

Selecting a LIMS for the Next-Generation Genomics Lab

The evolution of genomics has combined human ingenuity with right-place/right-time serendipity. Certainly the Human Genome Project would never have succeeded without the innovations envisioned and executed by the scientists responsible for the effort. But these successes in turn were made possible by advances in computer processing and storage that provided the bandwidth and throughput to power and run the project.

Today, the technologies and methods pioneered during the Human Genome Project have revolutionized the life-science industry. In fact, no other industry has seen processing speeds rise and costs drop as dramatically as genomics. The advent of next-generation sequencing has given organizations the ability to sequence entire genomes in less than a day for pennies per base pair.¹ Not surprisingly, organizations are now wondering how they will handle the data these techniques generate.

But how well are organizations sizing up the challenge? Most discussions of data management in next-generation sequencing focus on gathering and storing the reams of data generated by instruments. This challenge is not to be dismissed, but it should be considered in tandem with how contextual sample and project information will be used to inform downstream analysis and critical research decision points. Recent commentaries note that while the cost of sequencing has decreased, analysis costs remain high—over half again as much as sequencing alone.²

One critical obstacle to analysis is that instrument data is often siloed and stored separately from key contextual information that describes the experiment and samples being sequenced—what they are, where they came from, and how they are prepared for experimentation. To ensure that bioinformaticians have the information they need to conduct higher quality, faster, and more informed analyses, labs need to track and trace sample information from the point samples are acquired to when sequencing results on that sample are reported.

Lab information management systems (LIMS) are a mature class of life science software, and commercial systems are now available that are specifically designed for genomics. The best of these systems offer the following advantages to modern sequencing facilities:

- End-to-end sample traceability
- Scalability so that labs can get up, running, and producing results quickly
- Adaptability to help labs

accommodate changing technologies and methodologies

- Workflow management and operational reporting tools to ensure labs run efficiently and collaboratively

Achieving these benefits requires labs to assess available LIMS against their specific experimental needs and research workflows. This paper reviews three criteria that every lab should evaluate when selecting a LIMS for next-generation genomics. The choice will depend on

- The specific sequencing instrumentation a lab is running or plans to run
- What system customizations will be needed to enable the lab to meet research objectives
- What types of scientists the lab employs and how these scientists will need to interact with the LIMS

Next-generation sequencing: Bringing genomics into the mainstream

Prior to the first draft of the human genome by scientists at Celera in 2000, genomics was an elite discipline that required a tremendous investment of time and money. Completing the first draft required more than a decade and cost an

Selecting a LIMS for the Next-Generation Genomics Lab

estimated \$300 million annually. It also launched genomics into the mainstream. The technologies and methods developed over those initial two decades have continued to evolve, lowering sequencing costs and increasing data volume by several orders of magnitude. No other industry has seen such precipitous gains in throughput and drops in cost. Even the well-known Moore's Law (the suggestion first made in 1965 that computer processing speed doubles every 18 months) is growing slower than DNA sequencing (see Figure 1).³ The consequences are easy to track. It took three years and a total of \$3 billion for scientists to release the final draft of the human genome.⁴ Now, just 10 years after the publication of the first human genome draft, businesses are competing to sequence an individual's entire genome in a matter of weeks for about \$10,000, and experts are claiming that "the thousand-dollar genome is within sight."⁵

Genomics still requires an investment, but it's one that more and more labs can afford to make, particularly as the three primary manufacturers of next-generation sequencing instrumentation (Life Technologies, Illumina, and Roche) continue to drive the cost down while pushing their instrumentation to set new throughput records. Table 1 compares the leading systems for next-generation sequencing available in 2010. All of the systems reviewed by Deutsche Bank are capable of producing several gigabytes of data per run, and the leading systems are poised to sequence entire genomes for a total cost of less than \$10,000.

