Sequencing benchmarked

The Sequencing Quality Control 2 (SEQC2/MAQC-IV) project provides resources to aid sequencing reproducibility and highlights factors that can guide platform and software choice.

SEQC2 is the most comprehensive evaluation of major sequencing platforms to date. It not only provides reference samples and datasets related to inter- and intra-lab reproducibility, but also identifies factors that can influence the performance of next-generation sequencing (NGS) instruments and their computational pipelines. Although the most likely near-term benefit of SEQC2 will be to encourage best practices when setting up sequencing pipelines at core centers, the legacy of its parent group, the Microarray Quality Control (MAQC) consortium, may be to serve as a template for other community-wide efforts seeking to benchmark rapidly evolving technologies.

This month, *Nature Biotechnology* publishes the SEQC2 suite of papers. The reports in this Focus, and in other Nature Portfolio journals, describe analysis protocols and quality-control metrics for NGS platforms for basic researchers and those working in clinical and regulatory settings.

MAQC was initiated in 2005 at the FDA's National Center for Toxicological Research (NCTR) as a response to the agency's Voluntary eXploratory Data Submission program and controversy surrounding the reliability of DNA microarrays in research. The program sought to assess emerging omics technologies, reach a consensus on how best to analyze massively parallel genomic data, and agree how such datasets should be interpreted in packages submitted to the agency.

Phase 1 of MAQC (MAQC-I) was published in 2006 and involved >100 researchers from six FDA centers, the Environmental Protection Agency, the National Institute of Standards and Technology, several leading microarray manufacturers, reagent and material suppliers, academic laboratories, drug companies and other stakeholders. Using two human reference RNA samples, MAQC-I assessed the precision and cross-platform and cross-laboratory comparability of microarray and quantitative RT-PCR datasets.

The results of MAQC-II were published four years later. Spurred by the FDA's need to handle product applications for genomic classifiers like Roche Diagnostic's AmpliChip CYP450 and Agendia's MammaPrint, it assessed the performance of various machine-learning and data-analysis methods in microarray-based predictive models and presented best practices for validating gene signatures representative of a phenotype or disease.

Around this time, the rapid adoption of NGS for RNA profiling prompted MAQC to again shift emphasis, this time to RNA-seq. The result was SEQC/MAQC-III, culminating in ten papers published in 2014 that investigated sources of bias and compared the performance of different RNA-seq platforms and DNA microarrays.

Now SEQC2—a final five-year effort by a coalition of >300 participants and >150 organizations—reports its efforts to benchmark sequencing platforms in several applications, including somatic and germline mutation analysis, single-cell RNA-seq, copy number variation, oncopanel sequencing and liquid biopsies of tumor samples. A complementary project by the Association of Biomolecular Resource Facilities investigates sources of bias and compares the performance of different NGS platforms using an Ashkenazi family trio, three individual bacterial strains and a metagenomic mixture of ten bacteria.

Taken together, these studies provide perhaps the most comprehensive assessment of NGS performance to date and a detailed analysis of different software options for alignment and mutation calling (in bulk sequencing) and preprocessing, normalization, batch correction and visualization (in single-cell RNA-seq). For those seeking to set up and benchmark the performance of a sequencing pipeline, SEQC2 provides standardized reference samples and model datasets, as well as information on experimental designs and spike-in controls.

MAQC has fundamentally changed the practice of genomic data analysis.

First, perhaps its most fundamental contribution—and at the time of MAQC-I its most controversial insight—was to challenge the reproducibility of omic data analyses driven solely by *P*-value magnitudes and false discovery rates. In the regulatory or clinical context, MAQC argued reproducibility (even in multiple testing methods) requires the combination of less stringent *P*-values (or false discovery rates) with minimum effect sizes specific to a particular analytical technology. For DNA microarrays, this meant proposing a minimum 1.5- to 2-fold change in gene expression as a cutoff for data in regulatory submissions.

Second, with the increasing pace of advances in biological research and analytical technology, the need for benchmarking efforts like MAQC has never been clearer. One need only look at the recent explosion of different methods for analyzing single-cell data, many of which are evaluated in differing contexts, to appreciate the difficulty researchers face in both parsing and keeping track of the plethora of changing options, the performance of which depends on many biological, experimental and technical variables. The MAQC model—comprising researchers from across academia, industry and government agencies—has proven a highly effective mechanism for benchmarking new analytical technology as it emerges.

Third, as highlighted in an accompanying Comment, insights from MAQC have directly contributed to regulatory practice. In 2007, MAQC findings were incorporated into draft FDA guidance for pharmacogenomics and in vitro diagnostics (IVDs), as well as the International Conference on Harmonization Tripartite Guideline; in 2018, they contributed to FDA guidance on the use of human genetic variant databases to support IVDs and IVDs for germline disease. In this light, MAQC has clearly had a direct impact on the practice of precision medicine.

Finally, the consortium has proven a wonderful example of community altruism and open science. Work was driven solely by small grants and contracts from national funding agencies and regulators and by in-kind contributions from companies of equipment and reagents; all involved gave their time for free, often using their own resources to pay for travel to meetings or workshops. This reflects both the community's commitment to reproducibility and the leadership and resourcefulness of NCTR division director Weida Tong, Fudan University's Leming Shi, Q2 Solutions' Wendell Jones and SAS Cary's Russ Wolfinger.

As MAQC completes its final phase, the MAQC Society takes up the gauntlet. Continued involvement of FDA leadership will be key to continued participation by academia and industry. As a host of other high-throughput technologies continue to come online and mature—artificial intelligence in clinical imaging, metagenomics, spatial transcriptomics and proteomics, to name a few—never have such benchmarking efforts been as important. ❐

https://doi.org/10.1038/s41587-021-01067-3