



Reference Materials for Improving Reliability of Multiomics Profiling

Luyao Ren¹ · Leming Shi^{1,2,3} · Yuanting Zheng¹

Received: 25 August 2023 / Revised: 18 December 2023 / Accepted: 22 December 2023
© The Author(s) 2024

Abstract

High-throughput technologies for multiomics or molecular phenomics profiling have been extensively adopted in biomedical research and clinical applications, offering a more comprehensive understanding of biological processes and diseases. Omics reference materials play a pivotal role in ensuring the accuracy, reliability, and comparability of laboratory measurements and analyses. However, the current application of omics reference materials has revealed several issues, including inappropriate selection and underutilization, leading to inconsistencies across laboratories. This review aims to address these concerns by emphasizing the importance of well-characterized reference materials at each level of omics, encompassing (epi-)genomics, transcriptomics, proteomics, and metabolomics. By summarizing their characteristics, advantages, and limitations along with appropriate performance metrics pertinent to study purposes, we provide an overview of how omics reference materials can enhance data quality and data integration, thus fostering robust scientific investigations with omics technologies.

Keywords Reference materials · Reference datasets · Quality control · Performance metrics · Reproducibility · Multiomics profiling · Molecular phenomics

Abbreviations

ABRF	Association of Biomolecular Resource Facilities	GeT-RM	Genetic Testing Reference Materials Coordination Program
ATCC	American Type Culture Collection	GIAB	Genome in a Bottle Consortium
CDC	Centers for Disease Control and Prevention	HPRC	Human Pangenome Reference Consortium
CNA	Copy number alteration	HUPO	Human Proteome Organization
CPTAC	Clinical Proteomic Tumor Analysis Consortium	IMMSA	International Microbiome and Multi-Omics Standards Alliance
CRM	Certified reference material	JRC	Joint Research Centre
ctDNA	Circulating tumor DNA	LDT	Laboratory developed tests
EBV	Epstein–Barr virus	LOD	Limit of detect
ERCC	External RNA Control Consortium	LTR	Long-term reference
FFPE	Formalin-Fixed Paraffin-Embedded	MAQC	MicroArray Quality Control Consortium
gDNA	Genomic DNA	MCC	Matthews correlation coefficient
		MetQual	Metabolomics Quality Assurance and Quality Control Program
		MMR	Mismatch repair
		mQACC	Metabolomics Quality Assurance and Quality Control Consortium
		MRM	Multiple reaction monitoring
		MS	Mass spectrometry
		NCI	National Cancer Institute
		NIM	National Institute of Metrology
		NIST	National Institute of Standards and Technology
		NMR	Nuclear magnetic resonance
		PCA	Principal component analysis
		PEA	Proximity Extension Assay

✉ Leming Shi
lemingshi@fudan.edu.cn

✉ Yuanting Zheng
zhengyanting@fudan.edu.cn

¹ State Key Laboratory of Genetic Engineering, School of Life Sciences and Human Phenome Institute, Fudan University, Shanghai 200438, China

² Shanghai Cancer Center, Fudan University, Shanghai 200032, China

³ International Human Phenome Institutes, Shanghai 200438, China

PRM	Parallel reaction monitoring
PSI	Proteomics Standards Initiative
PTM	Post-translational modification
QA	Quality assurance
QC	Quality control
RM	Reference material
RMSE	Root mean squared error
RT	Retention time
SEQC	Sequencing Quality Control Consortium
SNR	Signal-to-noise ratio
SNV	Single-nucleotide variations
SRM	Standard reference materials
SV	Structural variants
T2T	Telomere-to-Telomere
TCGA	The Cancer Genome Atlas
TMB	Tumor mutation burden
TNBC	Triple-negative breast cancer
UHRR	Agilent Universal Human Reference RNA material
VAF	Variant allele frequency
WES	Whole-exome sequencing
WGS	Whole-genome sequencing

Introduction

In recent years, the adoption of multiomics approaches in biomedical research and clinical application has increased significantly (Hasin et al. 2017; Hoadley et al. 2018). The integration of multiomics or molecular phenomics data (including genomics, epigenomics, transcriptomics, proteomics, and metabolomics) along with deep phenotypic data enables the discovery of correlations between the diverse levels of genetic and regulatory information and distinct phenotypic traits, fostering a more comprehensive understanding of biological processes and facilitating the identification of disease mechanisms, potential therapeutic targets, and disease biomarkers (Jiang et al. 2019; Mangiante et al. 2023; Martinez-Ruiz et al. 2023; Sammut et al. 2022). However, challenges exist in translating scientific research findings to clinical settings, particularly regarding the reproducibility of omics data (Bell et al. 2009; Foox et al. 2021; Khayat et al. 2021; Pan et al. 2022). The complexity of biological systems and potential technical artifacts from sample preparation, data generation and data analysis contribute to this challenge, which is further amplified in multiomics data integration (Krassowski et al. 2020). Implementing rigorous quality assurance (QA) and quality control (QC) measures is crucial to ensure the reliability of multiomics research (Bittremieux et al. 2018; Broadhurst et al. 2018; Zheng et al. 2022). QA involves processes and activities to prevent errors and ensure quality standards of the final product, whereas QC comprises activities to test and inspect the final product

or service to meet quality standards (International Organization for Standardization, ISO 9000:2015).

Reference materials (RMs) are essential for both QA and QC in multiomics research (Broadhurst et al. 2018; Hardwick et al. 2017; Jennings et al. 2017; Lipka et al. 2022). RMs are well-characterized samples with known properties that can be used to validate the accuracy and reliability of analytical methods, assess the comparability of data generated by different laboratories or instruments, and serve as a standard against which the accuracy and precision of measurements can be evaluated (Bell et al. 2009; Foox et al. 2021; Hardwick et al. 2017). Although the terms “certified reference materials (CRMs)”, “standard reference materials (SRMs)”, “reference materials”, “reference standards”, “reference samples”, and “quality control samples” are often used with the same or very similar meaning related to the calibration and validation of analytical methods, there are important differences between them at the level of characterization, traceability, and certification that they offer. CRMs and SRMs are typically considered the most reliable and accurate standards for analytical measurements, while the others may have more limited or uncertain properties. According to ISO Guide 30:2015, a CRM is “a reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure that establishes its traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence” (International Organization for Standardization, ISO Guide 30:2015). In other words, a CRM is a reference material that has been thoroughly analyzed and certified by an authorized organization to have a known value for one or more properties, along with its associated uncertainty and a statement of metrological traceability. The certification process ensures that the material meets established international standards for accuracy and traceability. Official governing bodies, such as the National Institute of Standards and Technology (NIST) in the United States, the National Institute of Metrology (NIM) in China, and the European Commission's Joint Research Centre (JRC) in Europe, can provide certification for CRMs. Other accredited organizations can provide certification for CRMs as well. The term SRM is a specific term used by NIST for those meeting additional NIST-specific certification criteria in accordance with ISO Guide 31:2000 (National Institute of Standards and Technology 2023a). A RM is defined as “a material or substance one or more of whose property values are sufficiently homogenous and well established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to materials” (International Organization for Standardization, ISO Guide 30:2015). RMs can also be referred to as “reference standards”, “reference samples”, or “quality control samples”. These materials can

be prepared in-house or purchased from commercial suppliers. While RMs are generally considered to be of high quality, they may not have undergone the rigorous testing and certification required for CRMs.

Omics RMs mentioned in this review refer to well-characterized and validated samples used as quality control tools in various omics technologies. A major difference between traditional RMs and omics RMs is the number of properties values they encompass. Traditional RMs typically comprise a limited number of well-defined and characterized property values, often associated with physical and chemical attributes, placing a strong emphasis on traceability. In contrast, omics RMs encompass a significantly larger number of property values, reflecting the complex and diverse nature of biological omics data. It is important to note that there is currently no internationally recognized CRMs for massive analysis technologies, because current omics RMs do not fulfill the conventional criteria for established traceability of the assigned property values. The signals detected by omics technologies, such as DNA or RNA sequencing reads, mass spectrometry (MS) peaks, or nuclear magnetic resonance (NMR) spectra, cannot be directly traced back to the international system of units (SI units). While omics RMs may not have the same level of rigor as CRMs, they can still serve as a useful tool for quality control and method validation in omics research. Many ongoing efforts have been made to develop omics RMs. These RMs are typically prepared by reputable laboratories using standardized protocols and characterized for their stability, homogeneity, and variability, with their properties traceable to a reference measurement system. Examples of ongoing efforts to establish omics RMs include the Genome in a Bottle Consortium (GIAB) (Zook et al. 2014), the MicroArray/Sequencing Quality Control (MAQC/SEQC) (MAQC Consortium 2006; Fang et al. 2021; Jones et al. 2021), the External RNA Control Consortium (ERCC) (Baker et al. 2005), Clinical Proteomic Tumor Analysis Consortium (CPTAC) (Tabb et al. 2016), the Metabolomics Quality Assurance and Quality Control Consortium (mQACC) (Lippa et al. 2022), and the Chinese Quartet Project for multiomics profiling (Yang et al. 2023; Zheng et al. 2023).

RMs are widely recognized as essential for ensuring data quality in omics research; however, their current application is inadequate due to issues such as inappropriate selection of RMs for the intended purpose and a lack of understanding about when and how to use them (Begley and Ioannidis 2015; Bowden et al. 2018; Chiva et al. 2021; Evans et al. 2020; Zhang et al. 2020). This review aims to provide a comprehensive overview of currently available omics RMs for high-throughput technologies in omics research, with a focus on quality assessment across batches, platforms, and laboratories. We will first summarize RMs for each omics level (DNA, RNA, protein, and

metabolite), including their intended usage, advantages and limitations (Table 1). We will then describe qualitative and quantitative properties of RMs that help determine accuracy and precision. Next, we will explain quality control metrics based on reference datasets or intrinsic relationships between reference sample groups. Finally, we will describe how to use RMs to improve biomarker discovery in omics studies and discuss considerations for utilizing appropriate RMs.

DNA Reference Materials

DNA RMs are designed to assess the accuracy of genetic variant detection using high-throughput DNA sequencing technologies (Fang et al. 2021; Zook et al. 2014). These materials are available in various formats to suit different research purposes. Biological DNA RMs are typically genomic DNA (gDNA) obtained from natural biological materials, such as Epstein–Barr virus (EBV)-immortalized lymphoblastoid cell lines or tumor cell lines. Immortalized cell lines are convenient and cost-effective sources of reference materials, as they can be readily proliferated through cell culturing, providing a renewable source of gDNA (Fang et al. 2021; Ren et al. 2023; Zook et al. 2014). To ensure that RMs are sufficiently homogenous and in a large quantity to be widely disseminated, they are usually extracted from a single large batch of cell culture and well-mixed. Although subtle genetic differences may exist among cells cultured in different dishes, each vial of reference materials contains the same mixture of genomes, because the cells and gDNA are thoroughly mixed. Biological DNA RMs represent the full size and complexity of the human genome, making them ideal for benchmarking thousands or even millions of variants detected through whole-genome (WGS) and whole-exome sequencing (WES). For genetic testing targeting specific disease-causing variants, patient genomes containing these variants can be provided as valuable reference materials (Kalman et al. 2007; Li et al. 2020). Alternatively, engineered DNA RMs with specific variants introduced into the genome using genome-editing technologies can be used (Lin et al. 2022; Suzuki et al. 2020). When assessing the performance of experimental and bioinformatics processes, DNA RMs derived from natural or engineered cell lines are typically analyzed in parallel alongside study samples, whereas synthetic spike-in controls are usually added to samples of interest as internal controls throughout the entire sequencing workflow to measure technical artifacts. To distinguish them from study samples, synthetic spike-in controls are composed of non-human artificial DNA sequences or contain unique molecular barcodes (Blackburn et al. 2019; Reis et al. 2020).

Table 1 Summary of reference materials for multiomics profiling

Omics	Category	Targeted molecule	Project/institute	Name	Sample source	Public avail- ability of materials	References
DNA	Biological DNA	Whole-genome germline variants	GIAB	RM8398/RM8292/ RM8393	gDNA from lymphoblastoid cell lines of a Utah woman with European Ancestry, an Ashkenazi Jewish trio and a Han Chinese trio	Yes	Chin et al. (2020), Wagner et al. (2022), Zook et al. (2014), Zook et al. (2019) and Zook et al. (2020)
		Genetic testing for germline variants	Quartet	GBW09900/ GBW09901/ GBW09902/ GBW09903	gDNA from B-lymphoblastoid cell lines of a Chinese Quartet family	Yes	Jia et al. (2023), Ren et al. (2023) and Zheng et al. (2023)
			GeT-RM	–	Immortalized cell lines from patients with pathogenic mutations or variants on pharmacogenetic loci	Yes	Amos Wilson et al. (2008), Bettinotti et al. (2018), Gaedigk et al. (2019), Kalman et al. (2007), Kalman et al. (2011), Pratt et al. (2009) and Pratt et al. (2016)
		Whole-genome somatic variants	MAQC/SEQC	HCC1395/ HCC1395BL	gDNA from a triple-negative breast cancer and its paired B-lymphoblastoid cell lines	Yes	Fang et al. (2021)
		Whole-exome germline and somatic variants	MAQC/SEQC	SampleA/SampleB	gDNA from ten cancer cell lines and a normal male cell line	Yes	Jones et al. (2021)
		Tumor mutation burden	FCR	–	gDNA from 36 tumor-normal matched FFPE samples	No	Merino et al. (2020), Stenzinger et al. (2019) and Vega et al. (2021)
	Engineered DNA	Cancer associated variants on genomic DNA	MDIC SPOT/Dx NCCL	– – –	Cancer related variants engineered into DNA sequence of HG002 by CRISPR/Cas9	No No No	Pfeifer et al. (2022) Wang et al. (2023)
		Cancer associated variants on ctDNA	Thermo Fisher SeraCare Horizon	AcroMatrix SeraSeq Multiplex	Engineered cancer actionable mutations mixed with genomic DNA background offered in series dilution of VAF	Yes Yes Yes	Thermo Scientific (2020) SeraCare (2023a, b) Horizon Discovery (2023)
	Synthetic DNA	Germline and somatic variants	GIMR	sequins	Synthetic mirrored human DNA sequences	Yes	Deveson et al. (2016, 2019)

Table 1 (continued)

Omics	Category	Targeted molecule	Project/institute	Name	Sample source	Public avail- ability of materials	References
RNA	Biological RNA	Whole transcriptome	MAQC/SEQC	Sample A/Sample B	Total RNA mixture from ten tumor cell lines and a human brain reference RNA from 23 donors	Yes	MAQC Consortium (2006, 2014)
			Quartet	GBW09904/ GBW09905/ GBW09906/ GBW09907	RNA from B-lymphocyte cell lines of a Chinese Quartet family	Yes	Yu et al. (2023)
		Single-cell RNA-seq	MAQC/SEQC	HCC1395/ HCC1395BL	Single-cell gene expression from a triple-negative breast cancer and its paired B-lymphoblastoid cell lines	Yes	Chen et al. (2021)
	Synthetic RNA	RNA	ERCC ¹	ERCC spike-in Mix	92 artificial RNA transcripts	Yes	Jiang et al. (2011) and Munro et al. (2014)
			GIMR	Sequins	Synthetic RNA isoforms with exons and introns	Yes	Hardwick et al. (2016)
		miRNA	ERCC ²	–	Synthetic RNA oligonucleotides	No	Giraldez et al. (2018)

Table 1 (continued)

Omics	Category	Targeted molecule	Project/institute	Name	Sample source	Public avail- ability of materials	References
Protein	Biological protein	Proteins from tissue	NIST	RM8462	Frozen human liver suite consists of normal, fatty and congested liver samples	Yes	Lippa et al. (2022)
		Protein from yeast	CPTAC	–	Digested Yeast Lysate	Yes	Beasley-Green et al. (2012) and Paulovich et al. (2010)
		Protein from cell line	Quartet	Quartet Protein	Peptides digested from B-lymphoblastoid cell lines of a Chinese Quartet family + 4 peptides spike-in	Yes	Tian et al. (2023)
			CPTAC	NCI-7	Proteins from a panel of seven cancer cell lines (NCI-H23, RPMI-8226, T47D, A549, COLO205, NCI-H226, CCRF-CEM)	No	Clark et al. (2018)
			Thermo Fisher	HeLa	Proteins digested from HeLa cell lines	Yes	Kocher et al. (2011)
		Proteins from plasma	NIST	RM8231	Frozen human plasma suite consists of diabetic, high triglyceride, young African-American and Normal plasma samples	Yes	Lippa et al. (2022)
	Synthetic protein	Purified human proteins	Sigma-Aldrich/ABRF	UPS1/2	48 human purified recombinant proteins	Yes	Andrews et al. (2006)
			Invitrogen/HUPO	–	20 human purified recombinant proteins	Yes	Bell et al. (2009)
			NIST/CPTAC	NCI-20	20 human purified recombinant proteins	Yes	Tabb et al. (2010) and Wang et al. (2014)
			Biognosys	iRT	13 peptides	Yes	Escher et al. (2012)
			ABRF/IPT	SpikeMix	1000 peptides	Yes	JPT Peptide Technologies (2023)

Table 1 (continued)

Omics Category	Targeted molecule	Project/institute	Name	Sample source	Public availability of materials	References	
Metabolite	Biological metabolite	Metabolites from tissue	NIST	RM 8462	Frozen human liver suite consists of normal, fatty and congested liver samples	Yes	Lippa et al. (2022)
		Metabolites from cell line	Quartet	Quartet Metabolite	Metabolites from B-lymphoblastoid cell lines of a Chinese Quartet family	Yes	Zhang et al. (2022)
		Metabolites from plasma	NIST	SRM 1950	Human plasma of 100 healthy individuals (50 males + 50 females)	Yes	Phinney et al. (2013)
				RM 8231	Human plasma of diabetic, hypertriglyceridemic, African-American and normal (SRM 1950) samples	Yes	Aristizabal-Henao et al. (2020)
				SRM 1949	Human prenatal serum suite consists of non-pregnant and three pregnant trimesters samples	Yes	Boggs et al. (2021) and Sempos et al. (2022)
Synthetic metabolite	Metabolites from urine			RM 8232	Human urine of female non-smokers, female smokers, male non-smokers and male smokers	Yes	Lippa et al. (2022)
	Synthetic metabolite	Synthetic metabolites	CIL	QRcSS	18 isotope-labeled metabolite mixes	Yes	Cambridge Isotope Laboratories, Inc. (2023)
				IROA-LTRS	Yeast extract isotope-labeled complex with universally labeled at both 5% and 95% U-13C and mixed 1:1	Yes	Evans et al. (2020)
				Lipidyzer	Unlabeled lipid mole cular species	Yes	Lippa et al. (2022)
				MxP/AbsoluteIDQ	Seven levels of calibration standards and isotope-labeled internal standards, and plasma-based quality control samples at three concentrations	Yes	Biocrates (2023)

This list provides examples of commonly used representative omics reference materials, rather than an exhaustive compilation of all existing omics reference materials

CPTAC, Clinical Proteomics Technology Assessment for Cancer; ERCC¹, External RNA Controls Consortium; ERCC², Extracellular RNA Controls Consortium; FCR, Friends of Cancer Research; FNIH, Foundation for the National Institutes of Health; GIAB, Genome in a Bottle; GeT-RM, Genetic Testing Reference Material Coordination Program; GIMR, Garvan Institute of Medical Research; IFCC, International Federation for Clinical Chemistry; MAQC/SEQC, MicroArray/Sequencing Quality Control; MDIC, Medical Device Innovation Consortium; NIST, National Institute of Standards and Technology; SPOT/Dx, Sustainable Predictive Oncology Therapeutics and Diagnostics

Biological DNA Reference Materials

Germline variants are inherited genetic changes that occur in either a sperm or an egg cell and are passed on to offspring at the time of conception. These variants are present in all cells of the body and are typically detected from a blood sample. The accurate and reliable detection of germline variants is crucial for identifying genetic causes of disease and developing personalized treatment strategies.

The GIAB consortium, hosted by NIST, is dedicated to creating reference materials, methods, and datasets that facilitate the clinical translation and regulation of human genome sequencing (National Institute of Standards and Technology 2023b). In 2015, NIST released the primary human genome DNA reference material, RM 8398, derived from HG001/NA12878, a healthy female of European ancestry. To improve the representation of human genetic diversity, NIST further developed DNA reference materials from different ethnic populations, including an Ashkenazi Jewish family trio (RM8392) and a Han Chinese son (RM8393) (Zook et al. 2016). These genomes were chosen from the Personal Genome Project because of the broad consent for public genome data sharing and commercial use of products based on these cell lines. This broad consent has enabled commercial reference materials to be based on the same cell lines characterized by the GIAB, including spike-in DNA mimicking challenging variants, somatic variants, and circulating tumor DNA (ctDNA), which are explained below. The seven genomes have been extensively characterized by the GIAB consortium for benchmarking germline variants, including single nucleotide variations (SNVs), small indels and structural variants (SVs) (Wagner et al. 2022; Zook et al. 2014, 2019, 2020). While all of GIAB's current benchmarks are focused on germline "normal" cell lines, the consortium is currently collaborating to develop new broadly consented tumor-normal cell line pairs for genomic RM development.

In recent years, several initiatives have made significant strides in developing biological DNA RMs to serve as benchmarks for whole-genome germline variants (Li et al. 2018). The Quartet Project, led by Fudan University in close collaboration with the National Institute of Metrology of China and other organizations, established four immortalized lymphoblastoid cell lines from a Chinese Quartet family, including a father, mother and two monozygotic daughters (Ren et al. 2023; Zhang et al. 2023). This family was recruited from the Fudan Taizhou cohort in Central China, thus possessing genetic features from both Northern and Southern Chinese populations (Wang et al. 2009). The four DNA RMs have been certified by China's State Administration for Market Regulation as the First Class of National Reference Materials. They have been extensively used for proficiency testing and methods validation in clinical and commercial laboratories and the sequencing datasets are

publicly available (Khayat et al. 2021; Pan et al. 2022). In addition to DNA RMs, the Quartet Project released corresponding RNA, protein and metabolite RMs derived from the same cell lines. Therefore, the Quartet RMs have three types of "truth" to assess the performance of variants calling results (Ren et al. 2023). The first is the characterized benchmark variants, which can be used to evaluate the performance of variants identified inside the benchmark regions. The second is the Mendelian inheritance law underlying the monozygotic twins and their parents. The third kind of "truth" is central dogma of multiomics RMs, which enables cross-omics validation of variant calls from multiomics datasets.

