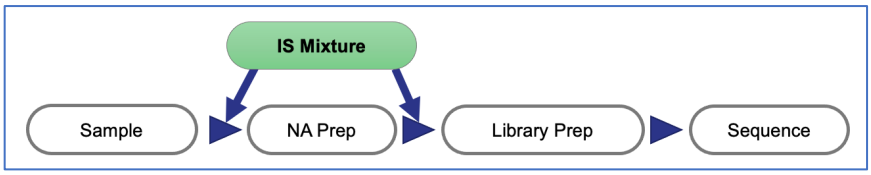
AccuGenomics
A higher standard of accuracy

How to achieve Optimal SNAQ™-SEQ Internal Standard Spike-in Levels

SNAQ™-SEQ internal standards (IS) are customized to each client's application and designed to biochemically mimic the desired native template targets (NT). The level of Internal Standard Mixture (ISM) added to each sample is adjusted during a one-time optimization run to minimize the impact of the ISM on reads per sample while providing optimized quality control concentrations. The optimal IS level depends on the application: quantification, limit of detection, or complexity capture.

The ISM biochemically mimic test targets and can be used as a near full process control when added at the sample lysate step. Alternatively, some clients prefer to add the IS at the library preparation step. ISM is adaptable to any work flow process, and are added in a single reagent addition to the workflow at the stages presented below (Figure 1). For example, ISM can be added as part of an automated method as part of the lysis buffer addition.

FIGURE 1:



Quantification / LOD. The use of ISM enables digital PCR (dPCR) like standardized abundance measurements on your existing NGS platform. Just like dPCR, the accuracy and reporting range depends on the number of molecules counted. However, *unlike* dPCR and current NGS abundance measurements, SNAQ abundance measurements use the ratio of native template (NT) to internal standard (IS) reads (this difference leads to >10-fold increase in sample testing throughput for transcript abundance applications). To maximize the reporting range while minimizing the IS input, the ISM has each individual IS tuned to the geometric mean of its target reporting range (e.g., if target "A" needs a reporting range from 10 to 10,000 copies then the IS for target "A" would be tuned to 316 copies in the ISM). For liquid biopsy measurements of somatic mutations per ml of plasma (SNV, fusions, etc..), when the reporting range is quite narrow (e.g., 3 to 300 copies), it is possible to use 30 IS input in the ISM. For our COVID or INFLUENZA testing products, the IS were designed to provide QC for low viral load samples, therefore the IS input level remains low so as to not interfere with low viral load detection (e.g., 2,000 IS), but may depend on your method's complexity capture and reads per sample.

Limit of Blank. SNV detection is a function of how well your method captures input template as sequence and the NGS background error rate. Typically, a crude cut off is established by the Dx manufacturer to ensure detected variants are well above NGS background error. Regions covered by the SNAQ™-SEQ IS are used to provide a per sample and a per variant NGS error measurement. The IS are a known (designed) sequence (with <1e-8 error rate) that biochemically mimic their target region, therefore any IS variation detected arose from testing error(s). How much ISM is required to provide a limit of blank estimate depends on a couple of factors. The Poisson Exact Test (PET) compares the ratios of alt_NT/coverage_NT and alt_IS/coverage_IS to determine if there was a significant difference. Assuming no IS variant was detected (i.e., zero IS alt count), the optimal IS:NT ratio (coverage_IS & coverage_NT) for a 5% alpha PET significance depends on the variant caller alt count PASS threshold.

AccuGenomics
A higher standard of accuracy

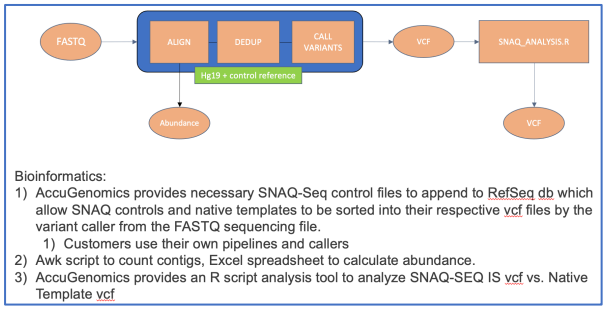
alt_NT	IS:NT reads
2	5.3
3	2.4
4	1.6
5	1.1

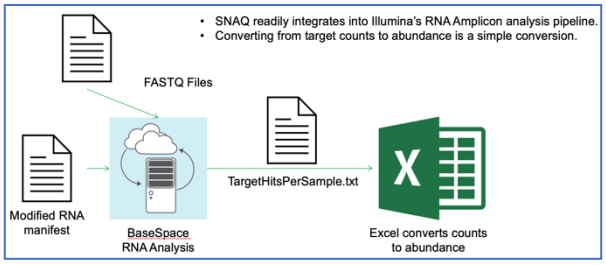
The table (inset, right) indicates that the required coverage *ratio* decreases as the required alt_NT increases. For larger NGS panels that were developed to cover hundreds of cancer genes, we recommend a more focused cancer specific IS mixture, which typically covers < 20 actionable mutations, to allow sufficient IS coverage while minimizing the IS reads per sample. When using fragmented or unfragmented ISM, users should perform a one-time optimization experiment to adjust the spiked-in ISM concentration to achieve expected IS:NT read ratios within their existing workflow.

Complexity Capture. AccuGenomics complexity capture controls (CC) are unique molecules that provide an indication of how much of a sample's original template is captured as sequence. Complexity controls are full process controls that indicate the loss of sample complexity, or detect issues that arise from UMI deduplication inaccuracies. A single optimization is required to establish the optimal spike-in CC level. Some suggested levels are 30,000 CC spiked into plasma, 10,000 CC spiked into amplicon viral testing, or 100,000 CC spiked into non-targeted sequencing. Ideally, you want to detect a few hundred unique CC as downstream sequencing reads.

Analysis. Briefly with regard to data analysis, internal standards (IS) are separated from the native template (NT) reads during alignment (they form their own contigs due to the base changes we design into the IS), this can occur using a number of methods. We also provide our customers with an example analysis script to assist their bioinformatics team to incorporate ISM into their workflow, two examples are shown below (Figure 2). ISM is easily adaptable to any informatics pipeline, and AccuGenomics can share their expertise as part of the design process with each client.

FIGURE 2:





The use of SNAQ™-SEQ internal control standards allows for greater accuracy in a wide range of targeted panels, improves sensitivity while maintaining specificity, enhances transcript abundance / MRD / CNV / fusion measurements, and provides an integral internal process control metric for NGS laboratories that run Clinical Dx or pathogen detection assays. AccuGenomics would welcome the opportunity to discuss IVD development of your targeted NGS panel, and share solutions to provide absolute accurate measurements for your NGS tests.

AccuGenomics – a higher standard of accuracy.

For more information:

- Message us for a confidential conversation
- Follow us on LinkedIn
- Visit our website