. OmicSoft's solution are already available on the Amazon cloud and the partnership with Microsoft allows the company to offer its customers an alternative. Moreover, it allows the company tap into a potential pool of customers who already have

solutions running on the Microsoft cloud and would prefer to keep all their software in a single system. "Those are customers where if we are on their marketplace and they are looking for omics data analysis, for storage and visualization, [and] for databases, it's going to be a natural fit," he said.

Microsoft's Miller said that the company continues to look for additional solution partners in the genomics space. New partnerships are in the works, but Miller declined to disclose details about who these partners are, adding that the company plans to make announcements about them in the coming months. Separately, Microsoft's genomics group is working on a method for analyzing long reads to identify structural variants in genomic data that they are implementing on field-programmable gate arrays.