Next-Generation Sequencing Somatic and Germline Assay Troubleshooting Guide Derived From Proficiency Testing Data

Valentina Nardi, MD; Karen D. Tsuchiya, MD; Annette S. Kim, MD, PhD; Lora J. H. Bean, PhD; Jaimie G. Halley, MLS(ASCP)CM; Thomas A. Long, MPH; Szabolcs Szelinger, PhD; Patricia Vasalos, BS; John A. Thorson, MD, PhD; Ann M. Moyer, MD, PhD; Joel T. Moncur, MD, PhD

• ***Context.***—**Next-generation sequencing–based assays are increasingly used in clinical molecular laboratories to detect somatic variants in solid tumors and hematologic malignancies and to detect constitutional variants. Proficiency testing data are potential sources of information about challenges in performing these assays.**

Objective.—**To examine the most common sources of unacceptable results from the College of American Pathologists Next-Generation Sequencing Bioinformatics, Hematological Malignancies, Solid Tumor, and Germline surveys and provide recommendations on how to avoid these pitfalls and improve performance.**

Design.—**The College of American Pathologists next-generation sequencing somatic and germline proficiency testing survey results from 2016 to 2019 were analyzed to identify the most common causes of unacceptable results.**

Results.—**On somatic and germline proficiency testing surveys, 95.9% (18 815/19 623) and 97.8% (33 890/34 641) of all variants were correctly identified, respectively. The most common causes of unacceptable results related to sequencing were false-negative errors in genomic regions that were difficult to sequence because of high GC content. False-positive errors occurred in the context of homopolymers and pseudogenes. Recurrent errors in variant annotation were seen for dinucleotide and duplication variants and included unacceptable transcript selection and outdated variant nomenclature. A small percentage of preanalytic or postanalytic errors were attributed to specimen swaps and transcription errors.**

Conclusions.—**Laboratories demonstrate overall excellent performance for detecting variants in both somatic and germline proficiency testing surveys. Proficiency testing survey results highlight infrequent, but recurrent, analytic and nonanalytic challenges in performing next-generation sequencing–based assays and point to remedies to help laboratories improve performance.**

(*Arch Pathol Lab Med.* 2022;146:451–461; doi: 10.5858/arpa.2020-0842-CP)

Next-generation sequencing (NGS) has become a mainstay for molecular diagnostic laboratories that support personalized medicine. Proficiency testing (PT) plays an important role in evaluating the ongoing performance of accredited laboratories that perform clinical NGS testing. PT also provides an opportunity for laboratories to compare their performance with that of their peers and to receive feedback to ensure the highest quality care for patients.

Since 2015 and 2016, the College of American Pathologists has offered PT for germline and somatic NGS testing, respectively. Publications on the performance of laboratories on NGS PT surveys have described excellent overall performance for the detection of single nucleotide variants (SNV) and small insertions and deletions.[1–3] As clinical practice has evolved to include more genes and variants,[4] so too has PT. PT samples have increasingly included less common and more challenging variants. While overall performance remains excellent, particularly for the detection

Accepted for publication April 27, 2021.
Published online August 23, 2021.
Supplemental digital content is available for this article at https://meridian.allenpress.com/aplm in the April 2022 table of contents.

From the Department of Pathology, Massachusetts General Hospital, Boston (Nardi); Department of Laboratories, Seattle Children's Hospital, Seattle, Washington (Tsuchiya); Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts (Kim); Department of Pathology, PerkinElmer Genomics, Pittsburgh, Pennsylvania (Bean); Proficiency Testing (Halley, Szelinger, Vasalos), Biostatistics (Long), College of American Pathologists, Northfield, Illinois; Department of Pathology, UC San Diego, Del Mar, California (Thorson); Department of Laboratory Medicine and Pathology, Mayo Clinc, Rochester, Minnesota (Moyer); and the Office of the Director, the Joint Pathology Center, Silver Spring, Maryland (Moncur).

The authors have no relevant financial interest in the products or companies described in this article.

All authors are current or past members of the College of American Pathologists Molecular Oncology Committee or the College of American Pathologists Biochemical and Molecular Genetics Committee. Halley, Long, Szelinger, and Vasalos are employees of the College of American Pathologists. The identification of specific products or scientific instrumentation is considered an integral part of the scientific endeavor and does not constitute endorsement or implied endorsement on the part of the authors, DoD, or any component agency. The views expressed in this article are those of the authors and do not reflect the official policy of the Department of Army/Navy/Air Force, Department of Defense, or U.S. Government. All others declare no potential conflicts of interest with the contents of this manuscript.