High-throughput sequencing technologies aim to scan variants on a whole-genome scale, while clinical genetic testing focuses on a few particular genetic variants associated with diseases. To support quality control for clinical genetic testing, the U.S. Centers for Disease Control and Prevention (CDC) led the Genetic Testing Reference Materials Coordination Program (GeT-RM) to develop DNA RMs, including those for rare inherited genetic diseases, human leukocyte antigen (HLA) testing and pharmacogenetics (Coriell Institute 2023). The GeT-RM obtained cell lines containing medically important mutations from the Coriell Cell Repositories, and then distributed genomic DNAs to multiple volunteering laboratories for genotyping and mutation confirmation using a variety of platforms and assays (Centers for Disease Control and Prevention 2019). The GeT-RM has characterized DNA RMs for a wide range of genetic disorders, such as cystic fibrosis (Pratt et al. 2009), Duchenne and Becker muscular dystrophy (Kalman et al. 2011), fragile X syndrome (Amos Wilson et al. 2008), Huntington disease (Kalman et al. 2007), and many others, including 11 human leukocyte antigen loci (Bettinotti et al. 2018) and pharmacogenetic loci (Gaedigk et al. 2019; Pratt et al. 2016). These reference materials represent specific mutations associated with diseases and are available for research, clinical test development, quality assurance and control, and proficiency testing to ensure the accuracy of clinical testing.

Somatic variants are genetic mutations that occur in non-germline cells. They are typically detected in tumors from sequencing datasets of paired tumor and normal samples, with normal samples used to remove germline variants. Accurate and reliable detection of somatic variants is crucial for gaining insights into cancer biology, guiding targeted therapies and improving patient outcomes in cancer treatment. DNA RMs used to benchmark somatic variants usually consist of matched tumor and normal genomes.

The MicroArray and Sequencing Quality Control (MAQC-IV/SEQC2) consortium recently completed its fourth project, which aimed to develop standard analysis protocols and quality control metrics for the use of

high-throughput DNA sequencing data in regulatory science research and precision medicine (MAQC Consortium 2021). The Somatic Mutation Working Group (WG1) of SEQC2 established paired tumor-normal DNA RMs and corresponding whole-genome reference datasets for small variants and structural variants (Fang et al. 2021; Talsania et al. 2022). The DNA RMs are gDNA derived from a triple-negative breast cancer (TNBC) cell line (HCC1395) and a B-lymphocyte-derived normal cell line (HCC1395BL) from the same donor, obtained from the American Type Culture Collection (ATCC). HCC1395 is a well-studied cell line with abundant somatic alterations, including approximately 40,000 SNVs, around 2000 indels, copy number alterations (CNAs) affecting 56% of the genome, 256 complex genomic rearrangements, and 138 experimentally confirmed fusion genes (Stephens et al. 2009). The SEQC2 WG1 later used DNA RMs to address challenges in accurately detecting somatic variants from WGS and WES by examining experimental and bioinformatic components affecting their reproducibility and accuracy, covering a wide range of topics, including library preparation protocols, DNA input amount, tumor purity, read coverage, and bioinformatic pipelines (Sahraeian et al. 2022; Talsania et al. 2022; Xiao et al. 2021).

Due to the heterogeneity of tumors and diverse mutational profiles among different types of cancer, a single reference material, such as HCC1395, may not fully represent breast or any other types of cancer genomes. Nevertheless, it is highly mutated and suitable for benchmarking, developing, and refining protocols and tools for somatic variant detection. To better capture the genetic diversity of tumor genomes, researchers have made efforts to establish DNA RMs from more tumor types. For instance, Craig et al. (2016) created DNA RMs from a metastatic melanoma (COLO829) and its paired B-lymphoblastoid normal cell line (COLO829BL).

While WGS and WES provide a more comprehensive view of the entire genome, targeted sequencing, also known as oncopanel sequencing, offers a more cost-effective and efficient approach by focusing on a limited number of cancer hotspot variants. It can detect variants with a variant allele frequency (VAF) as low as 0.5%. The Oncopanel Sequencing Working Group (WG2) of SEQC2 established two DNA RMs for oncopanel benchmarking (Jones et al. 2021). Sample A is an equal mass pooled gDNA sample of the same 10 cancer cell lines that were originally used for developing the Agilent Universal Human Reference RNA material (UHRR, Catalog #74000) (MAQC Consortium 2006, 2014), covering as many clinically related variants as possible to increase variant density in coding regions. Sample B is derived from a non-cancer male cell line (Agilent OneSeq Human Reference DNA, PN 5190-8848). To emulate the range of VAFs typically encountered in targeted sequencing and ctDNA sequencing, tumor Sample A was diluted by normal Sample B at different ratios to create a series

of tumor DNA reference materials with even lower VAFs of variants. The SEQC2 WG2 employed these DNA RMs to conduct cross-platform multi-laboratory evaluations of commercially available oncopanels, and developed actionable guidelines to improve the performance and consistency of oncopanel sequencing across different laboratories and platforms (Deveson et al. 2021a; Gong et al. 2021).

Apart from benchmarking individual somatic variant calls, DNA RMs have been developed to benchmark aggregated genomic biomarkers derived from somatic variants, such as tumor mutation burden (TMB). TMB is a promising biomarker for predicting response to pan-cancer immune checkpoint inhibitor therapy (Samstein et al. 2019; Yarchoan et al. 2017). The gold standard for measuring TMB is to perform tumor-normal paired WES and count the total number of non-synonymous mutations in the coding regions. However, WES is a relatively costly and time-consuming approach. To address this, researchers are exploring the use of less expensive targeted sequencing panels that focus on a small number of driver genes to estimate TMB. However, significant variability in TMB measurement has been observed (Buttner et al. 2019). To address the need to standardize and harmonize TMB assessment across assays and laboratories, many initiatives have developed DNA RMs, such as Friends of Cancer Research TMB Harmonization Project (Merino et al. 2020; Stenzinger et al. 2019; Vega et al. 2021) and SeraSeq (Seracare 2023a). These DNA RMs are established from Formalin-Fixed Paraffin-Embedded (FFPE) clinical samples or tumor cell lines. Contrived RMs are developed by mixing gDNA of tumor cell lines with matched normal cell lines at a series of proportions to mimic low VAF variants detected from liquid biopsy (Zhang et al. 2021).

Engineered DNA Reference Materials

Engineered DNA RMs are designed to assess the analytical performance of laboratory developed tests (LDTs) for oncology therapies by introducing desired cancer hotspot mutations into germline DNA RMs using gene editing systems such as CRISPR/Cas9 (Jia et al. 2018; Pfeifer et al. 2022; Suzuki et al. 2020; Zehnbaauer et al. 2017). Cell lines like HapMap cell lines (e.g., HG001 and HG002) and the Quartet cell lines are preferred for this purpose, because they can be easily cultured in large quantities and widely distributed due to their broad consent (Lin et al. 2022). Each variant is independently engineered into different cell lines, and the multiplexed DNA RMs are created by mixing genomes with engineered variants together (Medical Device Innovation Consortium 2019). The risk of unexpected off-target effects induced by genome editing is a major concern, which can result in variants at similar DNA sequences other than the intended on-target sites (Zhang et al. 2015). To ensure that the

original and engineered cell lines are isogenic at all locations except for the engineered variant sites, various computational and experimental methods are used to detect any off-target CRISPR/Cas9 activity, including Sanger sequencing, Food and Drug Administration (FDA)-approved clinically validated targeted gene panels, PacBio WGS, and circularization for high-throughput analysis of nuclease genome-wide effects by sequencing (CHANGE-seq) technology (Lazzarotto et al. 2020). Engineered DNA RMs can also be developed with abundant somatic variants by knocking down genes in the mismatch repair (MMR) pathway and proofreading systems, which are crucial for high fidelity of genome replication (Wang et al. 2023). Clones are then selected by flow cytometry and cultured to accumulate sufficient somatic variations. One major concern of engineered DNA RMs is that the engineered genomes are not able to mimic the complexity and heterogeneity of real cancer genomes.

Engineered DNA RMs can be used to benchmarking ctDNA assays (Deveson et al. 2021b; Horizon Discovery 2023; Seracare 2023b; Thermo Scientific 2020). To emulate low concentrations of ctDNA in plasma, the mutated genome is diluted by a background genome with wild-type alleles, resulting in somatic variants with low VAF. The DNA sequences are then fragmented to an average size of 150–170 bp to closely resemble highly degraded ctDNA extracted from human plasma. While researchers have attempted to digest DNA sequences using micrococcal nuclease (MNase) to preserve nucleosome core particles and trimmed nucleosomes (Zhang et al. 2017), further investigation is needed to determine if the contrived DNA RMs can fully represent the biological properties of ctDNA and perform equivalently, or even sufficiently, to clinical specimens.

Synthetic DNA Reference Materials

Synthetic DNA molecules are artificially created through chemical synthesis techniques and do not necessarily align with the human reference genome. The synthetic DNA RMs are developed to represent diagnostic features and to address the specific requirements of any next-generation sequencing (NGS) test, especially clinically relevant or difficult variants. Synthetic DNA RMs are often used as spike-in controls to be added into a study sample with a known quantity to measure sensitivity and precision of NGS libraries (Deveson et al. 2016, 2019). These spike-ins should be added at sufficient abundance to achieve matched sequencing coverage with the accompanying sample without sacrificing too many sequencing reads (Blackburn et al. 2019).

Synthetic DNA RMs enable researchers to evaluate the quantitative properties of DNA sequencing, such as VAF, limit of detect (LOD), and copy numbers. A pair of synthetic DNA RMs, which represent wild-type and mutant alleles, can be combined to simulate lower somatic VAF

and to establish LOD (Blackburn et al. 2019; Deveson et al. 2021b). Alternatively, sequence elements can be encoded in a single synthetic DNA molecule at known abundances (Reis et al. 2020).

RNA Reference Materials

RNA sequencing (RNA-seq) is typically used for identifying differentially expressed transcripts or genes between experimental groups and control groups; thus, RNA RMs are often provided as sample pairs or groups. As high-throughput sequencing technologies do not directly measure absolute abundances of RNA molecules reliably, differences or relative expression levels (ratios) between sample groups serve as a built-in truth to benchmark quantitative measurements, instead of absolute abundance for each RNA in a single sample. Biological RNA RMs are derived from single large batches of immortalized cell lines. Gene and transcript expression levels are characterized by multiple methods (MAQC Consortium 2006; Yu et al. 2023). Synthetic RNA RMs are exogenous or artificial RNA oligonucleotides with a wide range of known concentrations, which enable for the assessment of dynamic range, specificity, sensitivity and LOD that is not otherwise possible for biological RNA reference materials (Jiang et al. 2011; Munro et al. 2014). However, synthetic RNA RMs lack the desired complexity and diversity and, thus, could behave differently from biological samples.

Biological RNA Reference Materials

The MAQC/SEQC projects initiated by the US FDA utilized two human RNA RMs to comprehensively evaluate the comparability and accuracy of gene-expression measurements obtained through microarray and RNA-seq techniques across different laboratories and protocols (MAQC Consortium 2006, 2014). The two RNA RMs are the Agilent Universal Human Reference RNA composed of total RNA from 10 human tumor cell lines (termed “Sample A”) and the ThermoFisher Human Brain Reference RNA (HBRR, termed “Sample B”). A pair of RNA RMs can be combined together at known mixing ratios, thus enabling users to assess relative accuracy of each method based on the differentially expressed genes detected. Samples A and B were then mixed in 3:1 and 1:3 ratios, respectively, to generate Samples C and D. This combination of biologically different RNA sources and known titration differences enables assessment of relative accuracy based on the differentially expressed genes detected.

In addition to comparing the consistency of differentially expressed genes between query datasets and reference datasets, unsupervised clustering methods such as principal component analysis (PCA) were used to assess the performance of omics data to distinguish sample groups. Gene

expression profiles between MAQC samples A and B are significantly different with more than 15,000 differentially expressed genes. Successfully distinguishing sample groups with such large differences does not guarantee the ability to identify subgroups with subtle biological differences of clinical samples. The four Quartet RNA reference materials are much more similar to each other and require increased performance of a method for distinguishing them (Yu et al. 2023). There are about 2000 differentially expressed genes between any two of the Quartet samples, which are enriched in B cell mediated immunity. Triplicates of each Quartet RNA reference material were sequenced in a batch to benchmark transcriptomic profiling cross protocols and laboratories. The signal-to-noise ratio (SNR) metric, defined as the ratio of the inter-sample distance on the PCA plot over the intra-sample distance between technical replicates, is applied to assess the quality of query datasets. Higher SNR values indicate stronger power to discriminate sample groups, while lower SNR values indicate technical biases or sequencing failures of one or more replicates. In the Quartet benchmark studies, some experiments were found to have high replicate consistency but relatively low SNR values, indicating that these experiments may had systematic technical biases which can only be revealed by multi-sample reference materials.

Synthetic RNA Reference Materials

Synthetic RNA RMs are used as external spike-in controls to be added to the samples of interest for transcriptomic analysis. Laboratories have used different custom external spike-ins for specific platforms and assays, before the External RNA Controls Consortium developed the first generally accepted RNA spike-ins for various microarray and sequencing applications, which were later distributed by NIST as SRM 2374 (Baker et al. 2005). The ERCC spike-in controls comprise 92 polyadenylated RNA transcripts derived from bacterial sequences or in vitro transcription of synthetic DNA sequences. The 92 ERCC RNA control transcripts are categorized into four sub-pools, each containing 23 transcripts spanning a wide dynamic range in concentration. The sub-pools are combined to create two mixtures, Mix1 and Mix2, in four defined abundance ratios of 4:1, 1:2, 1:1.5, and 1:1 (Munro et al. 2014). Spike-in controls can be added to biological RNA RMs to combine the advantages of both types of reference materials into a single sequencing run. In the third phase of the MAQC project or SEQC, researchers conducted a broader analysis of RNA-seq performance evaluating the sensitivity and technical variation between different NGS methods and laboratories by complementing samples A and B with ERCC controls (MAQC Consortium 2014).

While ERCC spike-in transcripts are valuable for standardization and quantification in transcriptomic analysis, they have limitations in representing the diversity and complexity of endogenously expressed transcripts due to their lack of isoforms. To address this limitation, alternative RNA reference materials have been developed. Sequins, designed by the Garvan Institute of Medical Research, provide a comprehensive representation of alternative isoforms and the complex exon–intron architecture of human genes, allowing for more accurate assessment of gene fusion, alternative splicing, and transcript assembly (Hardwick et al. 2016). Additionally, the spike-in RNA variants (SIRV) developed by Lexogen consist of 69 synthetic transcript isoforms that comprehensively reflect variations of alternative splicing, alternative transcription start- and end-sites, overlapping genes, and antisense transcripts (Lexogen 2023). This set of RNA reference materials enables the evaluation of the performance of isoform-specific RNA-seq workflows, and thus provides a more comprehensive evaluation of RNA-seq performance.

RNA-seq can be used to sequence long RNAs, such as messenger RNAs, as well as short RNAs, such as microRNAs (miRNAs), that differ in length. The Extracellular RNA Communication Consortium led a benchmark study for miRNA quantification across multiple protocols and laboratories using small RNA-seq (Giraldez et al. 2018). They used diverse combinations of synthetic RNAs to evaluate sequence-specific biases and accuracy. An equimolar pool consisted of over 1000 chemically synthesized RNA oligonucleotides (15–90 nt) mixed at equal concentration was used to assess reproducibility of absolute RNA sequences abundance at counts per million (CPM) level. Two synthetic small RNA pools with RNAs varied in defined relative amount were used to assess the concordance for relative quantification. Synthetic pools with unedited and edited miRNA variants in different ratios were used to determine the accuracy of quantifying miRNA editing.

Protein Reference Materials

The proteome refers to the entire set of proteins expressed by a cell, tissue or organism at a particular time. Proteomics is the systematic, high-throughput study of the composition, functions, and interactions of all proteins. In proteomics research, where the sheer multitude of proteins presents a formidable challenge, mass spectrometry (MS) is commonly used for both qualitative and quantitative protein analysis. This process involves comparing detected peptide maps with protein sequences sourced from databases. However, the complexity of MS-based proteomics experiments and their potential for considerable variability can hinder the achievement of accurate and reproducible

results. To enhance the reliability and reproducibility of proteomics, numerous initiatives have been actively working for decades to establish community standards and guidelines. These efforts aim to ensure consistency, promote rigorous experimental practices, and facilitate the generation of reliable and comparable proteomic data across different laboratories and studies. For example, the Proteomics Standards Initiative (PSI) of the Human Proteome Organization (HUPO) standardized practices and guidelines for data reporting formats (Deutsch et al. 2017), data quality control framework (Bittremieux et al. 2017) and data interpretation (Omenn 2021). CPTAC, launched by the US National Cancer Institute (NCI), intends to improve MS-based proteomics measurement quality for biomarker discovery in cancer research (Tabb et al. 2016; Zhou et al. 2017). The Proteomics Standards Research Group (sPRG) of the Association of Biomolecular Resource Facilities (ABRF) develops and implements standards to reflect the accuracy and consistency of proteomics (Tabb et al. 2010).

The limitations of traditional methods have driven the development of new technologies tailored for highly sensitive protein biomarker discovery while demanding minimal quantities of biological materials (Eldjarn et al. 2023; Sun et al. 2023). Examples of this innovation include Olink's Proximity Extension Assay (PEA) (Petretera et al. 2021; Wik et al. 2021), SomaLogic's SomaScan Assay (Candia et al. 2017, 2022), and Seer's Proteograph (Blume et al. 2020). Among them, the Olink technology implements more stringent quality control procedures by integrating four internal controls into all samples and including external controls within each plate (Olink 2023). The Olink technology employs a detection principle where two specific proximity probes are coupled, generating an amplicon upon binding to the target protein, which is then quantified using quantitative real-time PCR (qPCR) or NGS (Wik et al. 2021). The internal controls encompass an incubation control, an extension control, and a detection control, each serving distinct roles in data quality assurance during the PEA. The incubation controls involve non-human antigens with matching antibodies for monitoring throughput the PEA process. The extension control consists of IgG antibodies coupled with matching oligo pairs, ensuring the constant proximity of DNA tags and supporting data normalization while monitoring the extension, amplification, and detection steps. The detection control, a synthetic double-stranded DNA, contributes to data quality control and aids in identifying potential issues in the final amplification and detection stages. Furthermore, each sample plate incorporates eight external controls, which consist of two pooled sample controls to estimate inter and intra consistency for each assay, three negative controls containing buffer to establish background levels and calculate the limit of detection, and three plate controls designed to

account for potential variations between runs and plates, further enhancing the reliability and accuracy of the results.

In the realm of proteomics studies, the inclusion of various types of RMs, such as internal standards, blank QC samples or negative controls, and pooled QC samples, equips researchers with the means to effectively manage variability, achieve precise quantification, and ensure the reliability of their findings (Bittremieux et al. 2018; Bunk 2010; Chiva et al. 2021). Within this review, our primary focus is on external RMs that can be analyzed alongside the study samples, enabling comparisons between runs, batches and studies, as well as ensuring the accuracy of detection through comparison with reference datasets or the categorization of different sample groups.

A diverse range of protein RMs has been developed to support both qualitative and quantitative proteomic measurements, serving as invaluable tools for method validation, quality control, and calibration within the field of proteomics research. These RMs serve as valuable tools for method validation, quality control, and calibration in proteomics research. These RMs vary in complexity, ranging from simple synthetic peptide mixtures to digests of bovine serum albumin and even complex human tissues (Bittremieux et al. 2018). Biological matrix protein RMs are derived from biological materials such as whole-cell lysate and biofluids, offering a higher level of sample complexity with tens of thousands of proteins. While biological protein RMs obtained from human tissues and plasma better resemble the biological features of their respective sample types, their availability is usually limited, and the reference datasets can differ between lots. In contrast, microorganisms and cell lines provide inexpensive and renewable protein RMs, ensuring sufficient quantities for the research community. Synthetic protein RMs, on the other hand, consist of mixtures of a limited number of purified recombinant human proteins or chemically synthesized peptides with defined molecular weights and concentrations. However, the composition of synthetic proteins is much simpler and does not resemble the complex biological characteristics of clinical samples.

Biological Protein Reference Materials

Protein identification is a critical component of proteomics and often leads to subsequent investigations in which these proteins are quantified. One-sample based RMs are often used to evaluate the performance of protein identification and absolute quantification, for example RM8461 released by NIST from a cryogenically homogenized and freeze-dried liver tissue (Davis et al. 2019), yeast *Saccharomyces cerevisiae* (later released by NIST as RM8323) (Beasley-Green et al. 2012; Paulovich et al. 2010), bacterium *Shewanella oneidensis* (Nakayasu et al. 2021), NCI-7

Cell Line Panel (Clark et al. 2018), HeLa cells (Kocher et al. 2011) and HEK293T cells (Collins et al. 2017). While these proteomes have been extensively characterized and served as a model proteome in numerous fundamental proteomic investigations, reference datasets of high-confidence proteins have not yet been established. As a result, the performance of protein identification and absolute quantification are currently being assessed primarily based on the consistency between technical replicates of the same RM.

Another important application of quantitative proteomic measurements is to determine differentially expressed proteins (Anwaier et al. 2022; Ku et al. 2023). Multi-sample RM suites are developed to facilitate relative quantitation assessment. CPTAC established a pair of patient-derived xenograft tumors as a comparative reference material (CompRef) to longitudinally monitor the reproducibility of differential proteomics across instruments and centers for the Cancer Genome Atlas (TCGA) (Tabb et al. 2016; Zhou et al. 2017). The CompRef represent basal (WHIM2) and luminal-B (WHIM16) breast cancer subtypes, having significantly different proteomic signatures. They were then utilized as standards to establish uniform analytical pipelines for other cancer types (colorectal and ovarian cancers) by the CPTAC Common Data Analysis Platform (CDAP) (Rudnick et al. 2016). NIST is currently developing protein RMs suites for the assessment of relative quantitation in proteomics, including RM 8462 Frozen Human Liver Suite (normal, fatty and congested liver samples) and RM 8231 Frozen Human Plasma Suite (Diabetic, high triglyceride, young African-American and Normal plasma samples). The Quartet Project also released biological protein RMs, which are extracted from four lymphoblastoid cell lines as the same with DNA and RNA reference materials mentioned above (Tian et al. 2023).