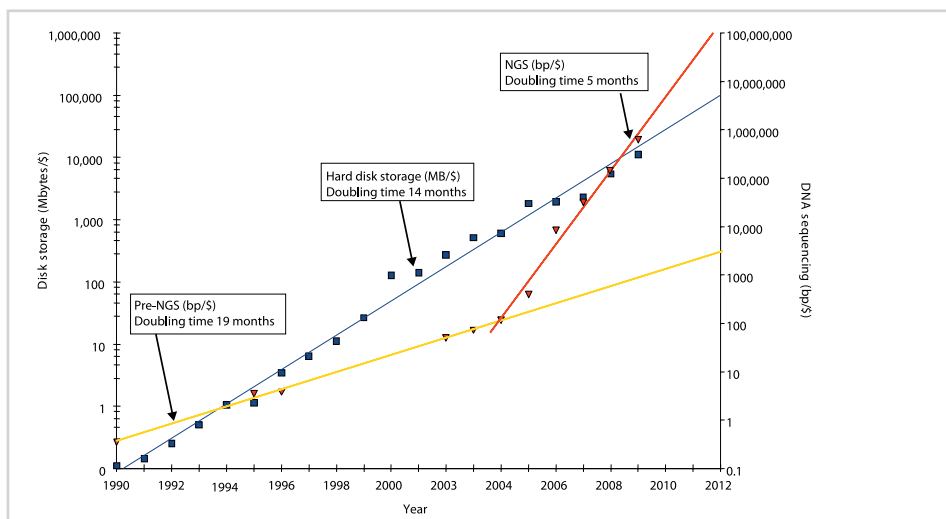


Figure 1. A logarithmic plot of historical trends in storage prices versus DNA sequencing costs. The advent of next-generation sequencing (NGS) in 2004 causes an inflection (red) in the exponential curve associated with DNA sequencing costs (yellow) to a doubling time of less than six months. (Source: Stein, *Genome Biology* 2010)

Leading systems for next-generation sequencing¹⁴

	THROUGHPUT	READ LENGTH	NO. READS	RUN TIME	LIST PRICE	COST / GENOME
ILLUMINA						
HiSeq 2000	~600 GB	2 x 100bps	6000 MM	10-11 days	\$690k	<5k (total)
HiSeq 1000	~300 GB	2 x 100bps	3000 MM	10-11 days	\$560k	<5k (total)
GAIIX	50-95 GB	2 x 100 bps	225-250 MM	9-10 days	\$205k	~40-45k (total)
LIFE TECHNOLOGIES						
5500 SOLiD	90 GB	75 bps x2	—	7 days	\$350	—
5500 xl SOLiD	180 GB	75 bps x2	—	7 days	\$600	—
SOLiD4	100 GB	2 x 50 bps	1400 MM	12-16 days	~\$500k	\$6k (reagent)
SOLiD3	50 GB	2 x 50 bps	1000 MM	12-14 days	~\$450k	\$10-15k (reagent)
SOLiD PI	upto 50 GB	2 x 75 bps	800 MM	12-14 days	\$230k	—
ROCHE						
GS FLX Ti	0.45 GB	400-500 bps*	1.25 MM	10 hours	~\$500k	—
GS Junior	0.035 GB	400 bps	0.1 MM	10 hours	~\$110k	—
ION TORRENT						
PGM™	0.1 GB	100 bps x2	1.5 MM	0.08 days	\$50k	—
HELICOS BIOSCIENCE						
Heliscope	25 GB	—	700 MM	8 days	\$1.35 MM	—
PACIFIC BIOSCIENCES						
RS2010	2.94 GB	1000 bps x2	2.94 MM	0.1 days	\$700	—

Table 1.

Selecting a LIMS for the Next-Generation Genomics Lab

These capabilities have resulted in unprecedented data production. Figure 2 tracks the number of genomes entered into GenBank each year, illustrating the rapid increase that has occurred since 2006. It's important to note that when this chart was published in 2009, 5,343 additional projects were underway and therefore not reflected in this data.⁶ More impressively, the 1000 Genomes Project, the first large project to capitalize on next-generation sequencing technologies, deposited twice as much raw sequencing data into the GenBank archives in its first six months of operation as had been deposited into GenBank in the 30 years since its inception.⁷

Yet organizations won't realize the promise of genomics merely by producing more and more data. Rather,

