Internal Standards for Limit Controls and Absolute Abundance Measurement of Oncogenic Fusions and Mutations

Tom Morrison,[a] Bryan Lo,[b,c] Sophie J. Deharvengt,[d,e] Nick Lazaridis,[a] and Gregory J. Tsongalis[d,e,*]

The success of precision oncology highly depends upon the timely production of accurate results for gene variants associated with therapeutic response. While some variants render tumor cells sensitive to a therapeutic drug, other variants provide a mechanism of resistance to that same therapeutic. Initially, laboratories were performing clinical testing of tumor tissues simultaneously for multiple variants in multiple genes using massively parallel or next-generation sequencing (NGS). This led to a variety of molecular profiling approaches that benefited from advances in sequencing technology and data analysis pipelines. However, the performance characteristics of current NGS assays tend to rely on both extrapolation and inference (1, 2). Currently, most clinical laboratories use NGS assays to generate qualitative results (i.e., variant detected or not detected), and clinical reports contain nonstandardized and noncalibrated values for parameters such as variant allele fractions (3, 4, 5). Positive and negative controls for NGS run in parallel with patient samples provide a crude process control but do not capture the variation that each sample experiences as it passes through the highly complex NGS test procedure (6, 7). Further, limit of detection, analytical sensitivity, and specificity are usually determined empirically with a combination of previously characterized specimens and synthetic nucleic acids that contain a fraction of the entirety of variants being interrogated. Lastly, NGS cancer targets such as circulating tumor DNA (ctDNA) or RNA fusions lack standardized reporting methods due to yield biases in NGS testing (8, 9). NGS cancer testing would benefit from methods that better address the analytical variability of NGS assays.

There is an opportunity to improve the quality control and performance characteristics of NGS assays by implementing spike-in internal standards (IS). Conceptually, the IS in NGS assays are no different than the IS used in quantitative reverse transcription polymerase chain reaction assays (10) and LC-MS methods (11) and would support the need for accuracy described in a recent ctDNA Draft Food and Drug Administration Guidance (12). The IS are simply added or spiked-in to the patient specimens prior to the running of the assay. Standardized Nucleic Acid

[a]Accugenomics, Wilmington, NC, United States; [b]Molecular Oncology Diagnostics Laboratory, Department of Pathology and Laboratory Medicine, The Ottawa Hospital, Eastern Ontario Laboratory Association, Ottawa, ON, Canada; [c]Ottawa Hospital Research Institute, Ottawa, ON, Canada; [d]Clinical Genomics and Advanced Technology, Department of Pathology and Laboratory Medicine, Dartmouth Health System, Lebanon, NH, United States; [e]The Geisel School of Medicine at Dartmouth, Hanover, NH, United States.
*Address correspondence to this author at: Department of Pathology and Laboratory Medicine, Dartmouth Health, 1 Medical Center Dr., Lebanon, NH 03756, United States. E-mail Gregory.j.tsongalis@hitchcock.org.
Received September 08, 2023; accepted October 30, 2023.
https://doi.org/10.1093/jalm/jfad108

Quantification for sequencing (SNAQ™-SEQ) is an IS method that distinguishes the signal generated by IS from the signal generated by the patient specimen or native template (NT) through the NGS sequencing of unique base changes engineered into the IS. Since the IS copy number can be calibrated to reference material specification by combining results of several independent quantification methods, the approach can be used to convert the NGS variant calls into a standardized copy number or abundance measurement. By selecting an appropriate IS copy number, a direct, rather than indirect, limit of detection can be calculated for the specimen to which IS are added. In theory, the use of IS can be applied to any NGS assay, but there are 2 categories of NGS assays: (*a*) cell-free DNA (cfDNA) or ctDNA NGS assays and (*b*) RNA NGS assays for gene fusion detection, where IS can be shown to be particularly meaningful.

INTERNAL STANDARDS IN CFDNA OR CTDNA NGS ASSAYS

The development of cfDNA assays introduced the potential for performing liquid biopsy assessments of the cancer patient at the time of diagnosis and for monitoring purposes using a simple blood draw as the specimen of choice. Once again, the field of precision oncology was to benefit from the technical advances made that would allow for the detection of cfDNA in patient blood but this time with a concern for the sensitivity, accuracy, and precision of these sequencing-based approaches.