Corresponding Author: Joel T. Moncur, MD, PhD, Office of the Director, the Joint Pathology Center, 606 Stephen Sitter Ave, Silver Spring, MD 20910 (email: joel.t.moncur.mil@mail.mil).

Table 1. Annual Enrollment in College of American Pathologists Next-generation Sequencing (NGS) Somatic and Germline Proficiency Testing Programs for 2015 to 2019 Based on B-mailing Enrollments for that Year

Proficiency Test	Enrollment				
	2015	2016	2017	2018	2019
Somatic					
NGS Bioinformatics (NGSB1 and NGSB2)	-	63	62	71	52
NGS Hematologic Malignancies (NGSHM)	-	76	101	129	154
NGS Solid Tumor (NGSST)	-	140	210	242	265
Germline					
NGS	161	146	186	209	210

of SNVs, some variants are more challenging for NGS-based methodologies.

The purpose of this study was to identify, describe, and discuss recurrent causes of unacceptable results on somatic and germline NGS PT surveys. By documenting and discussing these challenges, we aimed to raise awareness about these common pitfalls among laboratories performing the assay and those receiving the NGS results and provide the laboratories with remedies to help further improve performance.

MATERIALS AND METHODS

Data from the College of American Pathologists somatic and germline NGS PT surveys were analyzed to identify the most common causes of unacceptable results. This study included data from the following 3 different somatic surveys: the NGS Bioinformatics PT surveys (NGSB1 and NGSB2) from 2018 through 2019, and the NGS Hematologic Malignancies (NGSHM) and NGS Solid Tumor (NGSST) PT surveys from 2016 through 2019. For germline testing data, the NGS-Germline surveys from 2016 through 2019 were analyzed. Additional germline PT surveys were available but were not included in this study because they included participants who used NGS and non-NGS methods. For information about the design of the somatic and germline NGS PT surveys, see the supplemental digital content at https://meridian.allenpress.com/aplm in the April 2022 table of contents, containing data and 4 tables.

For the somatic NGS PT surveys, all summarizations and analyses were completed using SAS (version 9.4; Cary, North Carolina). All data were analyzed retrospectively; the analysis included all final participant-submitted data used in the participant summary report documents. For any given variant, data were included in the analysis when laboratories reported that their assay covered the variant and the provided materials contained the variant above the laboratory's reported limit of detection.

For the germline NGS PT surveys, all summarizations and analyses were completed using R (version 3.6.1, https://www.rproject.org/), and assessment of performance included responses received by the survey due date. The overall rate of acceptable responses for the detection of all variants was calculated for graded genomic positions. Owing to the challenges associated with free-text responses for variant annotations in this survey, only genes and/or chromosomal positions or intervals with unacceptable rates of 5% or more and involving at least 3 laboratories were reviewed. Insertions and duplications were combined for analysis because of the design of the result form. The percentage of laboratories that correctly used appropriate nomenclature (either preferred or acceptable) was assessed.

For both somatic and germline surveys, selected analysis of variant types (including deletions, duplications/insertions, and SNVs) was also performed to help identify the most common types of variants associated with unacceptable results.

RESULTS

As of 2019, the number of laboratories enrolled in the NGSB1/NGSB2, the NGSHM, and NGSST PT somatic surveys was 52, 154, and 265, respectively. As of 2019, 210 laboratories were enrolled in the germline NGS survey (Table 1).

The College of American Pathologists approaches NGS PT as an iterative cycle that is designed to support the adaptation and evolution of PT to match changes in clinical practice. The cycle involves the following: (1) collecting data about laboratory practices; (2) using those data to develop and adapt PT; (3) assessing laboratory performance; and (4) providing feedback and education to laboratories through participant summary reports, presentations, and publications (Figure 1).

