

# Controls let genomics experimenters drive with a dashboard

ERCC 2.0, spike-in controls and other metrics in genomics and metagenomics.

Vivien Marx

If your car lacks a dashboard, you have no way of knowing when gas is about to run out or if you are traveling above the speed limit. Without controls, experimenters are driving without a dashboard, says Marc Salit, who's long been involved in standards development<sup>1-4</sup>. He has set up the Joint Initiative for Metrology in Biology (JIMB) at Stanford University's SLAC National Accelerator Laboratory. JIMB includes researchers from 11 Stanford departments and is being expanded to include scientists at other institutions and companies. Salit recently left his post at the National Institute of Standards and Technology (NIST) to establish JIMB, which NIST funds. The goal is to foster a consensus for standards and controls in high-throughput genomics and metagenomics. This will help biologists offer more than a furrowed brow when asked, "How do you know it's not an artifact?" In his own experiments, Salit says, "I have seen myself be able to be fooled, thinking I have done everything right, but I've missed something."

Consensus isn't built quickly. But the alternative is a "sort of Wild West situation" in which it's hard to compare datasets and troubleshoot experiments, says Bob Setterquist, whose lab at Thermo Fisher Scientific focuses on extraction and sample prep of DNA, RNA and proteins and



Controls help with addressing the question, "How do you know it's not an artifact?" Credit: Mehmet Hilmi Barcin/E+/Getty

who codeveloped External RNA Controls Consortium (ERCC) standards (see Box 1). The 'ERCCs' are 96 transcripts 250 to 2,000 nucleotides in length with similar GC content as human RNAs. He and others are

keen on Salit's push toward "ERCC 2.0," an updated version of the ERCCs. "We need measurements we can trust," says Salit. As consensus-building moves ahead in fits and starts, academic labs and companies apply diverse approaches (see Box 1), which also deliver some cautionary tales.

## Box 1 | A few types of controls

**ERCC RNA Spike-in Mix** is a set of 92 RNA controls from the set of 96 developed by the External RNA Controls Consortium. It's sold by vendors such as Thermo Fisher Scientific and Lexogen.

**ERCC 2.0** is a set of standards and controls in the making, propelled by academics and vendors and advanced by the Joint Initiative for Metrology in Biology (JIMB) at Stanford University.

**Spike-in controls for sequencing, or sequins**, developed at the Garvan Institute of Medical Research, are synthetic sequences without homology to a known genome.

**Spike-in RNA Variant Controls (SIRVs)** are 69 synthetic transcripts developed and sold by Lexogen.

**Armored RNAs** are RNA spike-in controls packaged in noninfectious viral particles and sold by Asuragen.

ATCC and Zymo Research sell **characterized mock microbial communities** for use as controls in microbiome research.

## ERCC 1.0

NA12878 is the unassuming name of a genomic DNA control used in many labs. It's the sequenced genome of a Caucasian woman, a Utah resident who consented to have her genome used for research purposes. The genome is an integration of 14 datasets from five sequencing technologies, seven read mappers and three variant callers. "It's probably the most sequenced genome in the world," says Timothy Mercer, a computational biologist at the Garvan Institute of Medical Research in Sydney. Her genome belongs to the standards released by the international, public-private

Genome in a Bottle Consortium advanced by NIST<sup>4</sup>. Together these genomes are types of ‘measuring sticks’ shared with the research community to help move human genome analysis into clinical applications. An issue with these measuring sticks, says Mercer, is that they nudge labs to look at genomics “through the prism of this single genome.” Selecting a genome standard means that all subsequent comparisons and analyses depend on that initial decision, he says. The Genome in a Bottle set has been expanded to include references that reflect additional ethnic backgrounds. Reference standards must evolve, and that’s especially true for analysis of the “crazily complicated” transcriptome, says Mercer.