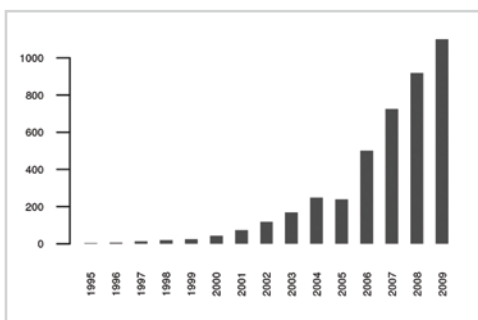


Figure 2. Number of genomes entered into GenBank by year as of September 2009.

their sequencing efforts will succeed or fail based on how well scientists exploit genomics data. At a minimum, scientists need to be able to effectively compare genomes across patient cohorts. Ultimately, though, studies aim

to catalog gene variation in humans and other species and identify the specific sequences or mutations that cause disease. In this context, too much disconnected information is the same thing as no information. Without a way to rapidly and intelligently interrogate these massive data sets and retrieve the associated sample information, scientists will be unable to make sense of the data and use it to drive research decisions.

The need for data and lab information management for next-gen sequencing

The consequence of next-generation sequencing's throughput potential is that labs can easily produce more data than they can effectively manage or analyze. Industry analysis that once focused on the costs associated with sequencing genome data now focus on the challenges of managing it. In a J.P. Morgan report conducted in 2010, 64% of lab directors cited data storage, data management, and informatics as the biggest collective hurdle to expanding next-generation sequencing operations.⁸ Third-generation sequencing systems, which some labs have begun to implement, are further shifting the data management burden upstream of the instrument run to place an even greater emphasis on sample handling in order to avoid cross- and background contamination.

The hurdle, therefore, is more of a steeplechase than a single, easily cleared obstacle. Storage was the first critical concern for most organizations as they confronted the reality that new machines



Selecting a LIMS for the Next-Generation Genomics Lab

running at capacity could generate in a single year more information than was deposited in GenBank by the beginning of 2008.⁹ In some cases, labs have changed their initial data handling strategies midstream to free up space; image files, for instance, are by far the bulkiest data types produced by sequencing, and some labs have opted not to store these file types.¹⁰ Other labs have realized that processing and storage power are relatively cheap. Their strategy has been to store everything and then figure out afterwards what they need for analysis.¹¹

A “store everything” approach shifts the bottleneck to analysis, which has emerged as the second major challenge for sequencing labs. While the total cost of sequencing a human genome has lowered exponentially, analysis costs remain high. The most generous estimates put analysis at half again as much as the cost of sequencing. Researchers at the National Center for Genome Resources said that the bulk of the costs in a quarter-million dollar sequencing project in 2009 is composed of analysis expenses. “An awful lot of manual analysis is required” according to this report. “It’s a very large amount of human effort.”¹²

Clearing the analysis hurdle requires more than an investment in hardware, infrastructure, and bioinformatics expertise. Organizations must completely revamp the workflows that support sequencing, many of which are based on manual, one-at-a-time processes and information stored in disconnected silos such as spreadsheets, emails or document-based communications, and paper lab notebooks. Sample preparation



often emerges as a critical area of emphasis for organizations seeking to streamline operations. Many of the most prestigious grants and research projects often require labs to be able to guarantee sample traceability—it’s essential when dealing with the often limited DNA supplies associated with certain clinical sample cohorts. Nevertheless, busy labs often struggle to ensure that samples received from clients and collaborators are appropriately labeled and that all vital experimental context is passed on efficiently and accurately to bioinformaticians. Clear sample taxonomy, tracked from the moment a sample enters a lab to the point at which results are reported, makes it easier for research scientists and bioinformaticians to set up

Many of the most prestigious grants and research projects often require labs to be able to guarantee sample traceability.