From 2016 through 2019, the assessment of laboratory performance for somatic and germline variants demonstrated excellent overall performance with 95.9% (18 815/19 623) and 97.8% (33 890/34 641) of all variants correctly identified, respectively. Despite the overall excellent performance of laboratories on NGS PT, recurrent causes of unacceptable results were revealed. In both somatic and germline PT surveys, there were analytic errors leading to false-negative

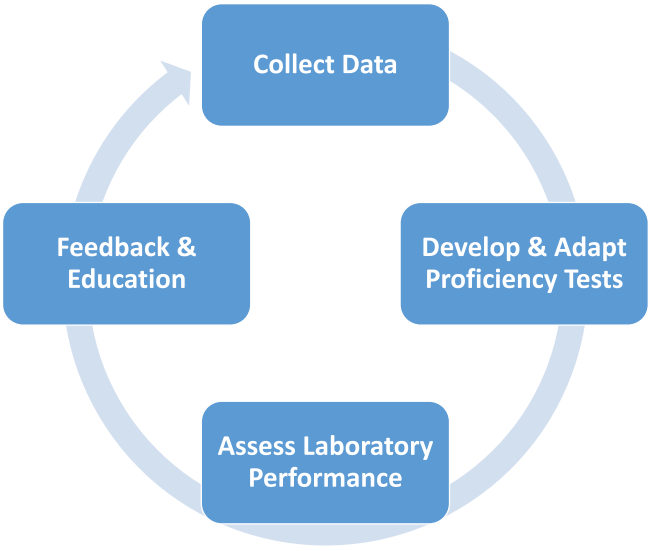

Figure 1. *The continuous proficiency testing cycle. The cycle of steps used by the College of American Pathologists to ensure next-generation sequencing proficiency testing adapts and evolves with changes in clinical practice.*

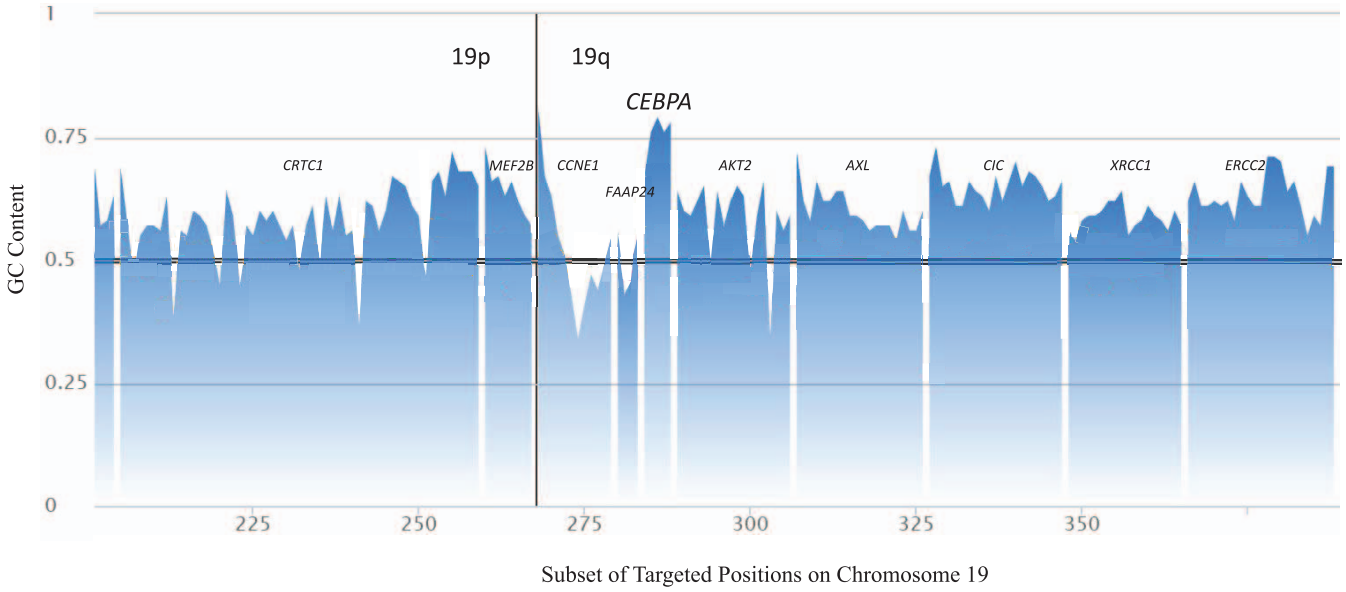

Figure 2. *GC content of* CEBPA. *GC content (ratio) of CEBPA compared with selected adjacent genes on chromosome 19. The x-axis represents the absolute number of targeted regions on chromosome 19 on this targeted panel (targeted regions 200–385 shown, including* CEBPA*).*

and false-positive results as well as errors in annotation. There were also nonanalytic errors involving the preanalytic and postanalytic phases of the testing process.