The importance of using a liquid biopsy to help monitor therapeutic outcome in solid tumors is unprecedented and has major advantages over the use of tissue biopsy. However, deep sequencing approaches must be quality controlled in a manner that produces high-quality and high-confidence results for variants that will directly impact patient care. SNAQ™-SEQ controls in NGS quality control can help monitor the testing process in individual samples, which is necessary for tumor molecular profiling from liquid biopsies, where using variant allele frequency (VAF) to monitor treatment response is sensitive up to a 2-log variation in background cfDNA. Figure 1A depicts the VAF and abundance results when a synthetic cfDNA sample was created with a constant variant input and a 10-fold range of normal cfDNA. As expected, the VAF ranged >10-fold due to the different amounts of the normal background genomic DNA. However, <25% CV of variant abundance per ml plasma for each variant resulted from calculations using the variant and IS sequence read alanine transaminase counts.

The SNAQ™-SEQ IS provides a near digital polymerase chain reaction-like accuracy to the sequencing platform, standardizing liquid biopsies variant reporting to abundance per ml of plasma.

INTERNAL STANDARDS IN RNA NGS ASSAYS

Gene fusions and exon skipping variants are relevant biomarkers in treatment decision algorithms and can be detected using NGS approaches. RNA-based targeted NGS panels present significant advantages, such as increased analytical sensitivity; however, they are also challenging to standardize. Some of these challenges derive from the relative lability of RNA compared to DNA and the additional step of reverse transcriptase complementary DNA synthesis. Current approaches make use of housekeeping gene expression to measure the intactness of the RNA being assayed and the efficiency of the reverse transcription step. Importantly, they depend on the assumption that the transcripts of gene fusions behave in a manner similar to the transcripts of housekeeping genes. Significantly, these approaches are hampered by the wide