Selecting a LIMS for the Next-Generation Genomics Lab

Overview of information required to track next-generation sequencing research

PRIMARY PEOPLE INVOLVED	STAGE IN WORKFLOW	KEY INFORMATION TO TRACK	
Lab Managers External collaborators Principal Investigators Bioinformaticians	Project Initiation & Sample Submission	<ul style="list-style-type: none"> Defined research goals Agreed upon experimental approach, number of samples <ul style="list-style-type: none"> Types of samples Sample taxonomy Type of sequencing analysis (single or paired end read, read length) Data analysis strategy Quotes, statements of work Contact and payment information 	
Lab Managers Lab Techs	Library Preparation	Project-Specific Experimental Details <ul style="list-style-type: none"> Sample identity Library strategy (genomic, mRNA-seq, ChIP-seq, mate paired, indexed) Average fragment length Gel images Quantitation results Quality measurements 	Information to Ensure and Improve Quality <ul style="list-style-type: none"> Kit versions Reagent lot numbers Protocol information Number of PCR cycles
Lab Managers Lab Techs	Cluster Generation	<ul style="list-style-type: none"> Sample loading pattern Kit versions Reagent lot numbers Protocol information Flow cell ID 	
Lab Managers Lab Techs	Sequencing	<ul style="list-style-type: none"> Location of data on network Kit versions Reagents lots numbers Protocol information Flow Cell ID 	
Bioinformaticians Lab Managers Technicians	Primary Data Analysis	<ul style="list-style-type: none"> Run quality metrics (%PF, first cycle, intensity, cluster density, etc.) Base calling algorithm Demultiplexed reads 	
Bioinformaticians PIs/External Collaborators Lab Managers	Secondary and Tertiary Data Analysis	<ul style="list-style-type: none"> Assembly/alignment algorithms Algorithm parameters Location of output result files (SAM/BAM files) Summary tables of SNP counts, InDels, etc. 	
Bioinformaticians PIs/External Collaborators Lab Managers	Results Reporting & Invoicing	<ul style="list-style-type: none"> Collating all work performed on samples Summarizing results and quality metrics Invoicing for work performed 	

Table 2.

Selecting a LIMS for the Next-Generation Genomics Lab

and validate experimental runs. It also speeds downstream analysis by ensuring that a sample's history and origin are tied directly to the results obtained (see Table 2).

Lab information and data management systems were first offered commercially in 1982, and today this mature class of software is being adapted to assist sequencing labs. Clearly, though, the unprecedented throughput, experimental complexity, and changeability associated with next-generation sequencing create unique challenges for traditional LIMS. The rapid timescales associated with sequencing require systems that can be quickly and easily configured to accommodate the specific instrumentation chosen by a lab. Bioinformaticians or scientific programmers must be able to easily adapt the system themselves to support changing technologies and protocols, either through on-the-fly configuration changes or scripting that programmers can undertake using application program interfaces (APIs). Finally, next-generation sequencing requires iterative, collaborative work that is performed by many different types of scientists. User-specific interfaces can ensure that these workers have access to all the information and only the information

they need to do their jobs effectively.

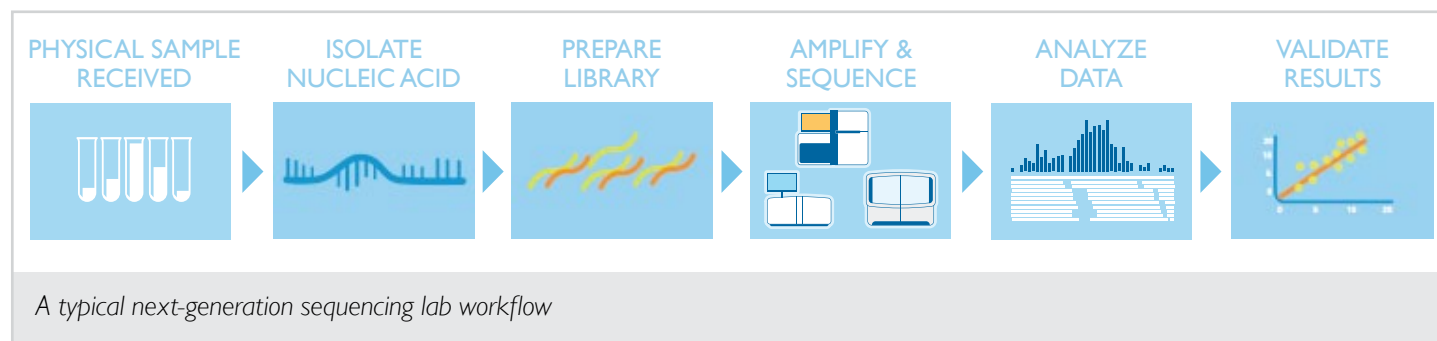
There are various parameters to be considered when determining which LIMS is best for a given type of lab or research organization. The rest of this paper discusses the three main criteria that organizations should consider in evaluating a LIMS: how well the system supports best practices in instrument configuration out of the box; how easy the system is to configure and customize; and whether the system provides user specific interfaces to streamline the work performed by the various types of users who will need to interact with the LIMS.