Analytic Errors

False-Negative Errors.—*Variants in Difficult to Sequence Genomic Regions with High GC Content.*—Polymerase chain reaction (PCR) amplification and sequenced reads alignment are challenging for high GC content targets. In the somatic NGSHM survey, a well-known gene that is challenging to sequence is *CEBPA*, an intron-less gene with approximately 75% GC content in the coding region and the presence of a trinucleotide repeat region (Figure 2). The nature of the recurrent variants in *CEBPA* also creates sequencing challenges, including complex variants and frequent occurrences of variants in mononucleotide repeat regions.[5] Laboratories using amplicon-based platforms to detect a 1-bp duplication in *CEBPA* (NM_004364.4:c.68dupC; p.His24fs*84) had a mean unacceptable rate of 28.3% over 3 surveys that included this variant (range, 16.7%–35.9%) despite engineering the mutation at a high variant allele fraction (range, 29.0%–50.0%) (Supplemental Table 1). Less than 1% of laboratories participating in this survey and using capture-based enrichment had false-negative results for this *CEBPA* mutation. The unacceptable responses were likely secondary to base quality and alignment issues and not to poor coverage. In fact, the average coverage by participant laboratories was high (×1275; ×1097; ×1451).

In the germline survey, 10 of 64 (15.6%) unique genomic positions associated with a false-negative rate of at least 5% were located within GC-rich regions. Similarly, among 32 targeted genomic positions that laboratories indicated they could not evaluate, 14 (43.8%) were within GC-rich regions.

False-Positive Errors.—*Variants from Homopolymer Regions.*—Homopolymers (HPs) in genomics are sequences of consecutive identical bases, also known as microsatellites, which can occur as mononucleotide repeats, or repeats of 2, 3, 4, or more nucleotides. HPs are prone to increased mutagenesis due to in vivo replication slippage,[6] but similar errors can occur in vitro during PCR amplification.[7] Therefore, distinguishing somatically acquired deletions or insertions occurring within the same repeated nucleotide(s) from in vitro artifact is particularly difficult. For this reason, genomic regions with HPs are prone to false-positive results.

The somatic NGSHM PT survey contains an example of this phenomenon. Low-level false-positive *ASXL1* mutations (NM_015338.5:c.1934dupG; p.Gly646Trpfs*12) were incorrectly reported by 7.0% (9 of 129) of laboratories in 2018 and 1.3% (2 of 154) in 2019 (Supplemental Table 2) with a variant allele fraction (VAF) between 4.3% and 14.5%. This variant is a duplication of a single guanine occurring within an 8-bp mononucleotide guanine repeat sequence (8G repeat) that extends from c.1927 to c.1934 (Figure 3, A). At low fraction (approximately ≤5%), it is known to be a recurrent artifact due to slipped strand mispairing,[8] both naturally and in vitro during enzymatic replication, and can result in both the duplication and deletion of a G (c.1934delG; p.Gly645fs) (typically deletion is more common than duplication). However, this same slippage can occur biologically as a pathogenic mutation (Figure 3, B).[9,10]

Although in the somatic survey errors in HP sequencing interpretation led to false-positive calls, in the germline survey, this same issue also resulted in false-negative interpretations when laboratories presumed mutations at these sites were artifactual. Positions with small insertions or duplications in an HP region have been included in some germline surveys. Among the unique 64 genomic positions associated with a false-negative rate greater than 5%, 8 (12.5%) were in regions of low genomic complexity (4 in HP regions, 4 in other repetitive regions). As an example, a deletion in *CEP290* in the NGS-Germline 2018-A survey, with an intended response of NM_025114.3:c.3574-9delT, was located in a stretch of 8 adenosine nucleotides (8A repeat). The detection rate of this homozygous deletion (with appropriate zygosity and variant description reported)