Synthetic Protein Reference Materials

Synthetic protein reference materials have been extensively used in benchmark studies of proteomic measurements to determine experimental and analytical variations by big consortia (Paulovich et al. 2010; Tabb et al. 2010). Notable examples of standard protein mixtures include: the Universal Proteomics Standards (UPS1 and UPS2), a mixture of 48 human recombinant proteins jointly developed by ABRF's sPRG and Sigma-Aldrich (Andrews et al. 2006); the HUPO Gold MS Protein Standard, a mixture of 20 human proteins, developed by the joint efforts of HUPO and Invitrogen (Bell et al. 2009); a mixture of 20 purified human proteins (NCI-20), produced by NIST and employed by CPTAC for intra- and inter-laboratory studies aiming at evaluating repeatability and comparability of qualitative proteomics (Tabb et al. 2010; Wang et al. 2014).

Chemical synthetic or modified peptide mixtures are also utilized as RMs. In comparison to protein mixtures, peptide mixtures have a simpler composition. However, it is important to note that they cannot fully capture the variability introduced during enzymatic digestion, as different laboratories may employ diverse proteolytic enzymes, chemicals, and conditions for digestion. Several synthetic peptide reference materials are commercially available, such as a mixture of 1000 heavy-label proteotypic peptides for conserved proteins across three species (human, mouse and rat), established by ABRF and JPT Peptide Technologies (2023). Synthetic peptides are especially important to evaluate the performance of targeted quantitative proteomic measurement, such as multiple reaction monitoring (MRM) and parallel reaction monitoring (PRM). They are often used to predict retention times (RTs) for large-scale scheduled liquid chromatography multiple reaction monitoring (LC-MRM) measurements with a single calibration run before the analytical runs. Biognosys has developed a mixture of 11 artificial synthetic peptides (iRT) to determine peptide retention time (RT) values and calibrate chromatographic systems for increasing the throughput (Escher et al. 2012). Additionally, well-defined synthetic protein reference materials can be added into biological protein reference materials or test samples to provide additional information of qualitative accuracy. An important consideration when spiking synthetic peptides into other samples is that these peptides should not overlap with the original sample content.

Metabolite Reference Materials

Metabolomics encompasses the extensive investigation of small molecules, known as metabolites, within cells, biological fluids, tissues, or organisms. It integrates the influences of factors from genomics, transcriptomics, proteomics, as well as environmental elements like diet and lifestyle. Since metabolites serve as indicators of the downstream effects of these factors on cellular functions, they closely represent the actual phenotypes of cells, tissues, or organisms, offering novel insights into metabolism and its regulation in physiological and pathological processes, including health, aging, and diseases. Metabolomics involves the simultaneous identification and quantification of various small molecule types, including amino acids, fatty acids, carbohydrates, and other products of cellular metabolic functions. In comparison to genomics, transcriptomics, and proteomics, the reliable identification and quantification of the metabolome are significantly more complex due to the chemical complexity and the presence of isomers—compounds with the same molecular formula but different structural arrangements—introducing challenges for precise identification and quantification.

To promote the advancement of metabolomics toward higher quality, several large research consortia have emerged in the field, aiming to enhance the reproducibility of metabolomics research results through comprehensive quality assurance and quality control measures. These consortia have undertaken various efforts, including the establishment of best practices, promotion of communication and education, and the advancement of the field toward higher-quality standards. The mQACC, consisting of experts in quality assurance and quality control, is focused on developing universal best practices and reporting standards to ensure the robustness and reproducibility of untargeted metabolomics research (Beger et al. 2019; Evans et al. 2020). The Metabolomics Society Data Quality Task Group (DQTG) aims to enhance the robustness of quality assurance and quality control in the metabolomics community through communication, advocacy, education, and the promotion of best practices (Kirwan et al. 2022). The Standard Metabolic Reporting Structures (SMRS) group is dedicated to standardizing metabolomics analysis and provides comprehensive reports and summaries on relevant key issues (Beckonert et al. 2007; Lindon et al. 2005). The ABRF Metabolomics Research Group aims to study the reproducibility of metabolomics research and propose best data analysis strategies by comparing analysis groups using the same dataset (Turck et al. 2020). Additionally, the ABRF plays a role in improving the core competencies of biotechnology laboratories through research, communication, and education (Cheema et al. 2015; Turck et al. 2020). The Metabolomics Consortium has proposed guidelines for achieving high-quality reporting of LC–MS-derived metabolomics data, including the identification and prioritization of test materials, assessment of useful indicators of data quality, and descriptions of common practices and variations in quality assurance and quality control workflows (Broadhurst et al. 2018).

Quality control samples can be categorized into three primary types based on their intended purposes. System suitability test samples serve as a quality assurance measure applied before data acquisition to instill confidence in the eventual high-quality results (Broadhurst et al. 2018; Kirwan et al. 2022). Typically, these samples consist of solutions containing a small number of authentic chemical standards, typically ranging from five to 10 analytes, with known concentrations. They play a critical role in instrument calibration and assessment of critical system parameters, including mass-to-charge (m/z) ratio and chromatographic characteristics such as retention time, peak area, and peak shape.

Blank quality control samples and matrix-matched quality control samples are essential components of quality control measures to ensure that the quality management process is fulfilled. Blank quality control samples consist of samples devoid of metabolites, serving to identify potential sample contamination or instrument-related background signals,

thereby eliminating interference from external contaminants or instrument-related background signals, thereby eliminating interference from external contaminants or the instrument itself (Kirwan et al. 2022). By comparing data from the actual samples to that from the blank samples, researchers can distinguish genuine metabolite signals from potential interferences or background noise. Within the category of matrix-matched quality control samples, the most commonly used are pooled samples. These samples are created by pooling a small amount of each analyzed biological sample within a study, representing both the sample matrix and metabolite composition. Pooled QC samples play a multifaceted role, conditioning the analytical platform, conducting intra-study reproducibility measurements, and mathematically correcting for systematic changes in parameter values (Broadhurst et al. 2018). A specific type of pooled QC sample can be used to assess data quality across different studies within the same laboratory, termed long-term reference (LTR) QC samples (Broadhurst et al. 2018). These samples are obtained either through the commercial purchase of the required sample types or by collecting representative samples from various studies within the laboratory. In this review, we focus on the use of external RMs for assessing performance across different laboratories, which are created and sold by a certified group.

Biological Metabolite Reference Materials

SRM 1950 released by NIST is one of the first developed metabolite reference materials, which is intended for quality control of identifying and quantifying metabolites in human plasma, such as fatty acids, electrolytes, vitamins, hormones, and amino acids (Phinney et al. 2013). It is a mixture of human plasma samples from 100 individuals reflecting a racial distribution in the US population at the time of implementation (77% white, 12% African-American or black, 2% American Indian or Askan Native, 4% Asian, 5% other, with about 15% Hispanic origin). A total of 90 metabolites are assigned with high confidence values of absolute concentrations by integrating several different analytical methods. SRM 1950 was initially designed for targeted metabolomics, and has been extensively used to benchmark platforms, protocols and workflows (McGaw et al. 2010; Misra and Olivier 2020; Siskos et al. 2017; Thompson et al. 2019). Recently, it has also been used in benchmark studies of untargeted metabolomics and lipidomics (Azab et al. 2019; Bowden et al. 2017; Cajka et al. 2017). NIST also released other standalone natural-matrix reference materials for organic contaminants from an assortment of biological materials, including frozen non-fortified human milk (SRM 1953), fortified human milk (SRM 1954), non-fortified human serum (SRM 1957), fortified human serum (SRM 1958) (Schantz et al. 2013), lyophilized human serum (SRM 909b and SRM

909c) (Aristizabal-Henao et al. 2021), smokers' human urine (SRM 3672), and non-smokers' urine (SRM 3673).

Like other quantitative omics, such as transcriptomics and proteomics, identifying differentially expressed metabolites between sample groups is one of the main purposes for metabolomics-based biomarker researches. RMs consisting of two or more sample groups can be used to assess the performance of distinguishing sample groups. The NIST Metabolomics Quality Assurance and Quality Control Materials (MetQual) Program released a suite of pooled plasma materials (RM 8231) comprising four different metabolic health states, including type 2 diabetes plasma, hypertriglyceridemia plasma, normal African-American plasma and normal human plasma (SRM 1950) (Met Qual Program Coordinators 2023). The MetQual Program is planning to conduct an inter-laboratory study to obtain consensus characterization of RM 8231 and assess measurement variability within the metabolomics community. NIST also developed several multi-sample metabolite reference materials from other biological resources. RM 971a consists of two serum mixtures: one from a pool of healthy, premenopausal adult females, and the other one from a pool of healthy adult males. It is intended to evaluate the accuracy of identify and quantify hormones in human serum (Aristizabal-Henao et al. 2021). SRM 1949 Frozen Human Prenatal Serum is a four-level material that was pooled from non-pregnant women and women during each trimester of pregnancy, aiming at quality control for the measurement of hormones and nutritional elements throughout pregnancy (Boggs et al. 2021; Sempes et al. 2022). A suite of human urine reference materials (RM 8232) is under development. The suite will consist of four pooled urine samples from female non-smokers, female smokers, male non-smokers and male smokers. Relative metabolite fold changes, percent differences for the top 20 metabolites and the identified top 30 abundant metabolites of the urine samples will be characterized by both LC-MS and nuclear magnetic resonance. RM 8462 Frozen Human Liver Suite mentioned in the protein reference materials section can be also used for metabolomics (Lippa et al. 2022).

The Quartet Project also developed a multi-sample metabolite RM suite by extracting metabolites from the four immortalized lymphoblastoid cell lines. Aiming at assessing the performance of detecting biological differences between different sample groups, reference datasets for fold changes of absolute abundance values between samples groups were constructed, by consensus across platforms, laboratories and replicates. The performance of quantitative metabolomics can be assessed not only by the consistency between fold changes of differentially expressed metabolites in query datasets and reference datasets, but also by SNR by measuring the ability to discriminate the intrinsic biological differences between the four sample groups.

Synthetic Metabolite Reference Materials

Synthetic metabolite reference materials are artificial substances that have identical chemical properties to naturally occurring metabolites in biological systems. They play an important role as calibration standards for analytical methods to allow accurate identification and quantification of metabolites. Synthetic metabolite RMs contain known concentrations of chemical components, which can be run separately or used as internal standards to perform system suitability tests, calibration, and metabolite quantification. These RMs can be prepared in individual laboratories to fit specific purposes for each study or can be purchased from vendors. They can be produced using chemical synthesis or enzymatic reactions, and they can be used for a range of applications, including targeted and untargeted metabolomics, and in the development and validation of new analytical methods. Synthetic metabolite RMs can also be used to assess the accuracy and precision of different analytical platforms and to facilitate inter-laboratory comparisons.

One example of a synthetic metabolite RM is the deuterated internal standards that are frequently used in MS-based metabolomics. These internal standards are made by incorporating deuterium into the metabolite of interest, allowing for accurate quantification of the metabolite in biological samples. Commercially available synthetic metabolite reference materials are typically mixtures of isotopically labeled or U-13C labeled metabolites that span a broad range of molecular weights, possess varied ionization propensities, and cover a distribution in class and retention time. Examples of commercially available synthetic metabolite reference materials include the QReSS kit from Cambridge Isotope Laboratories (CIL) (Cambridge Isotope Laboratories, Inc. 2023), the IROA-Long-Term Reference Standard (IROA-LTRS) from IROA Technologies (Evans et al. 2020), the Lipidzyer Platform kits from SCIEX (Lippa et al. 2022), and quantitative metabolic profiling kits from Biocrates (Biocrates 2023).

Multiomics Reference Materials

Multiomics integrates diverse omics data to better cluster and classify sample (sub)groups, and more comprehensively understand the mechanisms underlying biological processes by investigating molecular interaction across omics layers (Karczewski and Snyder 2018; Price et al. 2017; Schussler-Fiorenza Rose et al. 2019). Multiomics analysis inherits challenges from the single omics datasets and confronts new challenges in data harmonization and integration across different omics layers with varying numbers of features and statistical properties (Athieniti and Spyrou 2023; Sonia Tarazona 2021). Multiomics RMs derived from the same

source that incorporate multiple omics types and provide unbiased ground truth serve as crucial tools for assessing the performance of methods for normalizing and integrating multiomics datasets, conducting cross-omics validation, and imputing missing data (Krassowski et al. 2020; Zheng et al. 2023). They enable the validation and comparison of data integration methods across multiple omics layers, allowing for the identification of potential biases or discrepancies in the integration of data. Cross-omics validation is a critical step in multiomics research, involving comparing and validating findings across different omics layers. Multiomics RMs provide a standardized framework for conducting such validation, ensuring that results obtained from different omics techniques align and mutually reinforce each other.

The Quartet Project, aiming at quality control and data integration of multiomics profiling, has established a series of openly consented multiomics reference materials, including matched DNA, RNA, protein and metabolite, derived from the same batch of immortalized EBV infected B-lymphoblastoid cell lines from a healthy Chinese Quartet family with parents and monozygotic twin daughters (Fig. 1) (Zheng et al. 2023). Replicates of each Quartet reference material were analyzed in each batch for performance evaluation. Correctly classifying different Quartet samples based on multiomics features can be used for assessing the reliability of correlation-based multiomics network integration. The subtle known biological differences among the four reference samples may allow technical biases and batch effects to be discerned more efficiently when using multiple sample reference materials.

Recently, NIST has partnered with institutions around the world to form an open consortium, the International Microbiome and Multi-Omics Standards Alliance (IMMSA), to address multiomics measurement challenges for microbiome. The IMMSA has five working groups planning to develop microbial reference materials, benchmark bioinformatic tools, establish best practices for metabolomics, develop standards for documents and written, and develop standard methods for enumerating whole-cell reference materials. Human whole stool is one source of their candidate reference materials, because it can facilitate the understanding of biologically relevant properties of the human gut microbiome and identify new biomarkers that may serve as disease indicators.

Reference Datasets for Reference Materials

Establishing reference datasets for biological reference materials is crucial for evaluating the performance of high-throughput technologies. These reference datasets behave like "examination papers with right answers" to identify false-positive and false-negative results in a given test.

Reference datasets can be qualitative, which involves identifying the presence or absence of variants, transcripts, proteins, or metabolites in the sample. They can also be quantitative, which involves determining the concentration, abundance, or expression levels of these molecules. Since a single run of a sample can lead to errors or miss true analytes (substances), reference datasets need to be carefully established by integrating data from various measurement technologies and bioinformatic analysis pipelines to avoid biases toward a specific method or platform. This ensures that the reference datasets provide an unbiased ground truth for evaluating the performance of multiomics technologies.

Qualitative Reference Datasets

Reference datasets of variants for DNA reference materials typically include two components: benchmark variants and benchmark regions. Benchmark variants are a group of well-characterized and highly confident variants on the genomic sequences of a DNA reference material. These variants are developed with corresponding benchmark regions that include positions of the benchmark variants and homozygous reference positions.

To ensure the accuracy of benchmark variants, four approaches are commonly employed. First, benchmark variants are developed by integrating data from multiple sequencing technologies and bioinformatic algorithms to take advantage of the strengths of different methods, while carefully filtering out errors introduced from individual runs. To increase confidence in the benchmark variants, majority voting methods are often applied to select variants that are consistent among replicates. However, it is important to note that even reproducible variants may not always be true variants, as systematic errors shared across multiple methods can also be present (Robasky et al. 2014). Consensus genotype calls or *in silico* datasets can also be used to train machine learning models to find the optimal classification threshold to identify likely false positives. For example, the GIAB consortium used concordant genotype calls to train a simple one-class model for each dataset to determine whether each call from each dataset might be biased (Zook et al. 2019). Another example is the SEQC2 study, which spiked *in silico* SNVs and indels into normal replicates using BAMSurgeon to create "pseudo-tumors" (Fang et al. 2021). Variants detected by virtual tumor-normal pairs that were not spiked in were labeled as false positives. About 100 genomic and sequencing features were extracted to train adaptively boosted classifiers, which were used to classify variants called from real tumor-normal pairs into four confidence levels.

Second, pedigree information can be used to remove technical errors when establishing reference datasets for germline variants. Since the number of Mendelian inconsistent

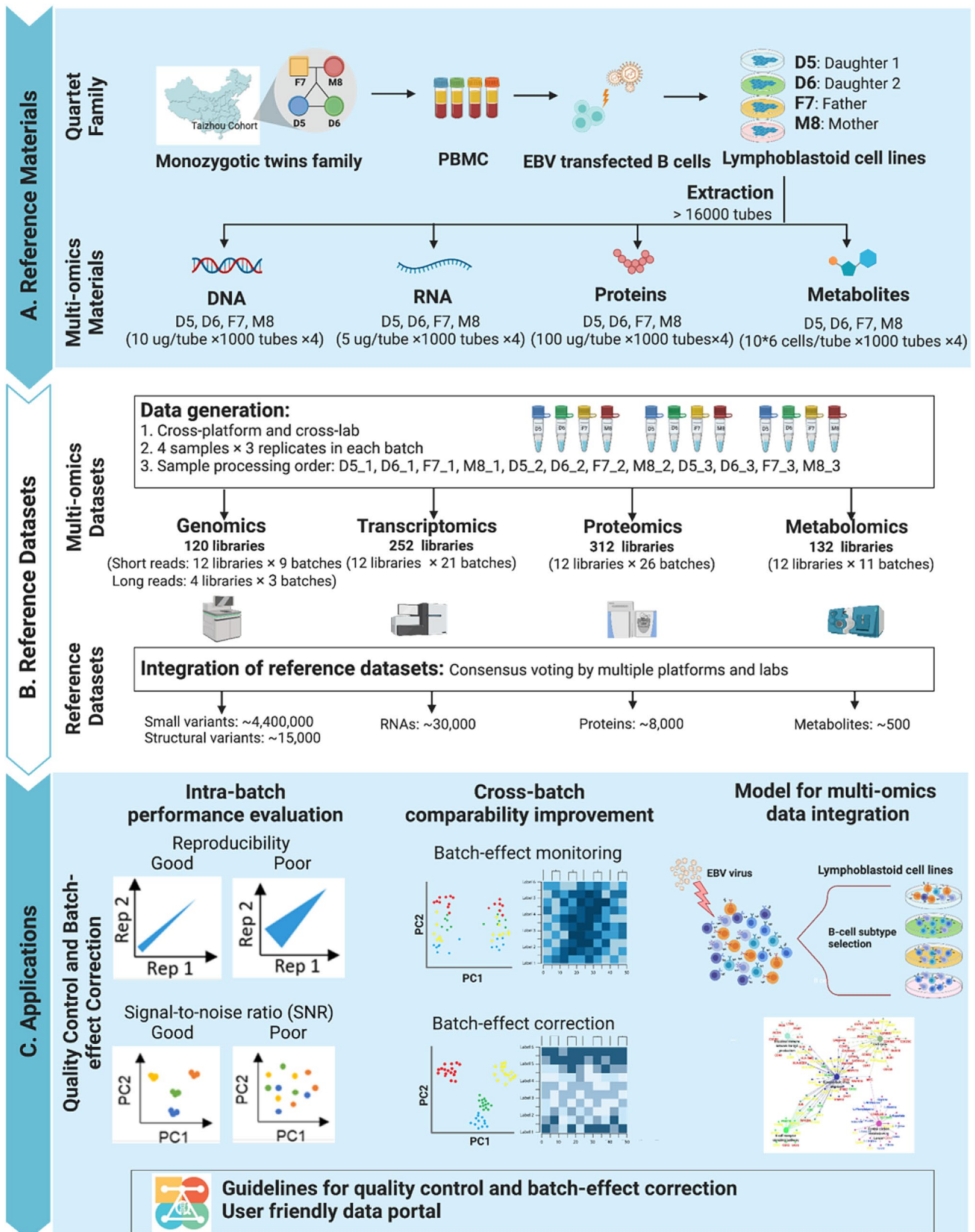


Fig. 1 The quartet project for quality control of multiomics profiling

variants is far more than that of de novo variants and somatic variants arisen somatically or from cell culture, they are potential technical artifacts (Conrad et al. 2011). Illumina released another version of small-variant benchmark calls for NA12878—“the Platinum Genome” (Eberle et al. 2017). Although the Platinum Genome was integrated from sequencing datasets generated from a single Illumina sequencing platform and called by multiple pipelines, its accuracy was validated by using haplotype inheritance information through a well-studied 17-member pedigree. Benchmark variants of the Quartet samples are required to be the same between the monozygotic twin daughters and follow Mendelian inheritance law with parents.

Third, validation of draft benchmark variants using orthogonal technologies with different principles is necessary to confirm their reliability. Sanger sequencing is widely recognized as the “gold standard” method for validating variant calls from high-throughput sequencing. Additionally, array-based genotyping and amplicon-based sequencing can be designed to validate variants in specific regions of interest. Given the impracticality of validating millions of variants across the entire human genome simultaneously, a small number of variants with different confidence levels are randomly selected for validation. Variants with the highest confidence level are expected to be fully supported by orthogonal technologies, and validation rate drops as the confidence level decreases (Fang et al. 2021). An alternative way is to focus on validating suspicious false positives. It should be noticed that discrepancies by orthogonal technologies do not necessarily indicate errors in benchmark variants, because variants on repetitive regions are unlikely to be easily characterized by sequencing technologies mentioned above, and long-read sequencing are more useful for such variants.

Fourth, manual inspection is necessary on discrepancies between benchmark variants and orthogonal validation. To ensure the accuracy and completeness of the benchmark sets, the GIAB consortium has established a process that involves sharing a draft benchmark with GIAB after initial evaluations at NIST, inviting volunteers with expertise in different technologies to contribute callsets, comparing these callsets to the draft benchmark, and randomly selecting putative false positives and false negatives for curation by the callset contributors (Wagner et al. 2022). Any sites identified as questionable or errors in the benchmark are re-curated by NIST. If the majority of false positives or false negatives are found to be errors or questionable in the benchmark, a new version of the benchmark is developed. This process ensures the continuous improvement and refinement of the benchmark sets, leading to more reliable and accurate variant calls.

Benchmark regions, also known as high-confidence genomic regions, are areas where accurate genotypes can be reliably derived, and sites within them are either benchmark calls or homozygous reference calls. Benchmark regions are

often integrated from callable regions of multiple sequencing datasets that have relatively high mapping quality and sequencing coverage. Low complexity regions and highly repetitive regions are typically excluded to avoid possible systematic mapping errors, and flanking regions of uncertain variants are also excluded. Concordant variants outside of the high-confidence regions are not considered as benchmark variants due to their lower confidence level. Benchmark regions can be continually updated with advances in sequencing technologies, assembly algorithms, and variant calling methods. Linked-read and long-read sequencing technologies have been utilized to provide better coverage of difficult-to-map and repetitive regions. The benchmark regions for small variants have been expanded from 77% to 92% of the autosomes and X chromosome of GRCh38 by including long reads (Chin et al. 2020; Wagner et al. 2022). GIAB has collaborated with the Human Pangenome Reference Consortium (HPRC) and the Telomere-to-Telomere Consortium (T2T) to expand benchmark sets by utilizing T2T assemblies of the HG002 genome, initially focusing on chromosomes X and Y and later expanding to the entire genome (Nurk et al. 2022).