Deep sequencing is a way to battle sequencing artifacts, and it might seem to obviate time-consuming controls. But, given today’s large-scale studies, standards and controls “provide confidence that your measurement system is working the way you want it to, the way vendors say it should and the way people expect it to,” says Setterquist. When studying a low-frequency oncogene variant that might be clinically relevant, “you better be sure that’s not an error,” he says. Controls add an extra layer of data confidence and raise a red flag when they do not work properly. Until a researcher knows what’s amiss, data are not trustworthy. His lab uses controls when monitoring effects of



Without controls and standards, it’s hard to compare datasets and troubleshoot experiments, says Bob Setterquist.

procedural changes such as how a different reagent affects nucleic acid extraction yield in sequencing library prep. When labs set up a new instrument, results might deviate from what was, and controls help to assess the deviation. “That’s just good science,” he says. He codeveloped the ERCCs, the ERCC spike-in mix, while at Ambion, a University of Texas spinout. Ambion was bought by Applied Biosystems, which was acquired by Life Technologies, which became part of Thermo Fisher Scientific. ERCC members hailed from academic labs and several companies. The spike-in mix is now used for different types of experiments such as RNA-sequencing (RNA-seq).

The controls emerged to address challenges labs faced when comparing results generated with microarrays from different manufacturers, says Setterquist. Labs were interrogating the same genes but with different probes. In making ERCCs, when short stretches hybridized to real genes in the array, they tweaked the RNAs. Companies also developed ERCC-related analysis software. As high-throughput sequencing emerged, labs began using the ERCC controls for RNA-seq and single-cell analysis, which they are not optimized for. In collaboration with several institutions in Seattle including the Fred Hutchinson Cancer Research Center, scientists at 10x Genomics did transcriptional profiling of around 250,000 blood cells from two people<sup>5</sup>. They modified the standard ERCC spike-in protocol to use it in their droplet-based system for single-cell mRNA quantification. Without assessing genotype, the team determined cell origin on the basis of variants identified in single-cell RNA-seq data. It’s shaped their view of ERCC.

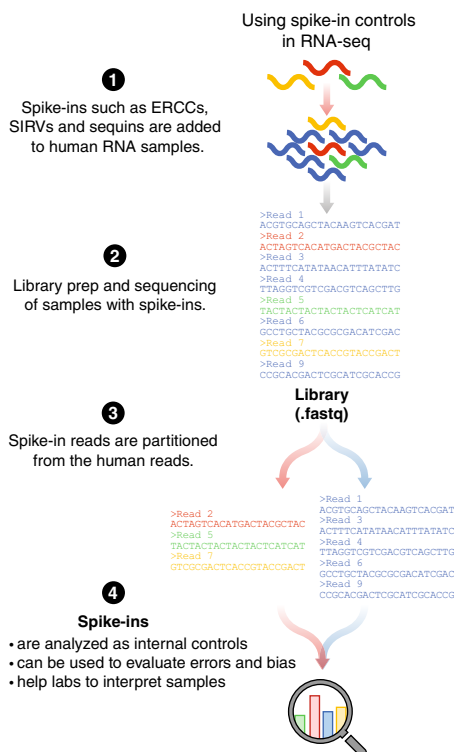
“We don’t have any specific criticisms of using ERCC in bulk RNA-seq experiments, but for use in single-cell experiments such as with 10x’s products, we do not recommend their use as a standard control,” says Michael Schnall-Levin, who directs product, R&D and strategy at 10x Genomics. In their platform, the spike-ins spread across the system’s partitions, including those lacking cells. “So you wind up wasting a fair amount of data from these empty partitions, and

you don’t get any feedback on the single-cell aspects of the system,” he says. For the 10x single-cell analysis platform, the ERCC spike-ins are not optimal, in the same way that it’s not optimal to measure an irregularly shaped dining room table with a meter stick, says Salit. “It’s the wrong tool.” Schnall-Levin and his team generally support standards development for emerging single-cell sequencing technologies such as theirs. They apply standards to guide technical decisions, both for their product’s physical and software aspects and to characterize their product’s performance. “We like when standards are available for comparing available technologies,” he says, which enables consistent comparison that helps them stay competitive in the marketplace.

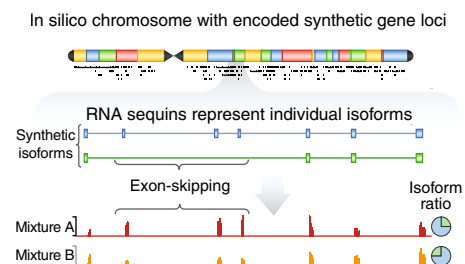
The rumblings about ERCC 2.0 have lasted for a few years but “we were never able to bring that off,” says Salit. Now, his renewed effort is under way to build the needed team and alliances. Setterquist looks forward to ERCC 2.0. “I think companies really do have a responsibility to be involved in some of that,” he says. At ERCC’s core was a “really nice collaboration” between companies and academic labs. Plasmids were sent around between the partners GE, Affymetrix, Agilent, Illumina and Ambion as they decided on the final mix, he says. They discussed such aspects as poly(A) tail lengths, whether RNAs should be capped and the manufacturing steps. Discussions have continued, for example, about inclusion of microRNAs and what to do about capped RNAs. In the past they decided against capping, says Setterquist, because doing so synthetically is challenging, as is accurately measuring the percentage of capped RNAs in the control mix.