Selection criterion #1: Does the LIMS enable labs to get up and running quickly?

Most labs recognize the value of lab and data management. To have any hope of parsing and acting on the enormous quantities of data produced by next-generation sequencing instruments, scientists minimally need the ability to assign unique IDs to samples, record information associated with each sample, and track this information across the experimental lifecycle. Gone are the days

when scientists could track experiments using a white board or labs could easily manage library preparation and instrument runs using Excel spreadsheets, Google docs, and paper lab notebooks. Data management and experimental tracking is even more difficult for labs using DNA indexing (also known as barcoding or tagging) to pool and multiplex samples from diverse, unrelated sources on a single flow cell lane. These techniques have in fact created a bottleneck at the library preparation step, where sheer throughput combined with the need to track which samples have been pooled in which runs delays the rate at which labs can get samples onto sequencers.

One of the primary selling points of LIMS since its inception has been its ability to integrate with laboratory instrumentation. As recently as 2009, LIMS users across all industries cited instrument integration as the most desired capability in a LIMS, with fully 70% of academic laboratories managers ranking it number one.¹³ In next-generation sequencing, however, LIMS must do more than simply run and interface with instrumentation—it must provide a framework to appropriately capture data and streamline and automate mundane,



Selecting a LIMS for the Next-Generation Genomics Lab

A LIMS gives organizations confidence in sample handling while saving time and removing tedious tasks from the overall next-gen workflow.

routine tasks to eliminate the bottlenecks that can slow or even stall sequencing workflows and analysis.

While each type of next-generation sequencing instrumentation comes with vendor-specified kits and protocols to optimize use and performance, the task of integrating with sequencing instrumentation encompasses three primary phases, each of which should be supported out of the box by a LIMS.

First, organizations must consider how they are collecting information about samples and associating them with runs. Traditionally, scientists have spent many hours poring over Excel spreadsheets to check sample preparation and run assignments. A LIMS, however, can automate the process of setting

up a run. Scientists simply specify the samples they wish to run, and the LIMS automatically generates the appropriate files for the lab's sequencing equipment. Conversely, next-generation sequencing instruments can be configured to hand-off information on completed runs directly to the LIMS, reducing hands-on time for lab staff. How the LIMS integrates with the instrumentation may differ; some LIMS may integrate more tightly with particular instrumentation, and organizations should verify the connectivity between LIMS and their preferred instrumentation. But any next-generation sequencing LIMS should provide some level of integration with major next-generation sequencing instrumentation. A well-designed LIMS gives organizations confidence in sample handling while saving time and removing error-prone and tedious tasks from the overall next-generation sequencing workflow.

The second phase of integrating instruments with a LIMS is configuring the LIMS to track the quality of sequencing data coming off instruments. Many sequencing instruments run for days on end, making it wasteful and inefficient for organizations to wait until runs are completed before evaluating the quality of the data obtained. In addition to monitoring the status of runs in progress, LIMS can also collect metrics, such as the total bases yielded from a run or the percentage of base calls with a PHRED quality score of more than Q30. Over time, these metrics can aid in assessing instrument performance. With data from sample runs archived and searchable in a centralized LIMS, labs can make



Illumina HiSeq 2000

Selecting a LIMS for the Next-Generation Genomics Lab

The best LIMS also provide simple ways to create automated workflows to demultiplex reads or generate sample sheets.

better, more informed decisions about which samples to rework, whether to request more samples for further experimentation, or how much time to spend on further analysis.