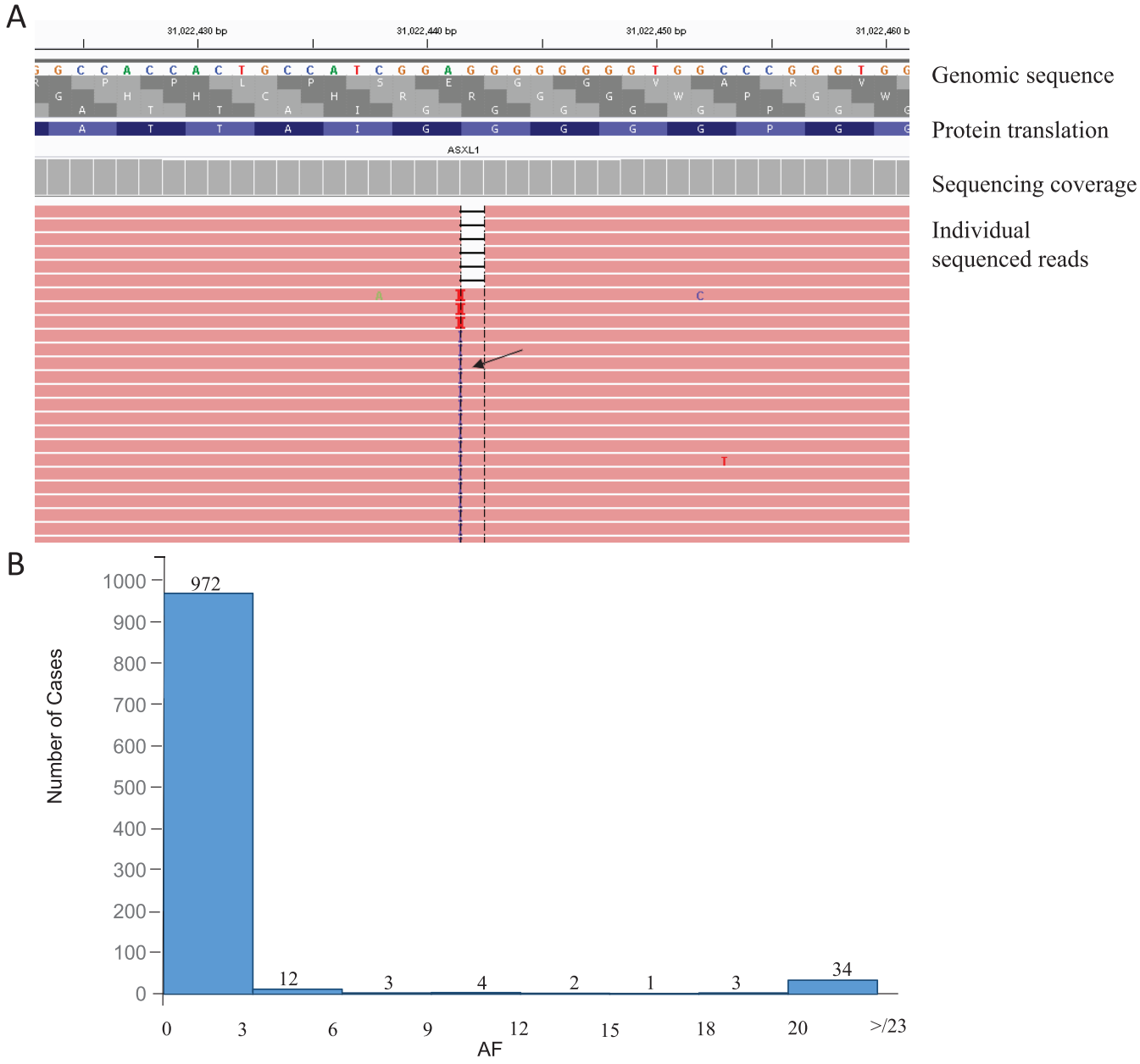

Figure 3. *Mutation in a homopolymer region. A, Screenshot of the Integrative Genomics Viewer showing a genomic region with a homopolymer (GGGGGGG) and sequenced reads showing a deletion (black lines) and 2 different insertion variants (red and purple "I"), the most frequent one (black arrow) consistent with the* ASXL1 *(NM_015338.5) c.1934dupG; p.Gly646Trpfs*12) (G646fs*). B, Histogram of 1000 randomly selected clinical cases showing the bimodal distribution of allelic fractions (AF) for* ASXL1 *variant c.1934dupG, shown on the x-axis versus the number of cases (Count) on the y-axis.*

was 84.4% (65 of 77) among the laboratories that could evaluate this region.