### Spreading the tools

Labs develop and distribute controls to others. Mercer, who is also a researcher at Altius Institute for Biomedical Sciences, developed ‘sequins’, which stands for ‘spike-in controls for sequencing



Adapted from ref. 7. Credit: T. Mercer



With synthetic sequences, or sequins, a mix of synthetic isoforms at different concentrations helps labs emulate alternative splicing. Adapted from ref. 6. Credit: T. Mercer



Lexogen's internal bioinformatics team triggered spike-in development to tune their RNA-seq pipeline, says Lukas Paul.  
Credit: P. Knibbe

experiments<sup>6</sup>. Sequins are entirely synthetic and their sequence is not found in people or any other organism. They can be used in many experiment types, says Mercer, such as when labs check whether their setup can find fusion genes or when they are hunting alternative splice sites. Sequins capture more of the transcriptome's complexity than the current ERCC spike-ins, he says.

The Garvan plans to use sequins internally for analyses of clinical genomics data, says David Barda, who directs the Garvan's business development and innovation. Reliable reference controls let researchers aggregate human genomic data from multiple sources and improve genotype–phenotype correlations. The need for reference controls will surface in areas such as metagenomics and immune repertoire profiling. “Accordingly we are developing sequins for these applications, too,” he says. Mercer also looks forward to ERCC 2.0. Even if consensus takes a while, controls are crucial. In the fast-moving high-throughput sequencing field, few labs have had the time to consider standards. “In the beginning, people don't really care about standards, people are just doing the biology and getting out the papers,” he says. Standards discussions begin as a field matures, as many labs enter a field and data mountains await<sup>7</sup>. In his observation, standards discussions with RNA emerged earlier than with DNA.

Lexogen, a biotech company in Vienna, sells ERCC controls and also in-house-developed RNA spike-in controls and analysis software. The ERCC spike-ins are good for dose–response correlation experiments, and their unique sequences help with testing an assay's sensitivity, says Lukas Paul, who manages the company's scientific collaborations. But ERCCs do less well with isoforms that are typical of eukaryotic genomes. The ERCCs lack spliced exons and alternative start-site exons, for example. “The whole complexity of the transcriptome wasn't mirrored in the ERCCs of the first generation,” says Paul. It was the company's internal bioinformatics team that triggered spike-in development to

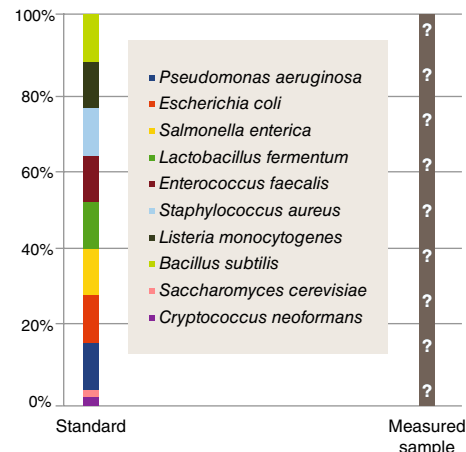
tune their RNA-seq pipeline with a kind of “external truth,” he says: well-defined spike-ins in terms of sequence and concentration.

The company began selling its Spike-in RNA Variant Controls (SIRVs) after beta-testing them and the bioinformatics tools with academic labs. SIRVs are 69 synthetic transcripts derived from human genes but with different sequences, and they contain tough-to-detect isoforms. “We have some really nasty ones in there,” says Paul. They differ by just a few nucleotides and would likely be missed in short-read RNA-seq experiments. SIRV modules are available in different molar mixtures and there is a combined set of SIRVs and ERCCs of which they are licensed distributors. Lexogen intends to be involved with ERCC 2.0. One Lexogen customer combined SIRVs with Pacific Biosciences and Oxford Nanopore long-read technology to get an inventory of isoforms and then used short-read high-throughput sequencing for transcript quantification, says Paul. Given Illumina's acquisition of PacBio, he believes SIRVs will be useful if a new sequencing technology emerges that combines short- and long-read sequencing. A group of Oxford Nanopore users have applied SIRVs in direct RNA-seq experiments, such as to achieve isoform-level analysis of NA12878 native poly(A) RNA sequence reads. Another class of SIRV-users do single-cell sequencing experiments, says Paul. In such instances, spike-ins deliver quick “technical feedback” about dataset quality. For that technical check, a lab doesn't need to first analyze all its single-cell data, he says.