The final consideration in integrating instrumentation effectively with a LIMS is results tracking. Most labs today have accumulated massive directories on their local area network dedicated to storing information associated with sequencing runs. Often this granular information surfaces in reports and summaries, while the underlying information is stored for future reference. Unfortunately, locating necessary detail can take staff hours or even days, leading some labs to re-run experiments rather than sift through directories for archived files. Multiplexing can also require an additional data management step: in some cases, those pooled samples must be “unpooled” or “demultiplexed” before the results can be analyzed and interpreted.

A LIMS can eliminate some of the most tedious aspects of next-generation sequencing for lab managers and bioinformaticians. Intuitive query tools enable labs to quickly collect information on sequencing runs, whether it was obtained last week or last year. The best LIMS also provide simple ways to

configure the system to create automated workflows to handle such tasks as demultiplexing reads, generating sample sheets for the sequencing instrument, or incorporating specific open source and commercial analysis pipelines. Freed from the need to sort through and organize data, lab staff can spend more time on analysis, decision making, publications, and innovation.

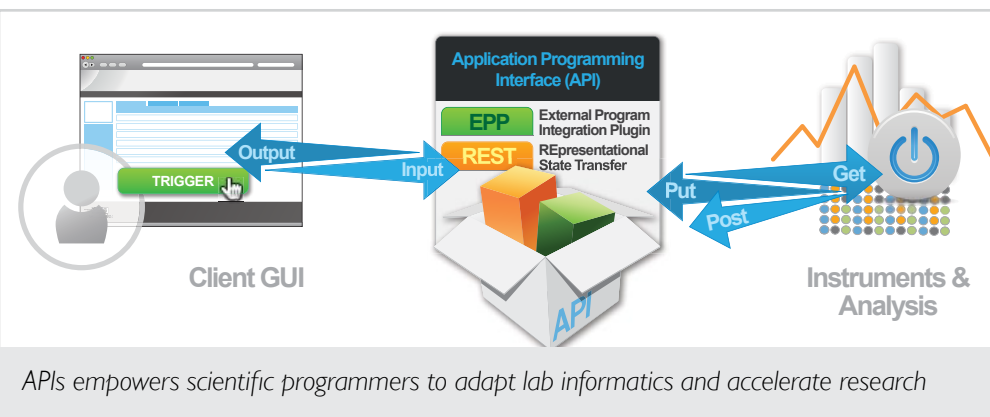
Selection criterion #2: How easy is the LIMS to configure and customize?

In cutting-edge research like next-generation sequencing, change is the operative word. Methods used one day are practically obsolete the next. Methods might not even exist for particular

applications; labs often develop and tweak protocols on the fly to handle the tasks given them by funding agencies or collaborators. In this environment, labs succeed by pushing the boundaries of innovation—and they cannot afford to be constrained in their vision by the software they implement to manage data and workflows.

No software package can effectively meet the needs of every lab—particularly in a field that’s evolving as quickly as next-generation sequencing. As a result, many labs often consider building their own lab and data management systems. Admittedly, home-grown systems do enable labs to design and implement exactly the system they want. But most labs fail to consider that when workflows and needs inevitably change, the system will also need to change quickly—and such change requires a critical investment of time, money, and personnel. Does a leading-edge next-generation sequencing lab also want to become an expert in software design and development?

While a commercial LIMS is a logical investment for labs that want to get up and running quickly, next-generation



Selecting a LIMS for the Next-Generation Genomics Lab

Using the API, programmers or bioinformaticians can tailor a LIMS to handle a range of tasks beyond those available out of the box.

sequencing requires adaptability and extensibility that can challenge commercial solutions. Several commercial LIMS have been specifically designed for next-generation sequencing, but these systems can be rigid and prescriptive about how work proceeds—and changes to the out-of-the-box configuration are discouraged and often impossible. Labs also have the option to work with broad enterprise LIMS vendors to build tailored systems using a combination of custom components coded specifically for the requesting lab and components developed by the vendor for other customers. These components, however, nearly always need to be custom designed to create a complete next-generation sequencing solution for each new customer. This approach not only makes the initial implementation costly and slow, but ensures that when a lab's needs change (and in next-generation sequencing, change is guaranteed), the vendor will need to update the system.