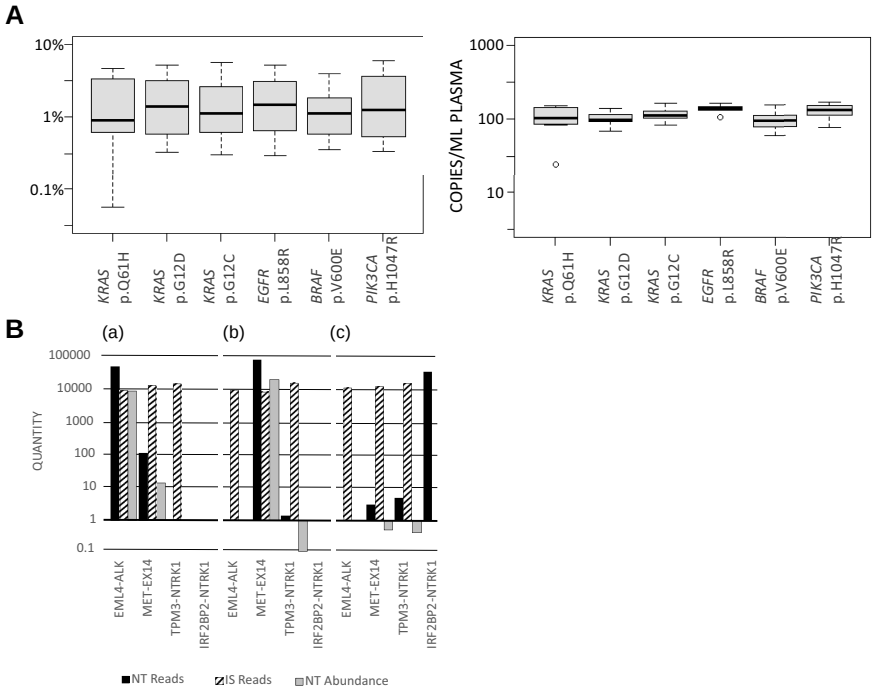

Fig. 1. (A), Boxplots of cancer variants in a sample with differing normal cell free genomic DNA. 30 000 SNAQ-SEQ IS (SNAQ-SEQ ONCO1LB; AccuGenomics) controls and 50 ng 5% ctDNA reference material (NT, SeraCare), which provided a single level of mutation template input, were added in increasing amounts (0 to 500 ng) of cfDNA normal (SeraCare) into 2 mL of DNA negative plasma. The IS mixture was synthetic DNA, fragmented to simulate cfDNA and distinguished from genomic DNA by unique nucleotide alterations every 50 bp. Targeted libraries were prepared from cfDNA extracted using the MagMAX™ Cell-Free Total Nucleic Acid Isolation Kit and the Oncomine™ Pan-Cancer Cell-Free assay reagents (Thermo Fisher Scientific). The indicated COSMIC mutation (*x*-axis) variant allele fraction (left panel, *y*-axis) were extracted from the Ion Reporter VCF. Each of the indicated COSMIC variants were associated with an IS containing a unique near adjacent variant. The COSMIC variant concentrations (right panel, *y*-axis) were calculated from the COSMIC variant alt count and associated IS alt count extracted from the Ion Reporter VCF using the formula NT variant alt count divided by adjacent IS variant alt count * 30 000 IS input copies divided by 2 mL plasma; (B), Thermo Fisher Omni Pan Cancer NGS Reads Converted to Abundance Measurements. The SNAQ™-SEQ single-stranded RNA IS mixture (110 copies each) was added to patient formalin-fixed, paraffin-embedded extraction prior to addition to column purification. Three patient RNA specimens positive for either fusion in *EML4*::*ALK* (a), *MET*-Exon 14 skipping (b), or *IRF2BP2*::*NTRK1* (c) were sequenced using the Oncomine™ Comprehensive Assay v3 panel (Thermo Fisher Scientific), with fusion reads extracted using a modified analysis workflow of Ion Reporter™ software (version 5.16). Read counts for *EML4*::*ALK*.E13A20, *MET*-Exon 14 skipping, and *TPM3*::*NTRK1*.T7N10 fusions were normalized to 1 million total reads (*y*-axis) for both native (black bars) and internal standard (hashed bar) and exported for analysis. Fusion abundance (copies per sample, grey bars) was calculated by NT_reads/IS_reads * IS_input_copies.

variation in the expression levels of housekeeping genes between different tumor tissues. In addition, cell lines engineered to harbor gene fusions are often used in parallel to control for the RNA extraction process and other aspects of RNA NGS assays.

To demonstrate the use of IS in RNA NGS assays, SNAQ™-SEQ ssRNA IS were designed for *EML4*::

ALK, MET-Exon14 skipping, and *TMP3*::*NTRK1* NTs; these standards included unique base changes flanking the fusion for bioinformatic separation. The IS mixture, 110 copies each, was added to patient formalin-fixed, paraffin-embedded RNA samples positive for fusions prior to addition to column purification, then sequenced using the Oncomine Comprehensive Assay v3 panel (Thermo Fisher Scientific), with fusion reads extracted using a modified analysis workflow of Ion Reporter software (version 5.16).

In these proof-of-concept experiments, we determined that the limiting SNAQ™-SEQ IS fusions were detected in each sample, demonstrating a per sample capability of detecting at least 110 fusion copies, i.e., a limit control (Fig. 1B). Further, the semiquantitative NT fusion reads were converted into absolute abundance.

The addition of SNAQ™-SEQ IS did not alter the patient fusion read counts. We believe that standardized abundance measurements could eliminate the less accurate read based thresholds, allowing NGS platforms to use established reporting range analytic validation like other quantitative RNA technologies. Since IS do not require a separate set of library preparation and sequencing reagents, the routine incorporation of IS in RNA NGS assays would be cost effective compared to conventional in parallel fusion controls which require a separate lane of sequencing. The standardization that IS provides will also be advantageous in the cross-validation of different RNA based NGS assays. For example, the ability to explore what levels of *MET*- Exon 14 skipping are relevant biologically will benefit from inter laboratory standardized abundance measurements. When combined with housekeeping gene abundance, the normalized abundance results will control for was added to patient formalin-fixed, paraffin-embedded differential RNA degradation, extraction bias and sequencing bias.

As more clinical decision-making becomes dependent on accurate NGS molecular profiling of DNA and RNA, the introduction of robust quality control measures becomes imperative. We introduce a method, SNAQ^{TM}-SEQ IS, to alleviate much of the variability in interpretation of VAFs, especially when used in the liquid biopsy testing of ctDNA/RNA. A major advantage of the SNAQ™-SEQ technology is that it reduces the potential for false-positive and false-negative results of any quantitative sequencing test for genetic variants. Reduced false positives by using established analytic methods to establish limits of detection and reduced false negatives by ensuring sensitivity performance on a per sample basis are critical to liquid biopsy testing. The disadvantage of using an IS is that the IS must "cover" each hotspot region, making it better suited for characterization of known actionable mutations. However, the incorporation of synthetic IS into each sample allows direct calculation of molecules of the variants per ml of plasma for any sample, on any panel, on any NGS platform. These IS are an appropriate set of reference materials (standards) to allow for comparability across multiple minimal residual disease assays and studies and across a longitudinal timeline to monitor disease progression or response to therapy.