The company's modules let labs separate various types of sensitivity-detection applications from ones in which isoform detection is key. They might explore their samples with several SIRV mixtures to find the differential between samples with highly abundant isoforms and a few rare variants, says Paul. He suggests labs keep in mind that for gene expression quantification, they need to count the “right” isoforms. More SIRV modules are in development at the company, such as ones that include different poly(A) lengths, capped RNAs, microRNAs,



Mock microbial communities help labs see, “Did I get out what I put in?” says Mindy Goldsborough.  
Credit: ATCC



Characterized mock microbial communities can serve as experimental standards. Credit: C. Cruz, Zymo Research

variations of GC content and various types of RNA modifications. Lexogen plans to accelerate this work with grants and academic partnerships. Potentially, SIRVs can be used for clinical applications for monitoring gene expression changes, and spike-ins can be a kind of sample “fingerprint.”

Other types of controls are emerging such as armored RNA, which Asuragen sells. RNases cannot get to the RNAs that are packaged inside the capsules of noninfectious viruses. This stable packaging won't degrade, says Setterquist, and can, for example, be spiked into a blood sample for analysis. Some labs feel they lack the time to be involved in standards development, he says. Others prefer sticking to their own controls. “It's not such that everybody has to be using the same one,” he says. A small lab might make a batch and freeze the rest. He recommends carefully characterizing such controls before use. Labs running large-scale projects may find a commercially manufactured and quality-controlled product to be a time-saver.

### Microbiome standards

Controls are still emerging in the microbiome field. Zymo and ATCC, for example, sell mock communities with published microbial ratios as controls. ATCC offers mock microbial communities of either whole cells or purified nucleic acids so labs have spike-ins for preparing sequencing libraries, for amplification or other types of analysis. The standard comes with a software license key for data analysis tools, so labs can see “did I get out what I put in?” says ATCC's chief science and technology officer Mindy Goldsborough. The controls involve variety—for example,



Gram-negative and Gram-positive microbes. “They all react differently to being ruptured in sequencing,” she says.

ATCC’s first mock community was general; in the works are ones geared toward different microbial populations: human skin, the oral cavity, vagina. To date, ATCC has seen uptake of these controls in the research community and with platform manufacturers who, for example, use controls to test their extraction kits.

At Zymo Research, the ZymoBIOMICS Microbial Community Standards grew out of a concern about discrepancies in published microbiome profiling and metagenomics studies. For example, in one instance there was divergent analysis of one fecal sample by two microbiome projects, American Gut and uBiome, says Shuiquan Tang, a Zymo Research scientist focused on microbiomics. Another example: different fecal DNA extraction protocols used by the Human Microbiome Project and the Metagenomes of Human Intestinal Tract analysis of one fecal sample led to “big variations” in the relative abundance of Bacteroidetes, a major phylum in the human gut microbiome<sup>8</sup>. “Because of these, I would say researchers should be extremely careful about comparing microbiome data across labs or projects,” he says. Zymo’s controls are documented and characterized mock microbial communities. One is a whole-cell version, and the other is isolated DNA. The controls do not mimic a certain sample type such as soil or feces. These are microbial mixes of bacteria and several yeast strains that include a variety of GC content, as well as Gram-positive and Gram-negative strains, says Emily Putnam, who is assistant manager of the company’s services team, which handles contract research. Internally, the company uses its own controls. The company also offers a mock community with a log abundance distribution for labs assessing detection limits. “That’s been titrated to allow to check for even very minimal levels of bacterial presence,” she says.