Identification of transcripts, proteins and metabolites is a critical step in interpreting omics datasets. In fact, the number of identified features (protein or metabolites) has been routinely used as a performance metric. As we have no idea what substances are expected to be present in biological reference materials, features are retained in the reference datasets if they are consistently detected among multiple replicates, platforms and algorithms with small coefficient of variation (CV) of quantitative measurements (Aristizabal-Henao et al. 2021; Davis et al. 2019). For reference materials with two or more sample groups, the number and identities of differentially expressed features between sample pairs are additional qualitative properties for quantitative omics.

Quantitative Reference Datasets

Quantitative properties fall into two categories: absolute and relative quantification. Absolute quantification is to determine absolute copy numbers of transcripts or absolute concentrations of proteins and metabolites, which can be achieved by using a standard curve created by dilution series of internal standards for each substance. In the MAQC study, the expression levels of 1044 genes in samples A and B were measured by TaqMan qPCR assays, which were later used as orthogonal gold standard to assess the accuracy of microarray and RNA-seq (MAQC Consortium 2006, 2014). In the Quartet study, absolute quantification of Quartet protein reference materials was performed by using C^{13} stable isotope-labeled concatenated peptides. They randomly selected 33 proteins with intensity-based absolute quantification values distributed in four orders of magnitude as anchor proteins for

absolute quantification and calibration of the copy numbers for the whole proteome. The abundance of more than 4000 proteins in each of four Quartet samples were quantified by aligning to the anchor proteins (Tian et al. 2023). NIST reported mass fraction, mass concentration and amount-of-substance concentration of metabolites for SRM 1950 by combining different extraction procedures, analytical methods, chromatographic separations and LC detection modes (Simon-Manso et al. 2013).

Absolute quantification is difficult to achieve in RNA-seq and untargeted metabolomics and proteomics. The original signals from a single sample, such as fragments per kilobase of transcript per million mapped reads (FPKM) in transcriptomics, fraction of total (FOT) in MS-based proteomics and relative peak areas in metabolomics, from different platforms, labs and batches are incomparable due to technical biases and batch effects (Zheng et al. 2023). These technologies are usually applied to compare expression profiles between control and test groups to identify differentially expressed features. Reference datasets of relative quantification, which determines fold changes in expression levels between different sample groups, is vital for performance assessment of such technologies (MAQC Consortium 2006, 2014). Recently, the Quartet Project established the first ratio-based quantitative reference datasets for

transcriptomics, proteomics and metabolomics, through converting the original signals to relative quantitative measurements by dividing the expression profiles of study samples by those of a universal reference material on a feature-by-feature basis (Zheng et al. 2023).

Performance Metrics

Performance metrics are essential for evaluating the quality of a test dataset or experiment. These metrics are categorized into two groups: reference dataset-dependent metrics and reference dataset-independent metrics (Fig. 2). Reference dataset-dependent metrics assess the performance of variant calls and expression profiles by comparing them to a reference dataset. On the other hand, reference dataset-independent metrics evaluate the performance of the experiment without using a reference dataset, usually by the reproducibility of replicates or built-in truth of multi-sample reference materials.

Reference Datasets Dependent

The evaluation of variant calls can be accomplished by comparing them to benchmark variants in benchmark regions.

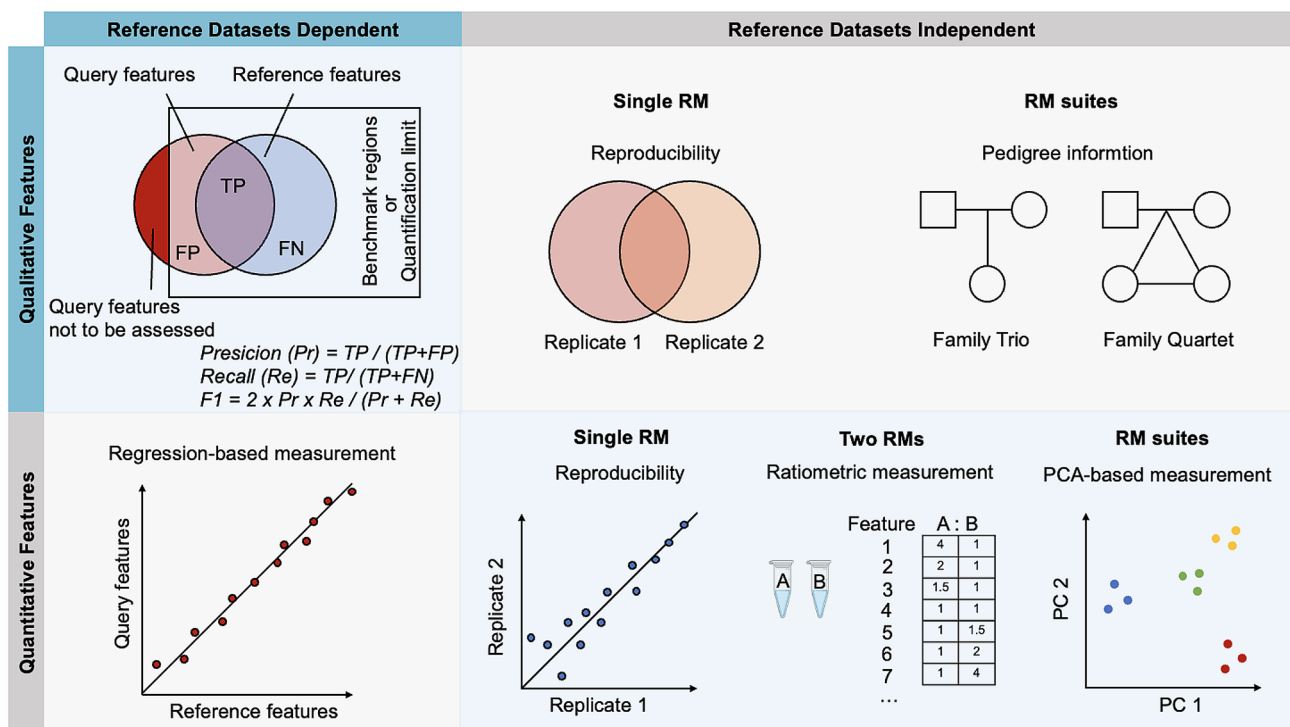


Fig. 2 Performance metrics based on reference materials in multiomics profiling. When reference datasets are available, the performance of a query dataset can be assessed by comparing its consistency with the reference datasets. In cases where reference datasets are not avail-

able, the performance of a query dataset can be assessed by evaluating the reproducibility between replicates or utilizing built-in truth within multi-sample reference materials

Two commonly used measures are precision, which refers to the fraction of called variants that are also benchmark variants, and recall or sensitivity, which is the fraction of benchmark variants that have been detected. Precision and recall are often in a trade-off relationship; improving one measure may result in a decrease in the other. The recall rate is not affected by false positives and can be increased by relaxing variant filtration thresholds, which may lead to detecting more putative variants. Therefore, it is necessary to assess both precision and recall to avoid over-inflating the recall rate. The F1-score is a commonly used measure that takes into account both false positives and false negatives by computing a harmonic mean of precision and recall. Specificity, which describes how many of all homozygous reference sites were correctly detected as non-variant sites, is less useful for variant calling performance evaluation because there is typically a three-orders-of-magnitude difference in the number of variants and the number of bases in a genomic region.

Benchmark variants developed from existing technologies and platforms are often limited to more easily detected variants and regions, which may lead to overestimating the overall performance of an assay when it is used to call variants outside of the high-confidence or benchmark regions. Variants outside of benchmark calls are often located in complex genomic regions and are less concordant between different callsets. In particular, the precision of an assay may be greatly inflated by using only benchmark regions.

As a variant or group of variants may be differently represented between variant call format (VCF) files, variant representation should be normalized before benchmarking to greatly reduce ambiguities. The Global Alliance for Genomics and Health (GA4GH) Benchmarking Team and GIAB have developed best practices and methods for benchmarking small germline variant calls (Kumaran et al. 2019), by providing guidance to match variant calls with different representations, define standard performance metrics, and stratify performance by variant type and genome context.

Reference datasets of transcriptomics, proteomics and metabolomics consist of a catalog of highly confident features or differentially expressed features (qualitative properties) and their corresponding expression levels or fold changes between sample groups (quantitative properties). Precision and recall are used to describe the proportion of detected features that are true and the proportion of reference features that are detected, respectively. Instead of F1-score, Matthews correlation coefficient (MCC) was employed as the main evaluation measure in the MAQC projects, because it takes true negatives into consideration and produces a high score only if good results are obtained in all of the four categories in a confusion matrix (Chicco and Jurman 2020). The Pearson correlation coefficient between the expression levels or fold changes of test datasets and reference datasets is used to describe the accuracy of quantitation. The root

mean squared error (RMSE) measures the deviation between predicted and actual expression values. It should be noticed that reference datasets of biological reference materials may be incomplete, as features can be removed if their expression levels are below the limit of quantification, or they can only be measured by specific extraction and analysis procedures that have not been included in the construction of the reference datasets.

Reference Datasets Independent

To diagnose the underlying causes of suboptimal datasets, it is necessary to combine performance metrics for various stages of the experiment including sample or library preparation, sequencing, raw datasets, and variant detection (Fig. 3). This approach can help identify the specific areas where improvements are needed to enhance the quality of the data. High-quality sequencing data improve the quality and comparability of profiling results. Confirming the quality of sequencing libraries before committing them to sequencing runs increases the chances of success. In next- and third generation sequencing-based genomics and transcriptomics, important performance metrics for run-time sequencing quality include the number of reads or bases produced in the run, the percentage of bases called incorrectly at any one cycle, the percentage of bases with a base quality score, cluster density on the flow cell, and library complexity (Patel and Jain 2012). Raw read quality control assessments help filter out low-quality reads and trim low-quality bases at both ends of a read. After the mapping of preprocessed reads, sequencing alignment performance metrics can be used to assist in detecting biases in sequencing and mapping processes. In MS-based proteomics and metabolomics, the quality of datasets is greatly affected by instrument performance. The better the instrument performance, the better the data. Key performance metrics of the instrument extracted from raw datasets include electrospray stability, cleanliness of the ion source, cleanliness of the inner components of the MS, fragmentation efficiency, and mass accuracy (Morgestern et al. 2021). For example, the uniformity of MS1 intensity distribution reflects the consistency of chromatographic spray and mass spectrometry sensitivity, while the uniformity of MS2 intensity distribution reflects the consistency of fragment ion detection sensitivity. There is no universally accepted standard for pre-analytical performance metrics. Appropriate thresholds depend on specific library preparation protocols, sequencers or instruments, and algorithms. While these metrics can be calculated for samples of interest, the use of widely adopted reference materials facilitates better understanding of performance across different assays and laboratories.

In cases where reference datasets are either unavailable or do not contain the features of interest, alternative

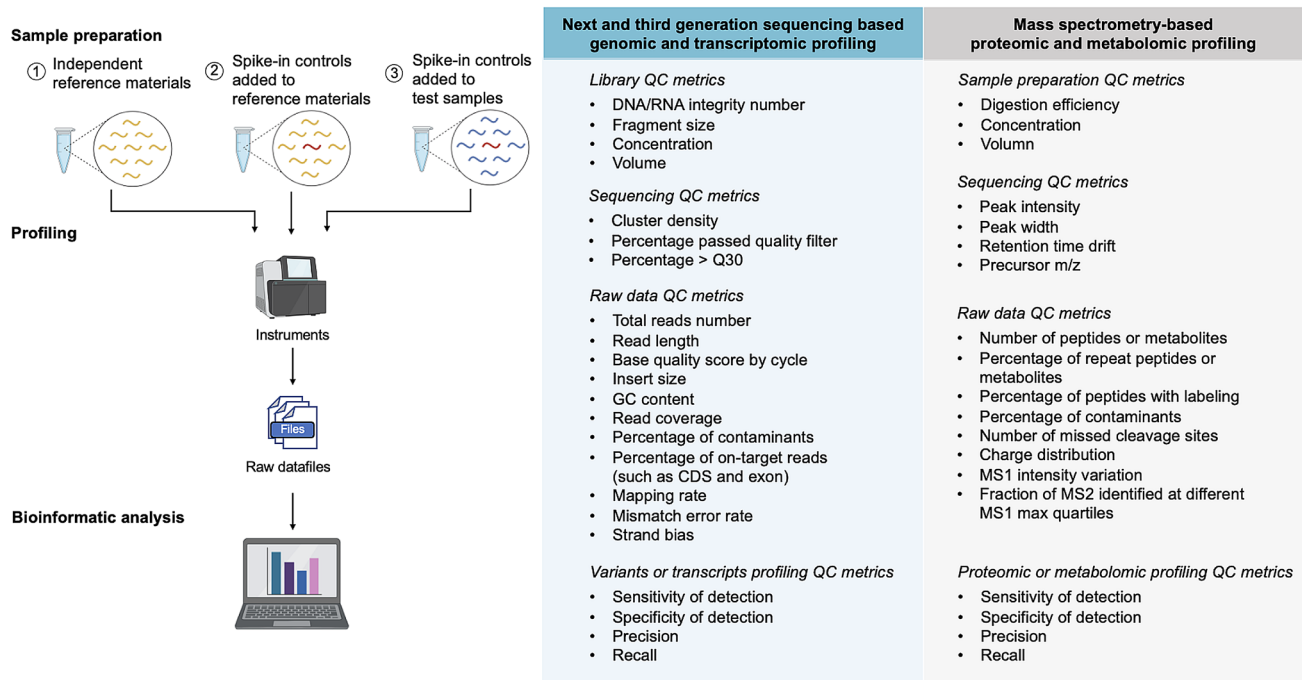


Fig. 3 Schematic overview of multiomics profiling and quality control workflows showing the use of reference materials. Omics reference materials can evaluate all steps of sequencing workflow,

including sample or library preparation, sequencing, raw reads, and profiling. Illustrated is the workflow of sequencing and key performance metrics for each step

methods can be employed to assess the performance. One such method involves evaluating the reproducibility of replicates, which compares the results of multiple measurements conducted on the same sample. Another approach is to utilize built-in truth from multi-sample reference materials, whereby a known standard is employed to evaluate the accuracy of the experiment.

The performance of variants calling results can be assessed by the repeatability and reproducibility of technical replicates or the Mendelian consistent ratio of family members. Technical replicates share the same variant calls and de novo mutations are rare; therefore, the majority of discordant variants is likely to represent genotyping errors (Veltman and Brunner 2012). The advantage of those reference datasets independent metrics is that they can evaluate the precision of variant calling on the whole genome without being restricted to the benchmark regions. However, these metrics cannot indicate how many true variants should be identified, or what the recall rate is.

To assess the accuracy of quantitative omics, three levels of reference dataset-independent metrics can be employed based on the number of available reference materials (Fig. 2). If a single reference material is available, the reproducibility between technical replicates is used to assess the performance of profiling results. However, a high correlation between two replicates of the same sample is not enough to ensure accuracy in detecting differences between sample

groups, because the replicates may share the same technical biases. If a pair of RMs is available, fold changes of features between sample pairs are expected to be the same as the designed expression signal ratios. If three or more RMs are available in a suite, PCA-based metrics can be used to assess the performance of distinguishing the intrinsic biological differences between sample groups.

Utilization of Reference Materials

Identifying reliable biomarkers that can accurately predict disease risk or response to treatment is a critical goal of omics-based cohort studies. Large cohort studies that involve collecting samples over a long period of time and profiling the samples with multiple platforms at multiple labs may suffer from issues related to data incomparability and batch effects, which add difficulties for biomarker discovery. In this section, we discuss how omics RMs can be integrated in large cohort studies to enhance the rigor and reproducibility of biomarker discovery (Fig. 4).

To ensure accurate and reliable results from large-scale analysis of precious cohort samples, it is important to assess the suitability of experimental and analytical pipelines using reference materials prior to initiating the data generation process. The first important step is to choose the suitable RMs based on study design and instruments available. Points of

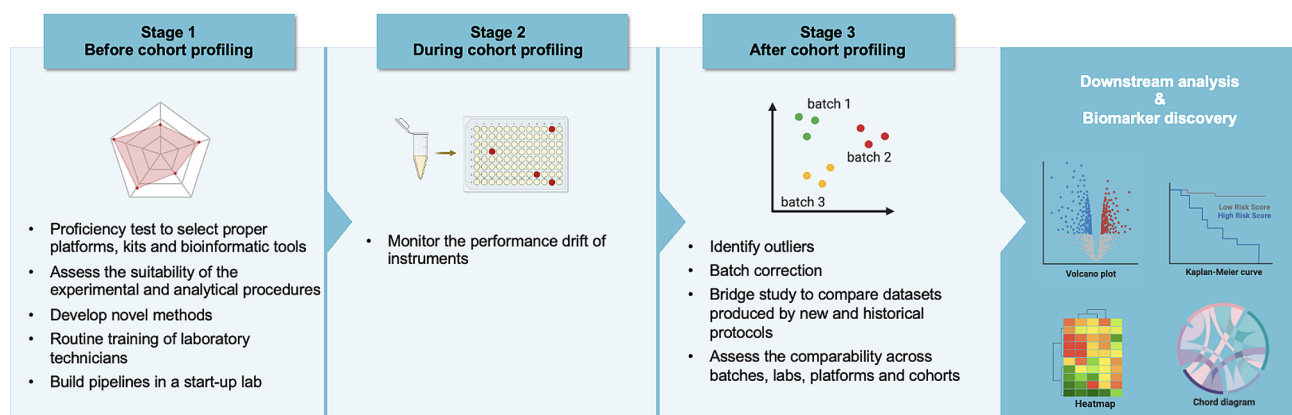


Fig. 4 Utilization of reference materials in large cohort study to enhance the rigor and reproducibility of biomarker discovery

consideration include the availability of RMs, their comparability to the test material, and whether the assigned property values and their confidence levels include the features of interest. The matrix composition of RMs is a critical consideration in the QC process of LC–MS. The performance indicators, such as calibration effectiveness, extraction efficiency, column performance, and ion suppression level, are directly influenced by the composition of the sample matrix. To ensure accurate and reliable performance assessment, it is recommended to employ RMs with a matrix composition as similar as possible to that of the study samples.

At each omics level, a variety of sample preparation methods, data generation platforms, and bioinformatic tools are available. By utilizing RMs in benchmark studies and proficiency test, researchers can gain insights into the strengths and limitations of various methods and technologies. This knowledge facilitates the selection of appropriate experimental and analytical procedures tailored to the specific goals, samples types, and available resources.

RMs can also be effectively used to optimize protocols and parameters by identifying and troubleshooting potential issues. For example, in genomics and transcriptomics by NGS, sequencing performance is influenced by the insert fragment size, which is associated with DNA shearing time. Longer shearing time produces shorter DNA fragments, and the insert fragment sizes must be measured to ensure that they fall within the expected molecular weight range (Fang et al. 2021). In MS-based proteomics and metabolomics, RMs can be used to assess mass-to-charge (m/z) ratio and chromatographic characteristics such as retention time, peak area, and peak shape, by comparing them to predefined acceptance criteria (Nakayasu et al. 2021). If the acceptance criteria are not met, corrective maintenance of the instruments or verification of reagents should be performed until the system suitability meets the requirements. Furthermore, reference materials can be used to train laboratory technicians to perform optimally in daily practice.

When conducting large-scale profiling of cohort samples, it's important to incorporate RMs into the experimental design to objectively monitor and evaluate the longitudinal stability of instruments and assays. In MS-based proteomic and metabolomic studies, RMs are typically run before and after a block of samples to monitor instrument performance drift and to ensure optimal settings are being used (Bittremieux et al. 2018). The block size is determined based on the expected performance drift over time and the separation length. In genomics and transcriptomics, DNA and RNA reference materials can be added to each batch of samples to be sequenced (Ren et al. 2023; Yu et al. 2023). This helps monitor the stability and comparability of analytical instruments across different batches, assays, and labs.

After large-scale profiling, datasets of both cohort samples and RMs are obtained. Datasets from the same RMs can reveal batch effects across labs, platforms, and time points. To eliminate batch effects, cohort sample datasets can be aligned to a common RM, removing unwanted variation and increasing comparability and statistical power, leading to greater confidence in biological insights from combined datasets of multiple batches. Sometimes, large-scale studies take a long time to complete and sequencing technologies can be updated, or new cohorts are needed to address important scientific questions. In such cases, bridge studies can be performed to compare the comparability of novel and historical protocols by using common RMs in both studies.

To ensure quality control in quantitative omics studies, the utilization of multi-sample RMs is essential, especially when investigating differences between cases and controls or various disease subtypes. In most proteomic and metabolomic studies, the reproducibility within technical replicates of a single reference material is commonly employed to evaluate dataset quality. Assessing the reproducibility of RMs is crucial for determining the stability and precision of an analytical method. However, relying solely on reproducibility may not be sufficient for accurately identifying

biological differences between sample groups, as technical biases can impact the absolute abundance of measurements without affecting the relative differences between sample groups (Yu et al. 2023; Zheng et al. 2023). To achieve statistical significance, at least three replicates for each RM are necessary. In a PCA plot, a high-quality dataset is expected to show both separation between sample groups and tight clustering of replicates from the same sample. This indicates that the technical variation is under control and that the biological differences between sample groups are significant. In contrast, a low-quality dataset may show overlapping or scattered sample groups, indicating a high level of technical variation or noise that could obscure the biological differences between the groups.

Challenges and Future Directions

High-throughput profiling technologies have revolutionized omics studies by enabling the generation of vast amounts of data in a relatively short period of time, allowing researchers to comprehensively study complex biological systems at an unprecedented level of resolution. However, performing high-throughput profiling is a highly complex and challenging process, and there are many potential sources of variability that can impact the results and reproducibility. Therefore, rigorous QA/QC is crucial to ensure confidence in the resulting data and biological discoveries. The use of RMs is an important aspect of QA/QC in high-throughput technologies to ensure accurate and reliable results. In this review, we aim to offer a comprehensive overview of the significance of utilizing well-characterized RMs across different levels of omics research, including genomics, transcriptomics, proteomics, and metabolomics. We provide insights into the characteristics, advantages, and limitations of RMs in each omics field, which are summarized in Table 2. Our goal is to assist researchers in making informed decisions when selecting suitable RMs for their specific research questions and analytical methods. Ultimately, the utilization of appropriate RMs can greatly enhance the accuracy and reliability of omics research outcomes.