Effective commercial LIMS should not require labs to call up the LIMS vendor every time they want to adapt or improve the software. Software should instead be configurable and customizable by the lab team. Unfortunately, the terms “configuration” and “customization” are

often conflated, particularly in software marketing. In software engineering, configuration refers to changes in existing software that can be made via the user interface without requiring any additional programming or changes to the underlying code. In a next-generation sequencing LIMS, configuration might be used to connect the system to preferred instrumentation, capture results, and set up the system to support general sample preparation and tracking.

Customization is when the actual code must be changed so that the software can do something new or different. Customization can have negative connotations, often implying additional (expensive) consulting services or other programming assistance that software vendors may sell in addition to out-of-the-box software. Practically, however, customization should be something that lab can undertake themselves armed with the appropriate expertise (programmers), software tools, and application programming interfaces (APIs) from their LIMS software provider.

While many LIMS offer programming interfaces, most of these can be constraining, particularly if they are built on proprietary code or scripting environments. An API should instead

accommodate both open-source and commercial bioinformatics tools, particularly scripting languages such as Groovy, PERL, or Python that are familiar to scientific programmers and bioinformaticians. Using the API, programmers or bioinformaticians can tailor a LIMS to handle a range of tasks beyond those available out of the box. Scripts can be developed to integrate and automate the system, for instance, to interface the LIMS with robotics or instruments, collect information into a LIMS, or initiate computational processing tasks. Some of these customizations are exceedingly complicated and powerful, yet can be developed and deployed entirely by programmers or bioinformaticians within a sequencing lab and without any vendor assistance.

A flexible and powerful API enables labs to make a commercial LIMS their own, on their own. It also offers flexibility that can enable a vendor to customize the system rapidly should a lab lack the resources to create its own scripts or demand more extensive system engineering. Ultimately, a flexible API benefits both the purchasing lab and the supplying vendor by ensuring that changes—whatever the scope—can be made swiftly and efficiently.

Selection criterion #3: Does the LIMS accommodate different users and workflows?

Next-generation sequencing labs require varied expertise to accomplish their objectives. Principal investigators, lab

Selecting a LIMS for the Next-Generation Genomics Lab

managers, lab technicians, scientists, bioinformaticians, and scientific programmers all contribute to keep experiments running quickly and efficiently. Additionally, many labs also provide “sequencing services” to other labs or collaborators, sometimes as part of a grant’s mandate and other times as a source of additional funding. This means that labs must also coordinate communications with external collaborators. All individuals have different responsibilities and priorities and correspondingly different ways that they wish to view and act on data.

For example, lab technicians are most interested in finding out what projects they need to work on now and what other work is happening in the lab that will impact their routine. They are interested in planning their schedule and being able to quickly record information pertinent to laboratory tasks. To be useful to lab technicians, a LIMS must simplify their work by offering a straightforward user-specific interface that minimizes repetitive tasks. This enables technicians to spend less time explaining what they have done and more time actually doing experiments.

It sounds clichéd, but in a dynamic, leading-edge next-generation genomics lab, one user interface does not fit all. To work effectively, users require access to all the information, and only the information, relevant to their job. Wading through a complicated interface only slows lab staff down. Targeted user interfaces in a next-generation genomics LIMS can help everyone in a lab work more productively. These interfaces

Targeted user interfaces in a next-generation sequencing LIMS can help everyone in a lab work more productively.

should provide a dashboard of relevant activities while also pulling appropriate data from the larger system and displaying it to those who need to act on it. Intelligent, targeted user interfaces can aid the following types of users:

Lab technicians

Scientists and technical staff require fast, efficient access to data that helps them track sample status, determine which samples can be prepared together, simplify creation of library pools for multiplexed sequencing runs, and access and review past work. Dashboards should help them answer such questions as, “What experiments do I need to carry out today?” “What work is coming my way so I can plan ahead?” “Which libraries can I pool together for a multiplexed run?” or “Am I getting good quality data off that run I just started?”