The mock communities include tough-to-lyse and easy-to-lyse microbes. DNA extraction approaches can be biased toward the easy-to-lyse microbes such as Bacteroidetes, says Tang. His company is developing additional standards to address the range of measurement challenges in microbiomics. Spike-in controls for quantification in microbiome studies are also in the works. “Unfortunately, interpreting sequencing data derived from the microbiome standards is not trivial,” says Tang, who helps customers with thorny issues. For example, one lab used the standards to compare different DNA extraction kits. Unlike the customer, Tang



Dashboards offer a systematic assessment of technical performance. They apply in transportation and in biology. Credit: Paulo Sousa/EyeEm/Getty

found none of those kits had satisfactory results. As it turns out, the customer’s mechanical lysis approach was introducing bias. The company is setting up online services to help customers analyze their results when using standards. ERCC spike-in controls appeal to Tang as a way to control RNA analysis from library prep all the way to bioinformatics analysis. Microbiome standards should have multiple formats: whole cells, DNA or RNA. “But whole-cell controls definitely play an important role here, because uneven microbial cell lysis during DNA extraction accounts for a lot of variations observed in this field,” he says. Reproducibility issues will not be solved by standardizing workflows or committing labs to the same workflow, says Tang. “I think it is too risky to talk about standardization in such a young field,” he says. High-throughput sequencing has let microbiome studies and technology advance. It will help the field establish guidelines about good practices, such as the inclusion of negative controls and positive controls that can be used to optimize workflows and improve measurement accuracy, he says.

### Metrology, biology

Now that scientists can look at cells with “atomic granularity,” says Salit, biology labs can celebrate “living in paradise where there’s much to learn.” But the life of biologists in paradise should involve the principles of metrology, such as traceability, which is the ability to report data, their ability to characterize measurement uncertainty and their ability to validate their method. These factors will help them, for example, trust the way they can compare disease to wild-type. ERCCs do not place a lab’s data on a calibrated axis, nor does any

lab report results as “gene x is expressed at ‘4 times ERCC44.’” They are controls, as are Mercer’s sequins and Lexogen’s SIRVs, that can be placed on a dashboard<sup>3</sup> of performance measures, he says.

A car’s dashboard with its readout of speed, for example, is a systematic assessment of technical performance. In biology, a dashboard tells a lab how well a given experiment is going and helps to assure artifacts are not coloring results. Given the limitations of ERCC 1.0, ERCC 2.0 will be new types of controls, says Salit. “Should we make new molecules? Absolutely,” he says. They need to not interfere with endogenous molecules and must work across all technologies. In developing ERCC 2.0 and the companion analytical tools, he wants everyone at the table. “I absolutely need to mix up .com, .edu, .gov in the same room,” he says, to get new controls developed and to create force for adoption in the community. The age of high-throughput genomics, he says, is in some ways akin to the early days of the automobile when people drove without a dashboard and their engine boiled over because they had no inkling of their engine’s temperature. High-throughput genomics needs dashboards and standards as they have long existed in chemistry and physics. In those fields, standards and controls have taken time.

For example, the mole was not included in the International System of Units until 1968. Since the original definition of the gram in 1795, the kilogram has been repeatedly redefined. The kilogram has long been a physical object: a cylinder of iridium and platinum sitting in a vault in the Louis XIV Pavillon de Breteuil of the International Bureau of Weights and Measures on the outskirts of Paris. Salit recently attended the nearby weighty vote at the General Conference on Weights and Measures to redefine the kilogram in terms of physical constants. “My goodness, if we’re still figuring out the kilogram, let’s talk about the transcriptome,” he says. □

Vivien Marx

Technology editor for *Nature Methods*.  
e-mail: [v.marx@us.nature.com](mailto:v.marx@us.nature.com)

Published online: 20 December 2018  
<https://doi.org/10.1038/s41592-018-0265-y>

### References

1. Baker, S. C. et al. *Nat. Methods* **2**, 731–734 (2005).
2. Salit, M. *Methods Enzymol.* **411**, 63–78 (2006).
3. Munro, S. A. et al. *Nat. Commun.* **5**, 5125 (2014).
4. Zook, J. M. et al. *Nat. Biotechnol.* **32**, 246–251 (2014).
5. Zheng, G. X. et al. *Nat. Commun.* **8**, 14049 (2017).
6. Hardwick, S. A. et al. *Nat. Methods* **13**, 792–798 (2016).
7. Hardwick, S. A., Deveson, I. W. & Mercer, T. R. *Nat. Rev. Genet.* **18**, 473–484 (2017).
8. Wesolowska-Andersen, A. et al. *Microbiome* **2**, 19 (2014).