By incorporating well-characterized RMs into omics research, researchers can overcome various challenges and limitations. RMs provide a standardized reference point that enables calibration and quality control throughout the experimental workflow. They serve as valuable tools for method optimization, validation, and troubleshooting, allowing researchers to assess the performance of their analytical methods and identify any potential biases or errors. Furthermore, the use of RMs facilitates inter-laboratory comparisons and promotes data harmonization, enabling the integration and comparison of results across different studies and platforms. Although the profiling of RMs may

entail additional costs, implementing a thorough QA/QC methodology is important for evaluating and monitoring the performance of data generation processes. This upfront investment contributes to the long-term reliability and accuracy of the results, minimizing potential errors and ensuring the accuracy and reliability of the omics research.

The careful selection of RMs is crucial to ensure their relevance and applicability to the study at hand. Researchers should consider the intended use of the study and choose RMs that closely resemble the properties of the samples being investigated. Additionally, the selected RMs should be qualitatively and quantitatively representative of the entire collection of samples included in the study. This ensures that the RMs effectively mimic the characteristics of the biological samples, enabling accurate and meaningful comparisons and interpretations. When studying specific genetic or phenotypical features that vary among different ethnic groups, it is important to choose RMs that match the ethnicity of the study samples. This approach ensures that the RMs accurately reflect the characteristics of the study population, enabling the assessment of the detection performance of those specific genetic or phenotypical features (Hardwick et al. 2017).

As profiling methods continue to advance and new technologies emerge, the reference datasets for existing RMs will undergo continuous updates and refinements. One example of this is the utilization of long reads in genomic sequencing. Long reads are particularly valuable for profiling repetitive and complex regions, which are challenging to be mapped by short reads (Wenger et al. 2019). By incorporating long reads, benchmark variants in these regions can be better characterized (Wagner et al. 2022). Additionally, long-read technologies enable precise transcript detection and RNA modifications (Leger et al. 2021; Sonesson et al. 2019). In proteomics, MS techniques are extensively used to study post-translational modification (PTMs) of proteins (Zecha et al. 2022). The reference materials will expand to encompass more omics types along with the development of technologies. For example, reference datasets of DNA epigenomics for DNA RMs can be developed, RNA RMs can include small RNA profiling and RNA modification reference datasets, and protein RMs can incorporate PTM reference datasets.

Challenges persist in the global promotion and adoption of reference materials and reference datasets. First, regulatory challenges, especially across different regions of the world, can pose additional obstacles in adopting a universal RM (Guerrier et al. 2012; Krogstad et al. 2010). Biological RMs, especially those intended for human genomics and transcriptomics, which are frequently derived from human specimens, require stricter adherence to informed consent principles and governmental controls. Currently, there is no single, comprehensive international model for governing

Table 2 Characteristics of genomic, transcriptomic, proteomic, and metabolomic reference materials

Genomic reference materials		Engineered	Synthetic
Characteristics	Biological		
Source	Genomic DNA from cancer cell lines or lymphoblastoid cell lines	Genomic DNA of cell lines with specific genomic variants engineered at predetermined positions	Artificially generated DNA sequences that are designed to mimic specific genomic regions or variants but cannot be aligned to human reference genome
Intended usage	Performance assessment of variant calling from whole-genome, whole-exome and targeted sequencing	Performance assessment of genetic testing	Spiked into study samples as internal controls; performance assessment of quantitative genomic features (VAF, copy number, etc.)
Reference datasets	Millions of benchmark variants are integrated from multiple platforms and bioinformatic pipelines, but comprehensive characterization on whole genome is challenging	Limited number of engineered variants	Known DNA sequences and must establish commutability with biological samples before use
Representativeness	Readily commutable with biological samples, but not representative of population genetic diversity or tumor genome heterogeneity	Readily commutable with patient samples	May not fully represent the complexity of natural genomes
Preparation	Genomic instability and drift in cell lines can lead to lot-to-lot variability	Can cause off-target mutations and unintended effects	Represent any variant or sequence of interest, limited only by the constraints of synthesis
Consent and privacy considerations	Yes	No	No
Transcriptomic reference materials			
Characteristics	Biological		Synthetic
Source	RNA from natural sources, such as tissues, cell lines, etc.		Artificially generated RNA sequences
Intended usage	Performance assessment of differential gene expression from RNA sequencing, microarray, ddPCR and qPCR	Performance assessment of differential gene expression, dynamic range from RNA sequencing	Spiked into study samples as internal controls Performance assessment of differential gene expression, sensitivity (lower LOD) and dynamic range from RNA sequencing
Reference datasets	Fold changes of differential expressed genes integrated from multiple technologies, platforms and labs Gene with relatively low expression may not be included	Known RNA sequences with no homology to natural reference genomes Defined fold changes between two synthetic controls	
Representativeness	Reflect the complexity and diversity of gene-expression patterns observed in biological samples, but do not fully capture the entire spectrum of transcriptomic variations and complexities present in different biological contexts	May not fully capture the complexity and biological context present in a natural system	
Preparation	RNA exhibits variations across batches, even derived from the same cell lines RNA is prone to degradation	Synthesized with high precision and homogeneity across batches	
Consent and privacy considerations	Yes	No	No

Table 2 (continued)

Proteomic reference materials	
Characteristics	Biological
Source	Proteins from natural sources, such as cell lysates, tissues, or bodily fluids
Intended usage	Performance assessment of identification and quantification for proteins
Reference datasets	The complete proteome may not be fully characterized, only high confident proteins are identified by multiple platforms and techniques, with information of the identity and concentrations of proteins
Representativeness	Reflect the complexity and diversity of the proteome in biological systems, but do not capture the entire range of proteomic variations and complexities observed in different biological contexts
Preparation	Proteins exhibit variations across batches, even derived from the same cell lines Protein reference materials prepared as peptides cannot assess the performance of digestion methods
Consent and privacy considerations	No
Metabolomic reference materials	
Characteristics	Biological
Source	Metabolites from natural sources, such as tissues, cells, bodily fluids or microbial cultures
Intended usage	Performance assessment of identification and quantification for metabolites
Reference datasets	The complete metabolome may not be fully characterized, only high confident metabolites are identified by multiple platforms and techniques, with the information of the identity and concentrations of metabolites
Representativeness	Represent snapshot of the metabolites present in a specific biological system, but do not capture the entire range of metabolome variations and complexities observed in different biological contexts and conditions
Preparation	Metabolites exhibit variations across batches, even derived from the same cell lines
Consent and privacy considerations	No
Synthetic	
Source	Chemical synthetic metabolites or isotopically labeled metabolites
Intended usage	Spiked into study samples as internal controls Performance assessment of identification and quantification for metabolites
Reference datasets	Specific metabolites with defined and known composition and concentration
Representativeness	May not fully capture the complexity and biological context present in a natural system
Preparation	Synthesized with high precision and homogeneity across batches
Consent and privacy considerations	No

human genetic resources. The distinct nature of informed consent across different countries, influenced by diverse cultures and social traditions, necessitated addressing legal, ethical, and logistical aspects related to genetic materials and data utilization while respecting each nation's sovereignty and cultural norms. International collaboration and agreements are imperative in addressing these challenges and ensuring the conscientious and equitable utilization of human genetic resources worldwide (Gainotti et al. 2016; van Belle et al. 2015).

Second, we strongly recommend that QC data should be made available alongside the study samples in databases or repositories that adhere to the FAIR principles (Findable, Accessible, Interoperable, and Reusable), which is crucial for enhancing data management and sharing (Conesa and Beck 2019; Wilkinson et al. 2016). Currently, QC information is often omitted from scientific publications, leading to uncertainty about the performance methodology used. In the future, guidelines may be developed to mandate the inclusion of QC metrics in data submissions to public repositories, similar to existing guidelines for other aspects of data reporting. Coupling comprehensive QC information to the experimental data will allow for quick assessment of the reliability of an experiment, which is crucial in light of recent reports of the general reproducibility crisis in various scientific fields (Anonymous 2021; Baker 2016; Shi et al. 2017). It is essential to prioritize and formalize QC practices to ensure the quality and reproducibility of high-throughput multiomics profiling results by fully utilizing well-characterized RMs and appropriate QC metrics.

Conclusion

In this review, we summarized reference materials across all levels of omics, including (epi-)genomics, transcriptomics, proteomics, and metabolomics. We have offered a comprehensive overview of leveraging omics reference materials to enhance data quality. This initiative is geared toward promoting robust scientific research and advancing our understanding of complex biological systems through the thoughtful application of omics technologies.

Acknowledgements This study was supported in part by Shanghai Sailing Program (22YF1403500), the National Natural Science Foundation of China (32300536, 31720103909 and 32170657), the National Key R&D Project of China (2018YFE0201603 and 2018YFE0201600), State Key Laboratory of Genetic Engineering (SKLGE-2117), and the 111 Project (B13016). Some of the illustrations in this paper were created with BioRender.com.

Authors' Contributions Manuscript writing: LR, LS, and YZ. All authors approved this manuscript.

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Conflict of Interest The authors declare no competing financial interests.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Ethical Approval Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amos Wilson J, Pratt VM, Phansalkar A, Muralidharan K, Highsmith WE Jr, Beck JC, Bridgeman S, Courtney EM, Epp L, Ferreira-Gonzalez A, Hjelm NL, Holtegaard LM, Jama MA, Jakupciak JP, Johnson MA, Labrousse P, Lyon E, Prior TW, Richards CS, Richie KL, Roa BB, Rohlf EM, Sellers T, Sherman SL, Siegrist KA, Silverman LM, Wiszniewska J, Kalman LV, Fragile Xperts Working Group of the Association for Molecular Pathology Clinical Practice C (2008) Consensus characterization of 16 FMR1 reference materials: a consortium study. *J Mol Diagn* 10(1):2–12. <https://doi.org/10.2353/jmoldx.2008.070105>
- Andrews PC, Arnott DP, Gawinowicz MA, Kowalak JA, Lane WS, Lilley KS, Martin LT, Stein S (2006) ABRF-sPRG 2006 study: a proteomics standard. ABRF 2006: Long Beach, CA, 2006
- Anonymous (2021) Replicating scientific results is tough—but essential. *Nature* 600(7889):359–360. <https://doi.org/10.1038/d41586-021-03736-4>
- Anwaier A, Zhu SX, Tian X, Xu WH, Wang Y, Palihati M, Wang WY, Shi GH, Qu YY, Zhang HL, Ye DW (2022) Large-scale proteomics data reveal integrated prognosis-related protein signatures and role of SMAD4 and RAD50 in prognosis and immune infiltrations of prostate cancer microenvironment. *Phenomics* 2(6):404–418. <https://doi.org/10.1007/s43657-022-00070-1>
- Aristizabal-Henao JJ, Jones CM, Lippa KA, Bowden JA (2020) Non-targeted lipidomics of novel human plasma reference materials: hypertriglyceridemic, diabetic, and African-American. *Anal Bioanal Chem* 412(27):7373–7380. <https://doi.org/10.1007/s00216-020-02910-3>
- Aristizabal-Henao JJ, Lemas DJ, Griffin EK, Costa KA, Camacho C, Bowden JA (2021) Metabolomic profiling of biological reference materials using a multiplatform high-resolution mass spectrometric approach. *J Am Soc Mass Spectrom* 32(9):2481–2489. <https://doi.org/10.1021/jasms.1c00194>
- Athieniti E, Spyrou GM (2023) A guide to multi-omics data collection and integration for translational medicine. *Comput Struct Biotechnol J* 21:134–149. <https://doi.org/10.1016/j.csbj.2022.11.050>

- Azab S, Ly R, Britz-McKibbin P (2019) Robust method for high-throughput screening of fatty acids by multisegment injection-nonaqueous capillary electrophoresis-mass spectrometry with stringent quality control. *Anal Chem* 91(3):2329–2336. <https://doi.org/10.1021/acs.analchem.8b05054>
- Baker M (2016) 1500 scientists lift the lid on reproducibility. *Nature* 533(7604):452–454. <https://doi.org/10.1038/533452a>
- Baker SC, Bauer SR, Beyer RP, Brenton JD, Bromley B, Burrill J, Causton H, Conley MP, Elespuru R, Fero M, Foy C, Fuscoe J, Gao X, Gerhold DL, Gilles P, Goodsaid F, Guo X, Hackett J, Hockett RD, Ikononi P, Irizarry RA, Kawasaki ES, Kaysser-Kranich T, Kerr K, Kiser G, Koch WH, Lee KY, Liu C, Liu ZL, Lucas A, Manohar CF, Miyada G, Modrusan Z, Parkes H, Puri RK, Reid L, Ryder TB, Salit M, Samaha RR, Scherf U, Sendera TJ, Setterquist RA, Shi L, Shippy R, Soriano JV, Wagar EA, Warrington JA, Williams M, Wilmer F, Wilson M, Wolber PK, Wu X, Zadro R, External RNACC (2005) The external RNA controls consortium: a progress report. *Nat Methods* 2(10):731–734. <https://doi.org/10.1038/nmeth1005-731>
- Beasley-Green A, Bunk D, Rudnick P, Kilpatrick L, Phinney K (2012) A proteomics performance standard to support measurement quality in proteomics. *Proteomics* 12(7):923–931. <https://doi.org/10.1002/pmic.201100522>
- Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, Lindon JC, Nicholson JK (2007) Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat Protoc* 2(11):2692–2703. <https://doi.org/10.1038/nprot.2007.376>
- Beger RD, Dunn WB, Bandukwala A, Bethan B, Broadhurst D, Clish CB, Dasari S, Derr L, Evans A, Fischer S, Flynn T, Hartung T, Herrington D, Higashi R, Hsu PC, Jones C, Kachman M, Karuso H, Kruppa G, Lippa K, Maruvada P, Mosley J, Ntai I, O'Donovan C, Playdon M, Raftery D, Shaughnessy D, Souza A, Spaeder T, Spalholz B, Tayyari F, Ubhi B, Verma M, Walk T, Wilson I, Witkin K, Bearden DW, Zanetti KA (2019) Towards quality assurance and quality control in untargeted metabolomics studies. *Metabolomics* 15(1):4. <https://doi.org/10.1007/s11306-018-1460-7>
- Begley CG, Ioannidis JP (2015) Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res* 116(1):116–126. <https://doi.org/10.1161/CIRCRESAHA.114.303819>
- Bell AW, Deutsch EW, Au CE, Kearney RE, Beavis R, Sechi S, Nilsson T, Bergeron JJ, Group HTSW (2009) A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nat Methods* 6(6):423–430. <https://doi.org/10.1038/nmeth.1333>
- Bettinotti MP, Ferriola D, Duke JL, Mosbrugger TL, Tairis N, Jennings L, Kalman LV, Monos D (2018) Characterization of 108 genomic DNA reference materials for 11 human leukocyte antigen loci: a GeT-RM collaborative project. *J Mol Diagn* 20(5):703–715. <https://doi.org/10.1016/j.jmoldx.2018.05.009>
- Biocrates (2023) Biocrates metabolomics technology. <https://biocrates.com/>
- Bittremieux W, Walzer M, Tenzer S, Zhu W, Salek RM, Eisenacher M, Tabb DL (2017) The human proteome organization-proteomics standards initiative quality control working group: making quality control more accessible for biological mass spectrometry. *Anal Chem* 89(8):4474–4479. <https://doi.org/10.1021/acs.analchem.6b04310>
- Bittremieux W, Tabb DL, Impens F, Staes A, Timmerman E, Martens L, Laukens K (2018) Quality control in mass spectrometry-based proteomics. *Mass Spectrom Rev* 37(5):697–711. <https://doi.org/10.1002/mas.21544>
- Blackburn J, Wong T, Madala BS, Barker C, Hardwick SA, Reis ALM, Deveson IW, Mercer TR (2019) Use of synthetic DNA spike-in controls (sequins) for human genome sequencing. *Nat Protoc* 14(7):2119–2151. <https://doi.org/10.1038/s41596-019-0175-1>
- Blume JE, Manning WC, Troiano G, Hornburg D, Figa M, Hesterberg L, Platt TL, Zhao X, Cuaresma RA, Everley PA, Ko M, Liou H, Mahoney M, Ferdosi S, Elgierari EM, Stolarczyk C, Tangeysh B, Xia H, Benz R, Siddiqui A, Carr SA, Ma P, Langer R, Farias V, Farokhzad OC (2020) Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nat Commun* 11(1):3662. <https://doi.org/10.1038/s41467-020-17033-7>
- Boggs ASP, Kilpatrick LE, Burdette CQ, Tevis DS, Fultz ZA, Nelson MA, Jarrett JM, Kemp JV, Singh RJ, Grebe SKG, Wise SA, Kassim BL, Long SE (2021) Development of a pregnancy-specific reference material for thyroid biomarkers, vitamin D, and nutritional trace elements in serum. *Clin Chem Lab Med* 59(4):671–679. <https://doi.org/10.1515/cclm-2020-0977>
- Bowden JA, Heckert A, Ulmer CZ, Jones CM, Koelmel JP, Abdullah L, Ahonen L, Alnouti Y, Armando AM, Asara JM, Bamba T, Barr JR, Bergquist J, Borchers CH, Brandsma J, Breitkopf SB, Cajka T, Cazenave-Gassiot A, Checa A, Cinel MA, Colas RA, Cremers S, Dennis EA, Evans JE, Fauland A, Fiehn O, Gardner MS, Garrett TJ, Gotlinger KH, Han J, Huang Y, Neo AH, Hyotylainen T, Izumi Y, Jiang H, Jiang H, Jiang J, Kachman M, Kiyonami R, Klavins K, Klose C, Kofeler HC, Kolmert J, Koal T, Koster G, Kuklennyik Z, Kurland JJ, Leadley M, Lin K, Maddipati KR, McDougall D, Meikle PJ, Mellett NA, Monnin C, Moseley MA, Nandakumar R, Oresic M, Patterson R, Peake D, Pierce JS, Post M, Postle AD, Pugh R, Qiu Y, Quehenberger O, Ramrup P, Rees J, Rembiesa B, Reynaud D, Roth MR, Sales S, Schuhmann K, Schwartzman ML, Serhan CN, Shevchenko A, Somerville SE, St John-Williams L, Surma MA, Takeda H, Thakare R, Thompson JW, Torta F, Triebel A, Troztmuller M, Ubhayasekera SJK, Vuckovic D, Weir JM, Welti R, Wenk MR, Wheelock CE, Yao L, Yuan M, Zhao XH, Zhou S (2017) Harmonizing lipidomics: NIST interlaboratory comparison exercise for lipidomics using SRM 1950-Metabolites in Frozen Human Plasma. *J Lipid Res* 58(12):2275–2288. <https://doi.org/10.1194/jlr.M079012>
- Bowden JA, Ulmer CZ, Jones CM, Koelmel JP, Yost RA (2018) NIST lipidomics workflow questionnaire: an assessment of community-wide methodologies and perspectives. *Metabolomics* 14(5):53. <https://doi.org/10.1007/s11306-018-1340-1>
- Broadhurst D, Goodacre R, Reinke SN, Kuligowski J, Wilson ID, Lewis MR, Dunn WB (2018) Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 14(6):72. <https://doi.org/10.1007/s11306-018-1367-3>
- Bunk DM (2010) Design considerations for proteomic reference materials. *Proteomics* 10(23):4220–4225. <https://doi.org/10.1002/pmic.201000242>
- Buttner R, Longshore JW, Lopez-Rios F, Merkelbach-Bruse S, Normanno N, Rouleau E, Penault-Llorca F (2019) Implementing TMB measurement in clinical practice: considerations on assay requirements. *ESMO Open* 4(1):e000442. <https://doi.org/10.1136/esmoopen-2018-000442>
- Cajka T, Smilowitz JT, Fiehn O (2017) Validating quantitative untargeted lipidomics across nine liquid chromatography-high-resolution mass spectrometry platforms. *Anal Chem* 89(22):12360–12368. <https://doi.org/10.1021/acs.analchem.7b03404>
- Cambridge Isotope Laboratories, Inc. (2023) Metabolomics QRSS™ Kit. <https://isotope.com/en-us/metabolomics-mixes-and-kits/metabolomics-qress-kit-msk-qress-kit>
- Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, Huang J, Stuccio S, Zingone A, Ryan BM, Tsang JS, Biancotto A (2017) Assessment of variability in the SOMAscan assay. *Sci Rep* 7(1):14248. <https://doi.org/10.1038/s41598-017-14755-5>