Lab managers

It’s close to impossible for lab managers to keep tabs on every activity occurring in a dynamic next-generation sequencing lab. Management dashboards should provide a high level summary of everything happening in the lab—an overview of active project status and

instrument performance with the ability to drill down into activities to look at more specific results or metrics. Managers also require reporting and project management tools to manage client communications, invoicing, and administrative reporting. Interfaces should help managers answer such questions as, “What the quality of data coming off that new sequencer?” “What’s the status on the project we’ve been running for our new collaborator?” or “Where are the results from that experiment we did six months ago?”

External collaborators

A secure portal ensures that outside collaborators have immediate access to data relevant to their projects, while protecting the broader project data accumulated by the servicing lab. The portal should provide a centralized way for collaborators to initiate work requests, inquire about project status, and view project summaries. Through the interface, collaborators should be able to answer such questions as, “Is my project finished yet?” “Are there any results available to download?” or “I’ve got some additional details to provide—How can I get them to you?”

Selecting a LIMS for the Next-Generation Genomics Lab

Selecting the right tool for the task

Next-generation sequencing may be the newest, hottest technology in the life science space, but the software needed to manage and communicate next-generation sequencing data is mature and well understood. LIMS approaches have proven themselves across a range of industries for over 30 years. Many options exist—what system is best for a given facility will depend on that facility's size, scope, and research goals. The unique

demands of next-generation sequencing, however, make certain issues imperative. Can the LIMS help you get up and running quickly—out of the box, with minimal to no vendor intervention? How easy is it to adapt and augment the system when your needs—inevitably—change? And does the LIMS provide actionable information to specific users so that they can do their jobs better and faster? A thorough examination of these questions will help organizations select a LIMS that meets their lab and data information management needs now and in the future.

www.genologics.com

Telephone: 1-866-457-5467

Email: info@genologics.com

Footnotes

1. Garrity, GM, et al. *Studies on Monitoring and Tracking Genetic Resources*. Report from Convention on Biological Diversity, 2-8 April, 2009, Paris.
2. Karow, J. *The Cost of Sequencing a Human Genome? Answers Differ, Even for the Same Platform*. *Genome Web In Sequence*, September 15, 2009. <http://2429-genomeweb.voxcdn.com/sequencing/cost-sequencing-human-genome-answers-differ-even-same-platform?page=2> (Accessed December 14, 2010).
3. Stein, L. D. *The case for cloud computing in genome informatics*. *Genome Biology* 2010, 11, 207.
4. Garrity, GM, et al. *Studies on Monitoring and Tracking Genetic Resources*. Report from Convention on Biological Diversity, 2-8 April, 2009, Paris.
5. Batley, J.; Edwards, D. *Genome sequences data: Management, storage, and visualization*. *BioTechniques* 2009, 46, 333-336.
6. Martinez, D. A.; Nelson, M. A. *The Next-generation Becomes the Now Generation: Terabyte has become the New Gigabyte*. *PLoS Genetics* 2010, 6.
7. Stein, L. D. *The case for cloud computing in genome informatics*. *Genome Biology* 2010, 11, 207.
8. J.P Morgan Equity Research. *Next-generation Sequencing Survey 2010*.
9. Holt, R. and Jones, S. *The new paradigm of flow cell sequencing*. *Genome Research* 2010, 18 (6).
10. Muken, R., et al. *Next-Gen Sequencing 201: Will A New Generation Create A Democracy or Oligarchy?* Deutsche Bank 17 May 2010.
11. Swanson, J. *Scaling Up Exome Sequencing with a Lab and Data Management System*. XGen Congress 2010, San Diego, CA.
12. Karow, J. *The Cost of Sequencing a Human Genome? Answers Differ, Even for the Same Platform*. *Genome Web In Sequence*, September 15, 2009. <http://2429-genomeweb.voxcdn.com/sequencing/cost-sequencing-human-genome-answers-differ-even-same-platform?page=2> (Accessed December 14, 2010).
13. *2009 Worldwide Survey of LIMS Users: Understanding Market Trends and End-User Attitudes*. Strategic Directions International, Los Angeles (2009).
14. *Genomics: High Throughput Sequencing, Next Gen Sequencing Buzz*, <http://ngsbuzz.blogspot.com> (Accessed March 7, 2011)

All of the product names may be trademarks or registered trademarks of their respective holders.

© Copyright 2011, GenoLogics Life Sciences Software Inc.