Nonstandard Abbreviations: NGS, next-generation sequencing; ctDNA, circulating tumor DNA; IS, internal standard; SNAQ^{TM}-SEQ, Standardized Nucleic Acid Quantification for sequencing; NT, native template; cfNDA, cell-free DNA; VAF, variant allele frequency.

Author Contributions: *The corresponding author takes full responsibility that all authors on this publication have met the following required criteria of eligibility for authorship: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d)*

agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved. Nobody who qualifies for authorship has been omitted from the list.

Tom Morrison (Conceptualization-Equal, Formal analysis-Equal, Investigation-Equal, Methodology-Equal, Project administration-Equal, Resources-Equal, Writing—original draft-Equal, Writing—review & editing-Equal), Bryan Lo (Investigation-Equal, Validation-Equal, Writing—review & editing-Equal), Sophie J. Deharnegt (Investigation-Equal, Validation-Equal), Nick Lazaridis (Conceptualization-Equal, Funding acquisition-Equal, Investigation-Equal, Methodology-Equal, Supervision-Equal, Writing—review & editing-Equal), and Gregory Tsongalis (Investigation-Equal, Methodology-Equal, Supervision-Equal, Writing—review & editing-Equal)

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form*.

Research Funding: The authors wish to acknowledge the support of the Pathology Shared Resource in the Laboratory for Clinical Genomics and Advanced Technology of the Dartmouth Health System and the Norris Cotton Cancer Center at Dartmouth with NCI Cancer Center Support Grant 5P30 CA023108-37. B. Lo and G.J. Tsongalis received materials and funding for sequencing from Accugenomics.

Disclosures: T. Morrison and N. Lazaridis are employees of Accugenomics.

REFERENCES

1. Robasaky K, Lewis NE, Chuch GM. The role of replicates for error mitigation in next generation sequencing. Nat Rev Genet 2014;15:56–62.
2. Hardwick SA, Deveson IW, Mercer TR. Reference standards for next generation sequencing. Nat Rev Genet 2017;18:473–84.
3. Lennon NJ, Adalsteinsson VA, Gabriel SB. Technological considerations for genome-guided diagnosis and management of cancer. Gen Med 2016;8:112–22.
4. Cleveland MH, Zook JM, Salit M, Vallone PM. Determining performance metrics for targeted next generation sequencing panels using reference materials. J Mol Diagn 2018;20:583–90.
5. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for validation of next generation sequencing based oncology panels. J Mol Diagn 2017;19:341–65.
6. Craig DJ, Morrison T, Khuder SA, Crawford EL, Wu L, Xu J, et al. Technical advance in targeted NGS analysis enables identification of lung cancer risk-associated low frequency TP53, PIK3CA, and BRAF mutations in airway epithelial cells. BMC Cancer 2019;19:1081.
7. Blomquist T, Crawford EL, Yeo J, Zhang X, Willey JC. Control for stochastic sampling variation and qualitative sequencing error in next generation sequencing. Biomol Detect Quantif 2015;5:30–7.
8. Kim H, Park KU. Clinical circulating tumor DNA testing for precision oncology. Cancer Res Treat 2023;55:351–66.
9. Li D, Kusko R, Ning B, Tong W, Johann DJ Jr, Xu J. FDA-led consortium studies advance quality control of targeted next generation sequencing assays for precision oncology. Precis Cancer Med 2021;4:32.
10. Zentlin L, Giacca M. Competitive PCR for precise nuclkeic acid quantification. Nat Protoc 2007;2:20902–2104.
11. Tan A, Awiye K, Li W, Zhang J, Tse FLS, Li W, Zhang J. Use of internal standards in LC-MS bioanalysis. In: Li W, Zhang J, Tse FLS, editors. Handbook of LC-MS bioanalysis: best practices, experimental protocols, and regulations. Hoboken (NJ): Wiley and Sons, Inc; 2013. p. 217–27.
12. U.S. Department of Health and Human Services, Food and Drug Administration. Use of circulating tumor DNA for early—stage solid tumor drug development—draft guidance for industry. 2022. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-circulating-tumor-deoxyribonucleic-acidearly-stage-solid-tumor-drug-development-draft-guidance (Accessed November 2023).