- Candia J, Daya GN, Tanaka T, Ferrucci L, Walker KA (2022) Assessment of variability in the plasma 7k SomaScan proteomics assay. *Sci Rep* 12(1):17147. <https://doi.org/10.1038/s41598-022-22116-0>
- Centers for Disease Control and Prevention (2019) Genetic testing reference materials coordination program. <https://www.cdc.gov/labquality/get-rm/index.html>
- Cheema AK, Asara JM, Wang Y, Neubert TA, Tolstikov V, Turck CW (2015) The ABRF metabolomics research group 2013 study: investigation of spiked compound differences in a human plasma matrix. *J Biomol Tech* 26(3):83–89. <https://doi.org/10.7171/jbt.15-2603-001>
- Chen W, Zhao Y, Chen X, Yang Z, Xu X, Bi Y, Chen V, Li J, Choi H, Ernest B, Tran B, Mehta M, Kumar P, Farmer A, Mir A, Mehra UA, Li JL, Moos M Jr, Xiao W, Wang C (2021) A multicenter study benchmarking single-cell RNA sequencing technologies using reference samples. *Nat Biotechnol* 39(9):1103–1114. <https://doi.org/10.1038/s41587-020-00748-9>
- Chicco D, Jurman G (2020) The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 21(1):6. <https://doi.org/10.1186/s12864-019-6413-7>
- Chin CS, Wagner J, Zeng Q, Garrison E, Garg S, Functammasan A, Rautiainen M, Aganezov S, Kirsche M, Zarate S, Schatz MC, Xiao C, Rowell WJ, Markello C, Farek J, Sedlazeck FJ, Bansal V, Yoo B, Miller N, Zhou X, Carroll A, Barrio AM, Salit M, Marshall T, Dilthey AT, Zook JM (2020) A diploid assembly-based benchmark for variants in the major histocompatibility complex. *Nat Commun* 11(1):4794. <https://doi.org/10.1038/s41467-020-18564-9>
- Chiva C, Mendes Maia T, Panse C, Stejskal K, Douche T, Matondo M, Loew D, Helm D, Rettel M, Mechtler K, Impens F, Nanni P, Shevchenko A, Sabido E (2021) Quality standards in proteomics research facilities: Common standards and quality procedures are essential for proteomics facilities and their users. *EMBO Rep* 22(6):e52626. <https://doi.org/10.15252/embr.202152626>
- Clark DJ, Hu Y, Bocik W, Chen L, Schnaubelt M, Roberts R, Shah P, Whiteley G, Zhang H (2018) Evaluation of NCI-7 cell line panel as a reference material for clinical proteomics. *J Proteome Res* 17(6):2205–2215. <https://doi.org/10.1021/acs.jproteome.8b00165>
- Collins BC, Hunter CL, Liu Y, Schilling B, Rosenberger G, Bader SL, Chan DW, Gibson BW, Gingras AC, Held JM, Hirayama-Kurogi M, Hou G, Krisp C, Larsen B, Lin L, Liu S, Molloy MP, Moritz RL, Ohtsuki S, Schlapbach R, Selevsek N, Thomas SN, Tzeng SC, Zhang H, Aebersold R (2017) Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat Commun* 8(1):291. <https://doi.org/10.1038/s41467-017-00249-5>
- Conesa A, Beck S (2019) Making multi-omics data accessible to researchers. *Sci Data* 6(1):251. <https://doi.org/10.1038/s41597-019-0258-4>
- Conrad DF, Keebler JE, DePristo MA, Lindsay SJ, Zhang Y, Casals F, Idaghdour Y, Hartl CL, Torroja C, Garimella KV, Zilversmit M, Cartwright R, Rouleau GA, Daly M, Stone EA, Hurles ME, Awadalla P, Genomes P (2011) Variation in genome-wide mutation rates within and between human families. *Nat Genet* 43(7):712–714. <https://doi.org/10.1038/ng.862>
- Coriell Institute (2023) Available samples for genetic testing reference materials coordination program. <https://www.coriell.org/1/NIGMS/Additional-Resources/Multiply-Confirmed-Mutations-GeT-RM>
- Craig DW, Nasser S, Corbett R, Chan SK, Murray L, Legendre C, Tembe W, Adkins J, Kim N, Wong S, Baker A, Enriquez D, Pond S, Pleasance E, Mungall AJ, Moore RA, McDaniel T, Ma Y, Jones SJ, Marra MA, Carpten JD, Liang WS (2016) A somatic reference standard for cancer genome sequencing. *Sci Rep* 6:24607. <https://doi.org/10.1038/srep24607>
- Davis WC, Kilpatrick LE, Ellisor DL, Neely BA (2019) Characterization of a human liver reference material fit for proteomics applications. *Sci Data* 6(1):324. <https://doi.org/10.1038/s41597-019-0336-7>
- Deutsch EW, Orchard S, Binz PA, Bittremieux W, Eisenacher M, Hermjakob H, Kawano S, Lam H, Mayer G, Menschaert G, Perez-Riverol Y, Salek RM, Tabb DL, Tenzer S, Vizcaino JA, Walzer M, Jones AR (2017) Proteomics standards initiative: fifteen years of progress and future work. *J Proteome Res* 16(12):4288–4298. <https://doi.org/10.1021/acs.jproteome.7b00370>
- Deveson IW, Chen WY, Wong T, Hardwick SA, Andersen SB, Nielsen LK, Mattick JS, Mercer TR (2016) Representing genetic variation with synthetic DNA standards. *Nat Methods* 13(9):784–791. <https://doi.org/10.1038/nmeth.3957>
- Deveson IW, Madala BS, Blackburn J, Barker C, Wong T, Barton KM, Smith MA, Watkins DN, Mercer TR (2019) Chiral DNA sequences as commutable controls for clinical genomics. *Nat Commun* 10(1):1342. <https://doi.org/10.1038/s41467-019-09272-0>
- Deveson IW, Gong B, Lai K, LoCoco JS, Richmond TA, Schageman J, Zhang Z, Novorodovskaya N, Willey JC, Jones W, Kusko R, Chen G, Madala BS, Blackburn J, Stevanovski I, Bhandari A, Close D, Conroy J, Hubank M, Marella N, Mieczkowski PA, Qiu F, Sebra R, Stetson D, Sun L, Szankasi P, Tan H, Tang LY, Arib H, Best H, Burgher B, Bushel PR, Casey F, Cawley S, Chang CJ, Choi J, Dinis J, Duncan D, Eterovic AK, Feng L, Ghosal A, Giorda K, Glenn S, Happe S, Haseley N, Horvath K, Hung LY, Jarosz M, Kushwaha G, Li D, Li QZ, Li Z, Liu LC, Liu Z, Ma C, Mason CE, Megherbi DB, Morrison T, Pabon-Pena C, Pirooznia M, Proszek PZ, Raymond A, Rindler P, Ringler R, Scherer A, Shakhovich R, Shi T, Smith M, Song P, Strahl M, Thodima VJ, Tom N, Verma S, Wang J, Wu L, Xiao W, Xu C, Yang M, Zhang G, Zhang S, Zhang Y, Shi L, Tong W, Johann DJ Jr, Mercer TR, Xu J, Group SOSW (2021a) Evaluating the analytical validity of circulating tumor DNA sequencing assays for precision oncology. *Nat Biotechnol* 39(9):1115–1128. <https://doi.org/10.1038/s41587-021-00857-z>
- Deveson IW, Gong B, Lai K, LoCoco JS, Richmond TA, Schageman J, Zhang Z, Novorodovskaya N, Willey JC, Jones W, Kusko R, Chen G, Madala BS, Blackburn J, Stevanovski I, Bhandari A, Close D, Conroy J, Hubank M, Marella N, Mieczkowski PA, Qiu F, Sebra R, Stetson D, Sun L, Szankasi P, Tan H, Tang LY, Arib H, Best H, Burgher B, Bushel PR, Casey F, Cawley S, Chang CJ, Choi J, Dinis J, Duncan D, Eterovic AK, Feng L, Ghosal A, Giorda K, Glenn S, Happe S, Haseley N, Horvath K, Hung LY, Jarosz M, Kushwaha G, Li D, Li QZ, Li Z, Liu LC, Liu Z, Ma C, Mason CE, Megherbi DB, Morrison T, Pabon-Pena C, Pirooznia M, Proszek PZ, Raymond A, Rindler P, Ringler R, Scherer A, Shakhovich R, Shi T, Smith M, Song P, Strahl M, Thodima VJ, Tom N, Verma S, Wang J, Wu L, Xiao W, Xu C, Yang M, Zhang G, Zhang S, Zhang Y, Shi L, Tong W, Johann DJ Jr, Mercer TR, Xu J, Group SOSW (2021b) Evaluating the analytical validity of circulating tumor DNA sequencing assays for precision oncology. *Nat Biotechnol*. <https://doi.org/10.1038/s41587-021-00857-z>
- Eberle MA, Fritzilas E, Krusche P, Kallberg M, Moore BL, Bekritsky MA, Iqbal Z, Chuang HY, Humphray SJ, Halpern AL, Kruglyak S, Margulies EH, McVean G, Bentley DR (2017) A reference data set of 5.4 million phased human variants validated by genetic inheritance from sequencing a three-generation 17-member pedigree. *Genome Res* 27(1):157–164. <https://doi.org/10.1101/gr.210500.116>
- Eldjarn GH, Ferkingstad E, Lund SH, Helgason H, Magnusson OT, Gunnarsdottir K, Olafsdottir TA, Halldorsson BV, Olason PI,

- Zink F, Gudjonsson SA, Sveinbjornsson G, Magnusson MI, Helgason A, Oddsson A, Halldorsson GH, Magnusson MK, Saevarsdottir S, Eiriksdottir T, Masson G, Stefansson H, Jonsson I, Holm H, Rafnar T, Melsted P, Saemundsdottir J, Norddahl GL, Thorleifsson G, Ulfarsson MO, Gudbjartsson DF, Thorsteinsdottir U, Sulem P, Stefansson K (2023) Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature* 622(7982):348–358. <https://doi.org/10.1038/s41586-023-06563-x>
- Escher C, Reiter L, MacLean B, Ossola R, Herzog F, Chilton J, MacCoss MJ, Rinner O (2012) Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* 12(8):1111–1121. <https://doi.org/10.1002/pmic.201100463>
- Evans AM, O'Donovan C, Playdon M, Beecher C, Beger RD, Bowden JA, Broadhurst D, Clish CB, Dasari S, Dunn WB, Griffin JL, Hartung T, Hsu PC, Huan T, Jans J, Jones CM, Kachman M, Kleensang A, Lewis MR, Monge ME, Mosley JD, Taylor E, Tayyari F, Theodoridis G, Torta F, Ubhi BK, Vuckovic D, Metabolomics Quality Assurance QCC (2020) Dissemination and analysis of the quality assurance (QA) and quality control (QC) practices of LC-MS based untargeted metabolomics practitioners. *Metabolomics* 16(10):113. <https://doi.org/10.1007/s11306-020-01728-5>
- Fang LT, Zhu B, Zhao Y, Chen W, Yang Z, Kerrigan L, Langenbach K, de Mars M, Lu C, Idler K, Jacob H, Zheng Y, Ren L, Yu Y, Jaeger E, Schroth GP, Abaan OD, Talsania K, Lack J, Shen TW, Chen Z, Stanbouly S, Tran B, Shetty J, Kriga Y, Meerzaman D, Nguyen C, Petitjean V, Sultan M, Cam M, Mehta M, Hung T, Peters E, Kalamegham R, Sahraeian SME, Mohiyuddin M, Guo Y, Yao L, Song L, Lam HYK, Drabek J, Vojta P, Maestro R, Gasparotto D, Koks S, Reimann E, Scherer A, Nordlund J, Liljedahl U, Jensen RV, Pirooznia M, Li Z, Xiao C, Sherry ST, Kusko R, Moos M, Donaldson E, Tezak Z, Ning B, Tong W, Li J, Duerken-Hughes P, Catalanotti C, Maheshwari S, Shuga J, Liang WS, Keats J, Adkins J, Tassone E, Zismann V, McDaniel T, Trent J, Foux J, Butler D, Mason CE, Hong H, Shi L, Wang C, Xiao W, Somatic Mutation Working Group of Sequencing Quality Control Phase IIC (2021) Establishing community reference samples, data and call sets for benchmarking cancer mutation detection using whole-genome sequencing. *Nat Biotechnol* 39(9):1151–1160. <https://doi.org/10.1038/s41587-021-00993-6>
- Foux J, Tighe SW, Nicolet CM, Zook JM, Byrska-Bishop M, Clarke WE, Khayat MM, Mahmoud M, Laaguiby PK, Herbert ZT, Warner D, Grills GS, Jen J, Levy S, Xiang J, Alonso A, Zhao X, Zhang W, Teng F, Zhao Y, Lu H, Schroth GP, Narzisi G, Farmerie W, Sedlazeck FJ, Baldwin DA, Mason CE (2021) Performance assessment of DNA sequencing platforms in the ABRF next-generation sequencing study. *Nat Biotechnol* 39(9):1129–1140. <https://doi.org/10.1038/s41587-021-01049-5>
- Gaedigk A, Turner A, Everts RE, Scott SA, Aggarwal P, Broeckel U, McMillin GA, Melis R, Boone EC, Pratt VM, Kalman LV (2019) Characterization of reference materials for genetic testing of CYP2D6 alleles: a GeT-RM collaborative project. *J Mol Diagn* 21(6):1034–1052. <https://doi.org/10.1016/j.jmoldx.2019.06.007>
- Gainotti S, Turner C, Woods S, Kole A, McCormack P, Lochmuller H, Riess O, Straub V, Posada M, Taruscio D, Mascalzoni D (2016) Improving the informed consent process in international collaborative rare disease research: effective consent for effective research. *Eur J Hum Genet* 24(9):1248–1254. <https://doi.org/10.1038/ejhg.2016.2>
- Giraldez MD, Spengler RM, Etheridge A, Godoy PM, Barczak AJ, Srinivasan S, De Hoff PL, Tanriverdi K, Courtright A, Lu S, Khoory J, Rubio R, Baxter D, Driedonks TAP, Buermans HPJ, Nolte-t Hoen ENM, Jiang H, Wang K, Ghiran I, Wang YE, Van Keuren-Jensen K, Freedman JE, Woodruff PG, Laurent LC, Erle DJ, Galas DJ, Tewari M (2018) Comprehensive multi-center assessment of small RNA-seq methods for quantitative miRNA profiling. *Nat Biotechnol* 36(8):746–757. <https://doi.org/10.1038/nbt.4183>
- Gong B, Li D, Kusko R, Novoradovskaya N, Zhang Y, Wang S, Pabon-Pena C, Zhang Z, Lai K, Cai W, LoCoco JS, Lader E, Richmond TA, Mittal VK, Liu LC, Johann DJ Jr, Willey JC, Bushel PR, Yu Y, Xu C, Chen G, Burgess D, Cawley S, Giorda K, Haseley N, Qiu F, Wilkins K, Arib H, Attwooll C, Babson K, Bao L, Bao W, Lucas AB, Best H, Bhandari A, Bisgin H, Blackburn J, Blomquist TM, Boardman L, Burgher B, Butler DJ, Chang CJ, Chaubey A, Chen T, Chierici M, Chin CR, Close D, Conroy J, Cooley Coleman J, Craig DJ, Crawford E, Del Pozo A, Deveson IW, Duncan D, Eterovic AK, Fan X, Foux J, Furlanello C, Ghosal A, Glenn S, Guan M, Haag C, Hang X, Happe S, Hennigan B, Hipp J, Hong H, Horvath K, Hu J, Hung LY, Jarosz M, Kerkhof J, Kipp B, Kreil DP, Labaj P, Lapunzina P, Li P, Li QZ, Li W, Li Z, Liang Y, Liu S, Liu Z, Ma C, Marella N, Martin-Arenas R, Megherbi DB, Meng Q, Mieczkowski PA, Morrison T, Muzny D, Ning B, Parsons BL, Paweletz CP, Pirooznia M, Qu W, Raymond A, Rindler P, Ringle R, Sadikovic B, Scherer A, Schulze E, Sebra R, Shakhovich R, Shi Q, Shi T, Silla-Castro JC, Smith M, Lopez MS, Song P, Stetson D, Strahl M, Stuart A, Supplee J, Szankasi P, Tan H, Tang LY, Tao Y, Thakkar S, Thierry-Mieg D, Thierry-Mieg J, Thodima VJ, Thomas D, Tichy B, Tom N, Garcia EV, Verma S, Walker K, Wang C, Wang J, Wang Y, Wen Z, Wirta V, Wu L, Xiao C, Xiao W, Xu S, Yang M, Ying J, Yip SH, Zhang G, Zhang S, Zhao M, Zheng Y, Zhou X, Mason CE, Mercer T, Tong W, Shi L, Jones W, Xu J (2021) Cross-oncopanel study reveals high sensitivity and accuracy with overall analytical performance depending on genomic regions. *Genome Biol* 22(1):109. <https://doi.org/10.1186/s13059-021-02315-0>
- Guerrier G, Sicard D, Brey PT (2012) Informed consent: cultural differences. *Nature* 483(7387):36. <https://doi.org/10.1038/483036a>
- Hardwick SA, Chen WY, Wong T, Deveson IW, Blackburn J, Andersen SB, Nielsen LK, Mattick JS, Mercer TR (2016) Spliced synthetic genes as internal controls in RNA sequencing experiments. *Nat Methods* 13(9):792–798. <https://doi.org/10.1038/nmeth.3958>
- Hardwick SA, Deveson IW, Mercer TR (2017) Reference standards for next-generation sequencing. *Nat Rev Genet* 18(8):473–484. <https://doi.org/10.1038/nrg.2017.44>
- Hasin Y, Seldin M, Lusic A (2017) Multi-omics approaches to disease. *Genome Biol* 18(1):83. <https://doi.org/10.1186/s13059-017-1215-1>
- Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, Shen R, Taylor AM, Cherniack AD, Thorsson V, Akbani R, Bowlby R, Wong CK, Wiznerowicz M, Sanchez-Vega F, Robertson AG, Schneider BG, Lawrence MS, Noushmehr H, Malta TM, Cancer Genome Atlas N, Stuart JM, Benz CC, Laird PW (2018) Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* 173(2):291–304 e296. <https://doi.org/10.1016/j.cell.2018.03.022>
- Horizon Discovery (2023) Multiplex I cfDNA reference standard set. <https://horizondiscovery.com/en/reference-standards/products/multiplex-i-cfdna-reference-standard-set>
- International Organization for Standardization, ISO Guide 30:2015—reference materials—selected terms and definitions. <https://webstore.ansi.org/standards/iso/isoguide302015>
- International Organization for Standardization, ISO 9000:2015—Quality management systems. <https://www.iso.org/standard/45481.html>
- Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, Temple-Smolkin RL, Voelkerding KV, Nikiforova MN (2017) Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the

- association for molecular pathology and college of American pathologists. *J Mol Diagn* 19(3):341–365. <https://doi.org/10.1016/j.jmoldx.2017.01.011>
- Jia S, Zhang R, Lin G, Peng R, Gao P, Han Y, Fu Y, Ding J, Wu Q, Zhang K, Xie J, Li J (2018) A novel cell line generated using the CRISPR/Cas9 technology as universal quality control material for KRAS G12V mutation testing. *J Clin Lab Anal* 32(5):e22391. <https://doi.org/10.1002/jcla.22391>
- Jia P, Dong L, Yang X, Wang B, Bush SJ, Wang T, Lin J, Wang S, Zhao X, Xu T, Che Y, Dang N, Ren L, Zhang Y, Wang X, Liang F, Wang Y, Ruan J, Xia H, Zheng Y, Shi L, Lv Y, Wang J, Ye K (2023) Haplotype-resolved assemblies and variant benchmark of a Chinese Quartet. *Genome Biol* 24(1):277. <https://doi.org/10.1186/s13059-023-03116-3>
- Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, Gingeras TR, Oliver B (2011) Synthetic spike-in standards for RNA-seq experiments. *Genome Res* 21(9):1543–1551. <https://doi.org/10.1101/gr.121095.111>
- Jiang YZ, Ma D, Suo C, Shi J, Xue M, Hu X, Xiao Y, Yu KD, Liu YR, Yu Y, Zheng Y, Li X, Zhang C, Hu P, Zhang J, Hua Q, Zhang J, Hou W, Ren L, Bao D, Li B, Yang J, Yao L, Zuo WJ, Zhao S, Gong Y, Ren YX, Zhao YX, Yang YS, Niu Z, Cao ZG, Stover DG, Verschraegen C, Kaklamani V, Daemen A, Benson JR, Takabe K, Bai F, Li DQ, Wang P, Shi L, Huang W, Shao ZM (2019) Genomic and transcriptomic landscape of triple-negative breast cancers: subtypes and treatment strategies. *Cancer Cell* 35(3):428–440 e425. <https://doi.org/10.1016/j.ccell.2019.02.001>
- Jones W, Gong B, Novoradovskaya N, Li D, Kusko R, Richmond TA, Johann DJ Jr, Bisgin H, Sahraeian SME, Bushel PR, Pirooznia M, Wilkins K, Chierici M, Bao W, Basehore LS, Lucas AB, Burgess D, Butler DJ, Cawley S, Chang CJ, Chen G, Chen T, Chen YC, Craig DJ, Del Pozo A, Foox J, Francescato M, Fu Y, Furlanello C, Giorda K, Grist KP, Guan M, Hao Y, Happe S, Hariyani G, Haseley N, Jasper J, Jurman G, Kreil DP, Labaj P, Lai K, Li J, Li QZ, Li Y, Li Z, Liu Z, Lopez MS, Miclaus K, Miller R, Mittal VK, Mohiyuddin M, Pabon-Pena C, Parsons BL, Qiu F, Scherer A, Shi T, Stiegemeyer S, Suo C, Tom N, Wang D, Wen Z, Wu L, Xiao W, Xu C, Yu Y, Zhang J, Zhang Y, Zhang Z, Zheng Y, Mason CE, Willey JC, Tong W, Shi L, Xu J (2021) A verified genomic reference sample for assessing performance of cancer panels detecting small variants of low allele frequency. *Genome Biol* 22(1):111. <https://doi.org/10.1186/s13059-021-02316-z>
- JPT Peptide Technologies (2023) SpikeMix for targeted proteomics. <https://www.jpt.com/products-services/peptide-pools/spikemix-targeted-proteomics/>
- Kalman L, Johnson MA, Beck J, Berry-Kravis E, Buller A, Casey B, Feldman GL, Handsfield J, Jakupciak JP, Maragh S, Matteson K, Muralidharan K, Richie KL, Rohlf s EM, Schaefer F, Sellers T, Spector E, Richards CS (2007) Development of genomic reference materials for Huntington disease genetic testing. *Genet Med* 9(10):719–723. <https://doi.org/10.1097/gim.0b013e318156e8c1>
- Kalman L, Leonard J, Gerry N, Tarleton J, Bridges C, Gastier-Foster JM, Pyatt RE, Stonerock E, Johnson MA, Richards CS, Schrijver I, Ma T, Miller VR, Adadevoh Y, Furlong P, Beiswanger C, Toji L (2011) Quality assurance for Duchenne and Becker muscular dystrophy genetic testing: development of a genomic DNA reference material panel. *J Mol Diagn* 13(2):167–174. <https://doi.org/10.1016/j.jmoldx.2010.11.018>
- Karczewski KJ, Snyder MP (2018) Integrative omics for health and disease. *Nat Rev Genet* 19(5):299–310. <https://doi.org/10.1038/nrg.2018.4>
- Khayat MM, Sahraeian SME, Zarate S, Carroll A, Hong H, Pan B, Shi L, Gibbs RA, Mohiyuddin M, Zheng Y, Sedlazeck FJ (2021) Hidden biases in germline structural variant detection. *Genome Biol* 22(1):347. <https://doi.org/10.1186/s13059-021-02558-x>
- Kirwan JA, Gika H, Beger RD, Bearden D, Dunn WB, Goodacre R, Theodoridis G, Witting M, Yu LR, Wilson ID, Metabolomics Quality A, Quality Control C (2022) Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management. *Metabolomics* 18(9):70. <https://doi.org/10.1007/s11306-022-01926-3>
- Kocher T, Pichler P, Swart R, Mechtler K (2011) Quality control in LC-MS/MS. *Proteomics* 11(6):1026–1030. <https://doi.org/10.1002/pmic.201000578>
- Krassowski M, Das V, Sahu SK, Misra BB (2020) State of the field in multi-omics research: from computational needs to data mining and sharing. *Front Genet* 11:610798. <https://doi.org/10.3389/fgene.2020.610798>
- Krogstad DJ, Diop S, Diallo A, Mzayek F, Keating J, Koita OA, Toure YT (2010) Informed consent in international research: the rationale for different approaches. *Am J Trop Med Hyg* 83(4):743–747. <https://doi.org/10.4269/ajtmh.2010.10-0014>
- Ku X, Wang J, Li H, Meng C, Yu F, Yu W, Li Z, Zhou Z, Zhang C, Hua Y, Yan W, Jin J (2023) Proteomic portrait of human lymphoma reveals protein molecular fingerprint of disease specific subtypes and progression. *Phenomics* 3(2):148–166. <https://doi.org/10.1007/s43657-022-00075-w>
- Kumaran M, Subramanian U, Devarajan B (2019) Performance assessment of variant calling pipelines using human whole exome sequencing and simulated data. *BMC Bioinformatics* 20(1):342. <https://doi.org/10.1186/s12859-019-2928-9>
- Lazzarotto CR, Malinin NL, Li Y, Zhang R, Yang Y, Lee G, Cowley E, He Y, Lan X, Jividen K, Katta V, Kolmakova NG, Petersen CT, Qi Q, Strelcov E, Maragh S, Krenciute G, Ma J, Cheng Y, Tsai SQ (2020) CHANGE-seq reveals genetic and epigenetic effects on CRISPR-Cas9 genome-wide activity. *Nat Biotechnol* 38(11):1317–1327. <https://doi.org/10.1038/s41587-020-0555-7>
- Leger A, Amaral PP, Pandolfini L, Capitanchik C, Capraro F, Miano V, Migliori V, Toolan-Kerr P, Sideri T, Enright AJ, Tzelepis K, van Werven FJ, Luscombe NM, Barbieri I, Ule J, Fitzgerald T, Birney E, Leonardi T, Kouzarides T (2021) RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat Commun* 12(1):7198. <https://doi.org/10.1038/s41467-021-27393-3>
- Lexogen (2023) Spike-in RNA variants (SIRV). <https://www.lexogen.com/sirvs/>
- Li H, Bloom JM, Farjoun Y, Fleharty M, Gauthier L, Neale B, MacArthur D (2018) A synthetic-diploid benchmark for accurate variant-calling evaluation. *Nat Methods* 15(8):595–597. <https://doi.org/10.1038/s41592-018-0054-7>
- Li J, Zhang L, Li L, Li X, Zhang X, Zhai S, Gao H, Li Y, Wu G, Wu Y (2020) Development of genomic DNA certified reference materials for genetically modified rice Kefeng 6. *ACS Omega* 5(34):21602–21609. <https://doi.org/10.1021/acsomega.0c02274>
- Lin G, Zhang K, Han Y, Peng R, Li J (2022) Preparation of multiplexed control materials for cancer mutation analysis by genome editing in GM12878 cells. *J Clin Lab Anal* 36(1):e24139. <https://doi.org/10.1002/jcla.24139>
- Lindon JC, Nicholson JK, Holmes E, Keun HC, Craig A, Pearce JT, Bruce SJ, Hardy N, Sansone SA, Antti H, Jonsson P, Daykin C, Navarange M, Beger RD, Verheij ER, Amberg A, Baunsgaard D, Cantor GH, Lehman-McKeeman L, Earll M, Wold S, Johansson E, Haselden JN, Kramer K, Thomas C, Lindberg J, Schuppe-Koistinen I, Wilson ID, Reily MD, Robertson DG, Senn H, Krotzky A, Kochhar S, Powell J, van der Ouderaa F, Plumb R, Schaefer H, Spraul M, Standard Metabolic Reporting Structures Working G (2005) Summary recommendations for standardization and reporting of metabolic analyses. *Nat Biotechnol* 23(7):833–838. <https://doi.org/10.1038/nbt0705-833>
- Lippa KA, Aristizabal-Henao JJ, Beger RD, Bowden JA, Broeckling C, Beecher C, Clay Davis W, Dunn WB, Flores R, Goodacre R,

- Gouveia GJ, Harms AC, Hartung T, Jones CM, Lewis MR, Ntai I, Percy AJ, Raftery D, Schock TB, Sun J, Theodoridis G, Tayyari F, Torta F, Ulmer CZ, Wilson I, Ubhi BK (2022) Reference materials for MS-based untargeted metabolomics and lipidomics: a review by the metabolomics quality assurance and quality control consortium (mQACC). *Metabolomics* 18(4):24. <https://doi.org/10.1007/s11306-021-01848-6>
- Mangiante L, Alcalá N, Sexton-Oates A, Di Genova A, Gonzalez-Perez A, Khandekar A, Bergstrom EN, Kim J, Liu X, Blazquez-Encinas R, Giacobi C, Le Stang N, Boyault S, Cuenin C, Tabone-Eglinger S, Damiola F, Voegelé C, Ardin M, Michallet MC, Soudade L, Delhomme TM, Poret A, Brevet M, Copin MC, Giusiano-Courcambeck S, Damotte D, Girard C, Hofman V, Hofman P, Mouroux J, Cohen C, Lacomme S, Mazieres J, de Montpreville VT, Perrin C, Planchard G, Rousseau N, Rouquette I, Sagan C, Scherpereel A, Thivolet F, Vignaud JM, Jean D, Ilg AGS, Olaso R, Meyer V, Boland-Auge A, Deleuze JF, Altmüller J, Nuernberg P, Ibanez-Costa A, Castano JP, Lantuejoul S, Ghantous A, Maussion C, Courtiol P, Hernandez-Vargas H, Caux C, Girard N, Lopez-Bigas N, Alexandrov LB, Galateau-Salle F, Foll M, Fernandez-Cuesta L (2023) Multiomic analysis of malignant pleural mesothelioma identifies molecular axes and specialized tumor profiles driving intertumor heterogeneity. *Nat Genet* 55(4):607–618. <https://doi.org/10.1038/s41588-023-01321-1>
- MAQC Consortium (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 24(9):1151–1161. <https://doi.org/10.1038/nbt1239>
- MAQC Consortium (2014) A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat Biotechnol* 32(9):903–914. <https://doi.org/10.1038/nbt.2957>
- MAQC Consortium (2021) Sequencing quality control 2 nature collection. <https://www.nature.com/collections/fjhdjcdcfj>
- Martinez-Ruiz C, Black JRM, Puttick C, Hill MS, Demeulemeester J, Larose Cadieux E, Thol K, Jones TP, Veeriah S, Naceur-Lombardelli C, Toncheva A, Prymas P, Rowan A, Ward S, Cubitt L, Athanasopoulou F, Pich O, Karasaki T, Moore DA, Salgado R, Colliver E, Castignani C, Dietzen M, Huebner A, Al Bakir M, Tanic M, Watkins TBK, Lim EL, Al-Rashed AM, Lang D, Clements J, Cook DE, Rosenthal R, Wilson GA, Frankell AM, de Carne TS, East P, Kanu N, Litchfield K, Birkbak NJ, Hackshaw A, Beck S, Van Loo P, Jamal-Hanjani M, Consortium TR, Swanton C, McGranahan N (2023) Genomic-transcriptomic evolution in lung cancer and metastasis. *Nature* 616(7957):543–552. <https://doi.org/10.1038/s41586-023-05706-4>
- McGaw EA, Phinney KW, Lowenthal MS (2010) Comparison of orthogonal liquid and gas chromatography–mass spectrometry platforms for the determination of amino acid concentrations in human plasma. *J Chromatogr A* 1217(37):5822–5831. <https://doi.org/10.1016/j.chroma.2010.07.025>
- Medical Device Innovation Consortium (2019) MDIC SRS report: somatic variant reference samples for NGS. <https://mdic.org/wp-content/uploads/2019/03/MDIC-SRS-Landscape-Analysis-Report-20190306.pdf>
- Merino DM, McShane LM, Fabrizio D, Funari V, Chen SJ, White JR, Wenz P, Baden J, Barrett JC, Chaudhary R, Chen L, Chen WS, Cheng JH, Cyanam D, Dickey JS, Gupta V, Hellmann M, Helman E, Li Y, Maas J, Papin A, Patidar R, Quinn KJ, Rizvi N, Tae H, Ward C, Xie M, Zehir A, Zhao C, Dietel M, Stenzinger A, Stewart M, Allen J, Consortium TMBH (2020) Establishing guidelines to harmonize tumor mutational burden (TMB): in silico assessment of variation in TMB quantification across diagnostic platforms: phase I of the Friends of Cancer Research TMB Harmonization Project. *J Immunother Cancer*. <https://doi.org/10.1136/jitc-2019-000147>
- Met Qual Program Coordinators (2023) The NIST metabolomics quality assurance and quality control materials (MetQual) program. <https://www.nist.gov/programs-projects/metabolomics-quality-assurance-and-quality-control-materials-metqual-program>
- Misra BB, Olivier M (2020) High resolution GC-orbitrap-MS metabolomics using both electron ionization and chemical ionization for analysis of human plasma. *J Proteome Res* 19(7):2717–2731. <https://doi.org/10.1021/acs.jproteome.9b00774>
- Morgenstern D, Barzilay R, Levin Y (2021) RawBeans: a simple, vendor-independent, raw-data quality-control tool. *J Proteome Res* 20(4):2098–2104. <https://doi.org/10.1021/acs.jproteome.0c00956>
- Munro SA, Lund SP, Pine PS, Binder H, Clevert DA, Conesa A, Dopazo J, Fasold M, Hochreiter S, Hong H, Jafari N, Kreil DP, Labaj PP, Li S, Liao Y, Lin SM, Meehan J, Mason CE, Santoyo-Lopez J, Setterquist RA, Shi L, Shi W, Smyth GK, Stralis-Pavese N, Su Z, Tong W, Wang C, Wang J, Xu J, Ye Z, Yang Y, Yu Y, Salit M (2014) Assessing technical performance in differential gene expression experiments with external spike-in RNA control ratio mixtures. *Nat Commun* 5:5125. <https://doi.org/10.1038/ncomms6125>
- Nakayasu ES, Gritsenko M, Piehowski PD, Gao Y, Orton DJ, Schepmoes AA, Fillmore TL, Frohnert BI, Rewers M, Krischer JP, Ansong C, Suchy-Dacey AM, Evans-Molina C, Qian WJ, Webb-Robertson BM, Metz TO (2021) Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. *Nat Protoc* 16(8):3737–3760. <https://doi.org/10.1038/s41596-021-00566-6>
- National Institute of Standards and Technology (2022) The 2022 NIST-hosted workshop on standards for microbiome and multi-omics measurements. <https://www.nist.gov/news-events/events/2022/08/2022-nist-hosted-workshop-standards-microbiome-and-multiomics>
- National Institute of Standards and Technology (2023a) NIST SRM definitions. <https://www.nist.gov/srm/srm-definitions>
- National Institute of Standards and Technology (2023b) Genome in a bottle. <https://www.nist.gov/programs-projects/genome-bottle>
- Nurk S, Koren S, Rhie A, Rautiainen M, Bizikadze AV, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, Aganezov S, Hoyt SJ, Diekhans M, Logsdon GA, Alonge M, Antonarakis SE, Borchers M, Bouffard GG, Brooks SY, Caldas GV, Chen NC, Cheng H, Chin CS, Chow W, de Lima LG, Dishuck PC, Durbin R, Dvorkina T, Fiddes IT, Formenti G, Fulton RS, Fungtammasan A, Garrison E, Grady PGS, Graves-Lindsay TA, Hall IM, Hansen NF, Hartley GA, Haukness M, Howe K, Hunkapiller MW, Jain C, Jain M, Jarvis ED, Kerpedjiev P, Kirsche M, Kolmogorov M, Korlach J, Kremitzki M, Li H, Maduro VV, Marschall T, McCartney AM, McDaniel J, Miller DE, Mullikin JC, Myers EW, Olson ND, Paten B, Peluso P, Pevzner PA, Porubsky D, Potapova T, Rogav EI, Rosenfeld JA, Salzberg SL, Schneider VA, Sedlazeck FJ, Shafin K, Shew CJ, Shumate A, Sims Y, Smit AFA, Soto DC, Sovic I, Storer JM, Streets A, Sullivan BA, Thibaud-Nissen F, Torrance J, Wagner J, Walenz BP, Wenger A, Wood JMD, Xiao C, Yan SM, Young AC, Zarate S, Surti U, McCoy RC, Dennis MY, Alexandrov IA, Gerton J, O’Neill RJ, Timp W, Zook JM, Schatz MC, Eichler EE, Miga KH, Phillippy AM (2022) The complete sequence of a human genome. *Science* 376(6588):44–53. <https://doi.org/10.1126/science.abj6987>
- Olink (2023) Olink data generation and QC. <https://olink.com/our-platform/our-pea-technology/data-generation-and-qc/>
- Omenn GS (2021) Reflections on the HUPO human proteome project, the flagship project of the human proteome organization, at 10 years. *Mol Cell Proteomics* 20:100062. <https://doi.org/10.1016/j.mcpro.2021.100062>

- Pan B, Ren L, Onuchic V, Guan M, Kusko R, Bruinsma S, Trigg L, Scherer A, Ning B, Zhang C, Glidewell-Kenney C, Xiao C, Donaldson E, Sedlazeck FJ, Schroth G, Yavas G, Grunewald H, Chen H, Meinholz H, Meehan J, Wang J, Yang J, Foox J, Shang J, Miclaus K, Dong L, Shi L, Mohiyuddin M, Pirooznia M, Gong P, Golshani R, Wolfinger R, Lababidi S, Sahraeian SME, Sherry S, Han T, Chen T, Shi T, Hou W, Ge W, Zou W, Guo W, Bao W, Xiao W, Fan X, Gondo Y, Yu Y, Zhao Y, Su Z, Liu Z, Tong W, Xiao W, Zook JM, Zheng Y, Hong H (2022) Assessing reproducibility of inherited variants detected with short-read whole genome sequencing. *Genome Biol* 23(1):2. <https://doi.org/10.1186/s13059-021-02569-8>
- Patel RK, Jain M (2012) NGS QC toolkit: a toolkit for quality control of next generation sequencing data. *PLoS ONE* 7(2):e30619. <https://doi.org/10.1371/journal.pone.0030619>
- Paulovich AG, Billheimer D, Ham AJ, Vega-Montoto L, Rudnick PA, Tabb DL, Wang P, Blackman RK, Bunk DM, Cardasis HL, Clauser KR, Kinsinger CR, Schilling B, Tegeler TJ, Variyath AM, Wang M, Whiteaker JR, Zimmerman LJ, Fenyo D, Carr SA, Fisher SJ, Gibson BW, Mesri M, Neubert TA, Regnier FE, Rodriguez H, Spiegelman C, Stein SE, Tempst P, Liebler DC (2010) Interlaboratory study characterizing a yeast performance standard for benchmarking LC–MS platform performance. *Mol Cell Proteomics* 9(2):242–254. <https://doi.org/10.1074/mcp.M900222-MCP200>
- Petrea A, von Toerne C, Behler J, Huth C, Thorand B, Hilgendorff A, Hauck SM (2021) Multiplatform approach for plasma proteomics: complementarity of olink proximity extension assay technology to mass spectrometry-based protein profiling. *J Proteome Res* 20(1):751–762. <https://doi.org/10.1021/acs.jproteome.0c00641>
- Pfeifer JD, Loberg R, Lofton-Day C, Zehnbauser BA (2022) Reference samples to compare next-generation sequencing test performance for oncology therapeutics and diagnostics. *Am J Clin Pathol* 157(4):628–638. <https://doi.org/10.1093/ajcp/aqab164>
- Phinney KW, Ballihaut G, Bedner M, Benford BS, Camara JE, Christopher SJ, Davis WC, Dodder NG, Epe G, Lang BE, Long SE, Lowenthal MS, McGaw EA, Murphy KE, Nelson BC, Prendergast JL, Reiner JL, Rimmer CA, Sander LC, Schantz MM, Sharpless KE, Sniegowski LT, Tai SS, Thomas JB, Vetter TW, Welch MJ, Wise SA, Wood LJ, Guthrie WF, Hagwood CR, Leigh SD, Yen JH, Zhang NF, Chaudhary-Webb M, Chen H, Fazili Z, LaVoie DJ, McCoy LF, Momin SS, Paladugula N, Pendergrast EC, Pfeiffer CM, Powers CD, Rabinowitz D, Rybak ME, Schleicher RL, Toombs BM, Xu M, Zhang M, Castle AL (2013) Development of a standard reference material for metabolomics research. *Anal Chem* 85(24):11732–11738. <https://doi.org/10.1021/ac402689t>
- Pratt VM, Caggana M, Bridges C, Buller AM, DiAntonio L, Highsmith WE, Holtegaard LM, Muralidharan K, Rohlf EM, Tarleton J, Toji L, Barker SD, Kalman LV (2009) Development of genomic reference materials for cystic fibrosis genetic testing. *J Mol Diagn* 11(3):186–193. <https://doi.org/10.2353/jmoldx.2009.080149>
- Pratt VM, Everts RE, Aggarwal P, Beyer BN, Broeckel U, Epstein-Baak R, Hujsak P, Kornreich R, Liao J, Lorier R, Scott SA, Smith CH, Toji LH, Turner A, Kalman LV (2016) Characterization of 137 genomic DNA reference materials for 28 pharmacogenetic genes: a GeT-RM collaborative project. *J Mol Diagn* 18(1):109–123. <https://doi.org/10.1016/j.jmoldx.2015.08.005>
- Price ND, Magis AT, Earls JC, Glusman G, Levy R, Lausted C, McDonald DT, Kusebauch U, Moss CL, Zhou Y, Qin S, Moritz RL, Brogaard K, Omenn GS, Lovejoy JC, Hood L (2017) A wellness study of 108 individuals using personal, dense, dynamic data clouds. *Nat Biotechnol* 35(8):747–756. <https://doi.org/10.1038/nbt.3870>
- Reis ALM, Deveson IW, Wong T, Madala BS, Barker C, Blackburn J, Marcellin E, Mercer TR (2020) A universal and independent synthetic DNA ladder for the quantitative measurement of genomic features. *Nat Commun* 11(1):3609. <https://doi.org/10.1038/s41467-020-17445-5>
- Ren L, Duan X, Dong L, Zhang R, Yang J, Gao Y, Peng R, Hou W, Liu Y, Li J, Yu Y, Zhang N, Shang J, Liang F, Wang D, The Quartet Project Team, Scherer A, Nordlund J, Xiao W, Xu J, Tong W, Hu X, Li J, Jin L, Shi L, Hong H, Wang J, Fan S, Fang X, Zheng Y (2023) Quartet DNA reference materials and datasets for comprehensively evaluating germline variants calling performance. *Genome Biol* 24:270. <https://doi.org/10.1186/s13059-023-03109-2>
- Robasky K, Lewis NE, Church GM (2014) The role of replicates for error mitigation in next-generation sequencing. *Nat Rev Genet* 15(1):56–62. <https://doi.org/10.1038/nrg3655>
- Rudnick PA, Markey SP, Roth J, Mirokhin Y, Yan X, Tchekhovskoi DV, Edwards NJ, Thangudu RR, Ketchum KA, Kinsinger CR, Mesri M, Rodriguez H, Stein SE (2016) A description of the clinical proteomic tumor analysis consortium (CPTAC) common data analysis pipeline. *J Proteome Res* 15(3):1023–1032. <https://doi.org/10.1021/acs.jproteome.5b01091>
- Sahraeian SME, Fang LT, Karagiannis K, Moos M, Smith S, Santana-Quintero L, Xiao C, Colgan M, Hong H, Mohiyuddin M, Xiao W (2022) Achieving robust somatic mutation detection with deep learning models derived from reference data sets of a cancer sample. *Genome Biol* 23(1):12. <https://doi.org/10.1186/s13059-021-02592-9>
- Sammur SJ, Crispin-Ortuzar M, Chin SF, Provenzano E, Bardwell HA, Ma W, Cope W, Dariush A, Dawson SJ, Abraham JE, Dunn J, Hiller L, Thomas J, Cameron DA, Bartlett JMS, Hayward L, Pharoah PD, Markowitz F, Rueda OM, Earl HM, Caldas C (2022) Multi-omic machine learning predictor of breast cancer therapy response. *Nature* 601(7894):623–629. <https://doi.org/10.1038/s41586-021-04278-5>
- Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, Barron DA, Zehir A, Jordan EJ, Omuro A, Kaley TJ, Kendall SM, Motzer RJ, Hakimi AA, Voss MH, Russo P, Rosenberg J, Iyer G, Bochner BH, Bajorin DF, Al-Ahmadie HA, Chaft JE, Rudin CM, Riely GJ, Baxi S, Ho AL, Wong RJ, Pfister DG, Wolchok JD, Barker CA, Gutin PH, Brennan CW, Tabar V, Mellinghoff IK, DeAngelis LM, Ariyan CE, Lee N, Tap WD, Gounder MM, D'Angelo SP, Saltz L, Stadler ZK, Scher HI, Baselga J, Razavi P, Klebanoff CA, Yaeger R, Segal NH, Ku GY, DeMatteo RP, Ladanyi M, Rizvi NA, Berger MF, Riaz N, Solit DB, Chan TA, Morris LGT (2019) Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* 51(2):202–206. <https://doi.org/10.1038/s41588-018-0312-8>
- Schantz MM, Epe G, Focant JF, Hamilton C, Heckert NA, Heltsley RM, Hoover D, Keller JM, Leigh SD, Patterson DG Jr, Pintar AL, Sharpless KE, Sjodin A, Turner WE, Vander Pol SS, Wise SA (2013) Milk and serum standard reference materials for monitoring organic contaminants in human samples. *Anal Bioanal Chem* 405(4):1203–1211. <https://doi.org/10.1007/s00216-012-6524-3>
- Schussler-Fiorenza Rose SM, Contrepolis K, Moneghetti KJ, Zhou W, Mishra T, Mataraso S, Dagan-Rosenfeld O, Ganz AB, Dunn J, Hornburg D, Rego S, Perelman D, Ahadi S, Sailani MR, Zhou Y, Leopold SR, Chen J, Ashland M, Christle JW, Avina M, Limcaoco P, Ruiz C, Tan M, Butte AJ, Weinstock GM, Slavich GM, Sodergren E, McLaughlin TL, Haddad F, Snyder MP (2019) A longitudinal big data approach for precision health. *Nat Med* 25(5):792–804. <https://doi.org/10.1038/s41591-019-0414-6>

- Sempos CT, Lindhout E, Heureux N, Hars M, Parkington DA, Dennison E, Durazo-Arvizu R, Jones KS, Wise SA (2022) Towards harmonization of directly measured free 25-hydroxyvitamin D using an enzyme-linked immunosorbent assay. *Anal Bioanal Chem* 414(27):7793–7803. <https://doi.org/10.1007/s00216-022-04313-y>
- Seracare (2023a) Seraseq gDNA TMB reference panel mix. <https://www.seracare.com/Seraseq-gDNA-TMB-Reference-Panel-Mix-0710-2463/>
- Seracare (2023b) Seraseq ctDNA reference materials. <https://www.seracare.com/Seraseq-ctDNA-Complete-Reference-Material-AF05-0710-0672/>
- Shi L, Kusko R, Wolfinger RD, Haibe-Kains B, Fischer M, Sansone SA, Mason CE, Furlanello C, Jones WD, Ning B, Tong W (2017) The international MAQC society launches to enhance reproducibility of high-throughput technologies. *Nat Biotechnol* 35(12):1127–1128. <https://doi.org/10.1038/nbt.4029>
- Simon-Manso Y, Lowenthal MS, Kilpatrick LE, Sampson ML, Telu KH, Rudnick PA, Mallard WG, Bearden DW, Schock TB, Tchekhovskoi DV, Blonder N, Yan X, Liang Y, Zheng Y, Wallace WE, Neta P, Phinney KW, Remaley AT, Stein SE (2013) Metabolite profiling of a NIST Standard Reference Material for human plasma (SRM 1950): GC–MS, LC–MS, NMR, and clinical laboratory analyses, libraries, and web-based resources. *Anal Chem* 85(24):11725–11731. <https://doi.org/10.1021/ac402503m>
- Siskos AP, Jain P, Romisch-Margl W, Bennett M, Achaintre D, Asad Y, Marney L, Richardson L, Koulman A, Griffin JL, Raynaud F, Scalbert A, Adamski J, Prehn C, Keun HC (2017) Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. *Anal Chem* 89(1):656–665. <https://doi.org/10.1021/acs.analchem.6b02930>
- Soneson C, Yao Y, Bratus-Neuenschwander A, Patrignani A, Robinson MD, Hussain S (2019) A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes. *Nat Commun* 10(1):3359. <https://doi.org/10.1038/s41467-019-11272-z>
- Sonia Tarazona AA-LAC (2021) Undisclosed, unmet and neglected challenges in multi-omics studies. *Nat Comput Sci* 1:395–402. <https://doi.org/10.1038/s43588-021-00086-z>
- Stenzinger A, Allen JD, Maas J, Stewart MD, Merino DM, Wempe MM, Dietel M (2019) Tumor mutational burden standardization initiatives: recommendations for consistent tumor mutational burden assessment in clinical samples to guide immunotherapy treatment decisions. *Genes Chromosomes Cancer* 58(8):578–588. <https://doi.org/10.1002/gcc.22733>
- Stephens PJ, McBride DJ, Lin ML, Varela I, Pleasance ED, Simpson JT, Stebbings LA, Leroy C, Edkins S, Mudie LJ, Greenman CD, Jia M, Latimer C, Teague JW, Lau KW, Burton J, Quail MA, Swerdlow H, Churchev C, Natrajan R, Sieuwerts AM, Martens JW, Silver DP, Langerod A, Russnes HE, Foekens JA, Reis-Filho JS, Van't Veer L, Richardson AL, Borresen-Dale AL, Campbell PJ, Futreal PA, Stratton MR (2009) Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* 462(7276):1005–1010. <https://doi.org/10.1038/nature08645>
- Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, Surendran P, Mahajan A, Robins C, Vasquez-Grinnell SG, Hou L, Kvikstad EM, Burren OS, Davitte J, Ferber KL, Gillies CE, Hedman AK, Hu S, Lin T, Mikkilineni R, Pendergrass RK, Pickering C, Prins B, Baird D, Chen CY, Ward LD, Deaton AM, Welsh S, Willis CM, Lehner N, Arnold M, Worheide MA, Suhre K, Kastenmuller G, Sethi A, Cule M, Raj A, Alnylam Human G, AstraZeneca Genomics I, Biogen Biobank T, Bristol Myers S, Genentech Human G, GlaxoSmithKline Genomic S, Pfizer Integrative B, Population Analytics of Janssen Data S, Regeneron Genetics C, Burkitt-Gray L, Melamed E, Black MH, Fauman EB, Howson JMM, Kang HM, McCarthy MI, Nioi P, Petrovski S, Scott RA, Smith EN, Szalma S, Waterworth DM, Mitnau LJ, Szustakowski JD, Gibson BW, Miller MR, Whelan CD (2023) Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* 622(7982):329–338. <https://doi.org/10.1038/s41586-023-06592-6>
- Suzuki T, Tsukumo Y, Furihata C, Naito M, Kohara A (2020) Preparation of the standard cell lines for reference mutations in cancer gene-panels by genome editing in HEK 293 T/17 cells. *Genes Environ* 42:8. <https://doi.org/10.1186/s41021-020-0147-2>
- Tabb DL, Vega-Montoto L, Rudnick PA, Variyath AM, Ham AJ, Bunk DM, Kilpatrick LE, Billheimer DD, Blackman RK, Cardasis HL, Carr SA, Clauser KR, Jaffe JD, Kowalski KA, Neubert TA, Regnier FE, Schilling B, Tegeler TJ, Wang M, Wang P, Whiteaker JR, Zimmerman LJ, Fisher SJ, Gibson BW, Kinsinger CR, Mesri M, Rodriguez H, Stein SE, Tempst P, Paulovich AG, Liebler DC, Spiegelman C (2010) Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* 9(2):761–776. <https://doi.org/10.1021/pr9006365>
- Tabb DL, Wang X, Carr SA, Clauser KR, Mertins P, Chambers MC, Holman JD, Wang J, Zhang B, Zimmerman LJ, Chen X, Gunawardena HP, Davies SR, Ellis MJ, Li S, Townsend RR, Boja ES, Ketchum KA, Kinsinger CR, Mesri M, Rodriguez H, Liu T, Kim S, McDermott JE, Payne SH, Petyuk VA, Rodland KD, Smith RD, Yang F, Chan DW, Zhang B, Zhang H, Zhang Z, Zhou JY, Liebler DC (2016) Reproducibility of differential proteomic technologies in CPTAC fractionated xenografts. *J Proteome Res* 15(3):691–706. <https://doi.org/10.1021/acs.jproteome.5b00859>
- Talsania K, Shen TW, Chen X, Jaeger E, Li Z, Chen Z, Chen W, Tran B, Kusko R, Wang L, Pang AWC, Yang Z, Choudhari S, Colgan M, Fang LT, Carroll A, Shetty J, Kriga Y, German O, Smirnova T, Liu T, Li J, Kellman B, Hong K, Hastie AR, Natarajan A, Moshrefi A, Granat A, Truong T, Bombardi R, Mankinen V, Meerzaman D, Mason CE, Collins J, Stahlberg E, Xiao C, Wang C, Xiao W, Zhao Y (2022) Structural variant analysis of a cancer reference cell line sample using multiple sequencing technologies. *Genome Biol* 23(1):255. <https://doi.org/10.1186/s13059-022-02816-6>
- Thermo Scientific (2020) AcroMatrix™ oncology hotspot control. <https://assets.thermofisher.com/TFS-Assets/CDD/manuals/MAN0010820-AMX-Oncology-Hotspot-Ctrl-EN.pdf>
- Thompson JW, Adams KJ, Adamski J, Asad Y, Borts D, Bowden JA, Byram G, Dang V, Dunn WB, Fernandez F, Fiehn O, Gaul DA, Huher AF, Kalli A, Koal T, Koeniger S, Mandal R, Meier F, Naser FJ, O'Neil D, Pal A, Patti GJ, Pham-Tuan H, Prehn C, Raynaud FI, Shen T, Southam AD, St John-Williams L, Sulek K, Vasilopoulou CG, Viant M, Winder CL, Wishart D, Zhang L, Zheng J, Moseley MA (2019) International ring trial of a high resolution targeted metabolomics and lipidomics platform for serum and plasma analysis. *Anal Chem* 91(22):14407–14416. <https://doi.org/10.1021/acs.analchem.9b02908>
- Tian S, Zhan D, Yu Y, Liu M, Wang Y, Song L, Qin Z, Li X, Liu Y, Li Y, Ji S, Li Y, Li L, Wang S, Analysis PM, Control Q, Zheng Y, He F, Qin J, Ding C (2023) Quartet protein reference materials and datasets for multi-platform assessment of label-free proteomics. *Genome Biol* 24:202. <https://doi.org/10.1186/s13059-023-03048-y>
- Turck CW, Mak TD, Goudarzi M, Salek RM, Cheema AK (2020) The ABRF metabolomics research group 2016 exploratory study: investigation of data analysis methods for untargeted metabolomics. *Metabolites*. <https://doi.org/10.3390/metabo10040128>
- van Belle G, Mentzelopoulos SD, Aufderheide T, May S, Nichol G (2015) International variation in policies and practices related to informed consent in acute cardiovascular research: results from a 44 country survey. *Resuscitation* 91:76–83. <https://doi.org/10.1016/j.resuscitation.2014.11.029>

- Vega DM, Yee LM, McShane LM, Williams PM, Chen L, Vilimas T, Fabrizio D, Funari V, Newberg J, Bruce LK, Chen SJ, Baden J, Carl Barrett J, Beer P, Butler M, Cheng JH, Conroy J, Cyanam D, Eyring K, Garcia E, Green G, Gregersen VR, Hellmann MD, Keefer LA, Lasiter L, Lazar AJ, Li MC, MacConaill LE, Meier K, Mellert H, Pabla S, Pallavajjalla A, Pestano G, Salgado R, Samara R, Sokol ES, Stafford P, Budczies J, Stenzinger A, Tom W, Valkenburg KC, Wang XZ, Weigman V, Xie M, Xie Q, Zehir A, Zhao C, Zhao Y, Stewart MD, Allen J, Consortium TMB (2021) Aligning tumor mutational burden (TMB) quantification across diagnostic platforms: phase II of the Friends of Cancer Research TMB Harmonization Project. *Ann Oncol* 32(12):1626–1636. <https://doi.org/10.1016/j.annonc.2021.09.016>
- Veltman JA, Brunner HG (2012) De novo mutations in human genetic disease. *Nat Rev Genet* 13(8):565–575. <https://doi.org/10.1038/nrg3241>
- Wagner J, Olson ND, Harris L, Khan Z, Farek J, Mahmoud M, Stankovic A, Kovacevic V, Yoo B, Miller N, Rosenfeld JA, Ni B, Zarate S, Kirsche M, Aganezov S, Schatz MC, Narzisi G, Byrskabishop M, Clarke W, Evans US, Markello C, Shafin K, Zhou X, Sidow A, Bansal V, Ebert P, Marschall T, Lansdorp P, Hanlon V, Mattsson CA, Barrio AM, Fiddes IT, Xiao C, Functammasan A, Chin CS, Wenger AM, Rowell WJ, Sedlazeck FJ, Carroll A, Salit M, Zook JM (2022) Benchmarking challenging small variants with linked and long reads. *Cell Genom*. <https://doi.org/10.1016/j.xgen.2022.100128>
- Wang X, Lu M, Qian J, Yang Y, Li S, Lu D, Yu S, Meng W, Ye W, Jin L (2009) Rationales, design and recruitment of the Taizhou Longitudinal Study. *BMC Public Health* 9:223. <https://doi.org/10.1186/1471-2458-9-223>
- Wang X, Chambers MC, Vega-Montoto LJ, Bunk DM, Stein SE, Tabb DL (2014) QC metrics from CPTAC raw LC–MS/MS data interpreted through multivariate statistics. *Anal Chem* 86(5):2497–2509. <https://doi.org/10.1021/ac4034455>
- Wang D, Zhang Y, Li R, Li J, Zhang R (2023) Consistency and reproducibility of large panel next-generation sequencing: Multi-laboratory assessment of somatic mutation detection on reference materials with mismatch repair and proofreading deficiency. *J Adv Res* 44:161–172. <https://doi.org/10.1016/j.jare.2022.03.016>
- Wenger AM, Peluso P, Rowell WJ, Chang PC, Hall RJ, Concepcion GT, Ebler J, Functammasan A, Kolesnikov A, Olson ND, Topfer A, Alonge M, Mahmoud M, Qian Y, Chin CS, Phillippy AM, Schatz MC, Myers G, DePristo MA, Ruan J, Marschall T, Sedlazeck FJ, Zook JM, Li H, Koren S, Carroll A, Rank DR, Hunkapiller MW (2019) Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat Biotechnol* 37(10):1155–1162. <https://doi.org/10.1038/s41587-019-0217-9>
- Wik L, Nordberg N, Broberg J, Bjorkesten J, Assarsson E, Henriksson S, Grundberg I, Pettersson E, Westerberg C, Liljeroth E, Falck A, Lundberg M (2021) Proximity extension assay in combination with next-generation sequencing for high-throughput proteome-wide analysis. *Mol Cell Proteomics* 20:100168. <https://doi.org/10.1016/j.mcpro.2021.100168>
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJ, Groth P, Goble C, Grethe JS, Heringa J, t Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater J, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3:160018. <https://doi.org/10.1038/sdata.2016.18>
- Xiao W, Ren L, Chen Z, Fang LT, Zhao Y, Lack J, Guan M, Zhu B, Jaeger E, Kerrigan L, Blomquist TM, Hung T, Sultan M, Idler K, Lu C, Scherer A, Kusko R, Moos M, Xiao C, Sherry ST, Abaan OD, Chen W, Chen X, Nordlund J, Liljedahl U, Maestro R, Polano M, Drabek J, Vojta P, Koks S, Reimann E, Madala BS, Mercer T, Miller C, Jacob H, Truong T, Moshrefi A, Natarajan A, Granat A, Schroth GP, Kalamegham R, Peters E, Petitjean V, Walton A, Shen TW, Talsania K, Vera CJ, Langenbach K, de Mars M, Hipp JA, Willey JC, Wang J, Shetty J, Kriga Y, Raziuddin A, Tran B, Zheng Y, Yu Y, Cam M, Jailwala P, Nguyen C, Meerzaman D, Chen Q, Yan C, Ernest B, Mehra U, Jensen RV, Jones W, Li JL, Papas BN, Pirooznia M, Chen YC, Seifuddin F, Li Z, Liu X, Resch W, Wang J, Wu L, Yavas G, Miles C, Ning B, Tong W, Mason CE, Donaldson E, Lababidi S, Staudt LM, Tezak Z, Hong H, Wang C, Shi L (2021) Toward best practice in cancer mutation detection with whole-genome and whole-exome sequencing. *Nat Biotechnol* 39(9):1141–1150. <https://doi.org/10.1038/s41587-021-00994-5>
- Yang J, Liu Y, Shang J, Chen Q, Chen Q, Ren L, Zhang N, Yu Y, Li Z, Song Y, Scherer A, Niehues A, Tong W, Hong H, Shi L, Xiao W, Zheng Y (2023) The Quartet Data Portal: integration of community-wide resources for multiomics quality control. *Genome Biol* 24:245. <https://doi.org/10.1186/s13059-023-03091-9>
- Yarchoan M, Hopkins A, Jaffee EM (2017) Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med* 377(25):2500–2501. <https://doi.org/10.1056/NEJMc1713444>
- Yu Y, Hou W, Wang H, Dong L, Liu Y, Sun S, Yang J, Cao Z, Zhang P, Zi Y, Li Z, Liu R, Gao J, Chen Q, Zhang N, Li J, Ren L, Jiang H, Shang J, Zhu S, Wang X, Qing T, Bao D, Li B, Li B, Suo C, Pi Y, The Quartet Project Team, Wang X, Dai F, Scherer A, Mattila P, Han J, Zhang L, Jiang H, Thierry-Mieg D, Thierry-Mieg J, Xiao W, Hong H, Tong W, Wang J, Li J, Fang X, Jin L, Shi L, Xu J, Qian F, Zhang R, Zheng Y (2023) Quartet RNA reference materials and ratio-based reference datasets for reliable transcriptomic profiling. *Nat Biotechnol*. <https://doi.org/10.1038/s41587-023-01867-9>
- Zecha J, Gabriel W, Spallek R, Chang YC, Mergner J, Wilhelm M, Bassermann F, Kuster B (2022) Linking post-translational modifications and protein turnover by site-resolved protein turnover profiling. *Nat Commun* 13(1):165. <https://doi.org/10.1038/s41467-021-27639-0>
- Zehnbauber B, Lofton-Day C, Pfeifer J, Shaughnessy E, Goh L (2017) Diagnostic quality assurance pilot: a model to demonstrate comparative laboratory test performance with an oncology companion diagnostic assay. *J Mol Diagn* 19(1):1–3. <https://doi.org/10.1016/j.jmoldx.2016.10.001>
- Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH (2015) Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther Nucl Acids* 4:e264. <https://doi.org/10.1038/mtna.2015.37>
- Zhang R, Peng R, Li Z, Gao P, Jia S, Yang X, Ding J, Han Y, Xie J, Li J (2017) Synthetic circulating cell-free DNA as quality control materials for somatic mutation detection in liquid biopsy for cancer. *Clin Chem* 63(9):1465–1475. <https://doi.org/10.1373/clinchem.2017.272559>
- Zhang K, Lin G, Han D, Han Y, Wang J, Shen Y, Li J (2020) An initial survey of the performances of exome variant analysis and clinical reporting among diagnostic laboratories in China. *Front Genet* 11:582637. <https://doi.org/10.3389/fgene.2020.582637>
- Zhang W, Wang R, Fang H, Ma X, Li D, Liu T, Chen Z, Wang K, Hao S, Yu Z, Chang Z, Na C, Wang Y, Bai J, Zhang Y, Chen F, Li M, Chen C, Wei L, Li J, Chang X, Qu S, Yang L, Huang J (2021) Influence of low tumor content on tumor mutational burden estimation by whole-exome sequencing and targeted panel

- sequencing. *Clin Transl Med* 11(5):e415. <https://doi.org/10.1002/ctm2.415>
- Zhang N, Zhang P, Chen Q, Zhou K, Liu Y, Wang H, Xie Y, Ren L, Hou W, Yang J, Yu Y, Zheng Y, Shi L (2022) Quartet metabolite reference materials for assessing inter-laboratory reliability and data integration of metabolomic profiling. *bioRxiv:2022.2011.2001.514762*. <https://doi.org/10.1101/2022.11.01.514762>
- Zheng Y, Liu Y, Yang J, Dong L, Zhang R, Tian S, Yu Y, Ren L, Hou W, Han J, Zhang L, Jiang H, Lin L, Lou J, Li R, Lin J, Liu H, Wang D, Dai F, Bao D, Cao Z, Chen Q, Chen Q, Chen X, Gao Y, Jiang H, Li B, Li B, Li J, Liu R, Qing T, Shang E, Shang J, Sun S, Wang H, Wang X, Zhang N, Zhang P, Zhang R, Zhu S, Scherer A, Gloerich J, Wang J, Wang J, Xu J, Hong H, Xiao W, Jin L, The Quartet Project Team, Ding C, Li J, Fang X, Tong W, Shi L (2023) Multi-omics data integration using ratio-based quantitative profiling with Quartet reference materials. *Nat Biotechnol*. <https://doi.org/10.1038/s41587-023-01934-1>
- Zhou JY, Chen L, Zhang B, Tian Y, Liu T, Thomas SN, Chen L, Schnaubelt M, Boja E, Hiltke T, Kinsinger CR, Rodriguez H, Davies SR, Li S, Snider JE, Erdmann-Gilmore P, Tabb DL, Townsend RR, Ellis MJ, Rodland KD, Smith RD, Carr SA, Zhang Z, Chan DW, Zhang H (2017) Quality assessments of long-term quantitative proteomic analysis of breast cancer xenograft tissues. *J Proteome Res* 16(12):4523–4530. <https://doi.org/10.1021/acs.jproteome.7b00362>
- Zook JM, Chapman B, Wang J, Mittelman D, Hofmann O, Hide W, Salit M (2014) Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. *Nat Biotechnol* 32(3):246–251. <https://doi.org/10.1038/nbt.2835>
- Zook JM, Catoe D, McDaniel J, Vang L, Spies N, Sidow A, Weng Z, Liu Y, Mason CE, Alexander N, Henaff E, McIntyre AB, Chandramohan D, Chen F, Jaeger E, Moshrefi A, Pham K, Stedman W, Liang T, Saghbini M, Dzakula Z, Hastie A, Cao H, Deikus G, Schadt E, Sebra R, Bashir A, Truty RM, Chang CC, Gulbahce N, Zhao K, Ghosh S, Hyland F, Fu Y, Chaisson M, Xiao C, Trow J, Sherry ST, Zaranek AW, Ball M, Bobe J, Estep P, Church GM, Marks P, Kyriazopoulou-Panagiotopoulou S, Zheng GX, Schnall-Levin M, Ordonez HS, Mudivarti PA, Giorda K, Sheng Y, Rypdal KB, Salit M (2016) Extensive sequencing of seven human genomes to characterize benchmark reference materials. *Sci Data* 3:160025. <https://doi.org/10.1038/sdata.2016.25>
- Zook JM, McDaniel J, Olson ND, Wagner J, Parikh H, Heaton H, Irvine SA, Trigg L, Truty R, McLean CY, De La Vega FM, Xiao C, Sherry S, Salit M (2019) An open resource for accurately benchmarking small variant and reference calls. *Nat Biotechnol* 37(5):561–566. <https://doi.org/10.1038/s41587-019-0074-6>
- Zook JM, Hansen NF, Olson ND, Chapman L, Mullikin JC, Xiao C, Sherry S, Koren S, Phillippy AM, Boutros PC, Sahraeian SME, Huang V, Rouette A, Alexander N, Mason CE, Hajirasouliha I, Ricketts C, Lee J, Tearle R, Fiddes IT, Barrio AM, Wala J, Carroll A, Ghaffari N, Rodriguez OL, Bashir A, Jackman S, Farrell JJ, Wenger AM, Alkan C, Soylev A, Schatz MC, Garg S, Church G, Marschall T, Chen K, Fan X, English AC, Rosenfeld JA, Zhou W, Mills RE, Sage JM, Davis JR, Kaiser MD, Oliver JS, Catalano AP, Chaisson MJP, Spies N, Sedlazeck FJ, Salit M (2020) A robust benchmark for detection of germline large deletions and insertions. *Nat Biotechnol* 38(11):1347–1355. <https://doi.org/10.1038/s41587-020-0538-8>