Pseudogene Interference.—Pseudogenes are genomic sequences that are similar to a gene but are considered to be nonfunctional. Owing to their sequence similarity to functional genes, pseudogenes can interfere with short-read NGS technology, resulting in mismapping of reads between the gene and pseudogene that can lead to either false-negative or false-positive calls. *PRSS1* encodes a trypsinogen and has 2 known pseudogenes, *PRSS3P1* and *PRSS3P2*. The NGS-Germline 2019-A survey included the genomic position chromosome 7:g.142460335 (NM_002769.4), which is located in *PRSS1*. Of 93 participants, 47.3% (44) responded "variant not detected," 51.6% (48) responded that an SNV was detected, and 1.1% (1) responded that the locus could not be evaluated. This lack of consensus was thought to be due to pseudogene interference (Figure 4, A and B), supported by 1 participating laboratory that confirmed by Sanger sequencing that the variant was not present in the gene.

Errors in Annotation.—*Dinucleotide Variants.*—According to the Human Genome Variation Society (HGVS) guidelines,[11] a substitution changes 1 nucleotide into 1 other nucleotide; thus, 2 sequential nucleotide changes (dinucleotide changes) are not considered substitutions but rather deletion–insertion (delins) variants. Therefore, dinucleotide

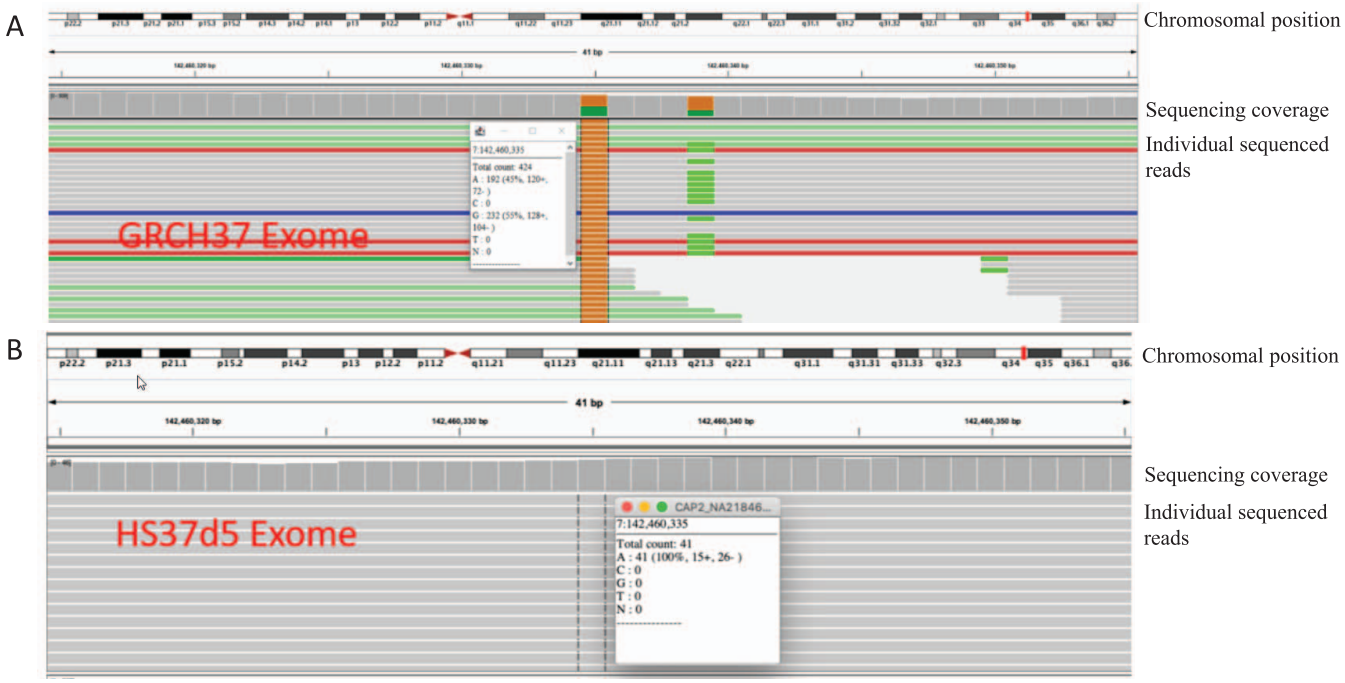

Figure 4. *Pseudogene interference. A, A variant is detected (highlighted by the colored position in the sequencing coverage plot and in the individual sequenced reads) in the Next-Generation Sequencing-Germline 2019-A survey in* PRSS1 *when reads are aligned to the standard hg 19/GRCh37 genome. The total reads including the reference and variant nucleotides are presented in the box. B, When the reads are aligned to the hs37d5 genome, the variant is not detected.*

changes should be reported as a single delins variant that is merged for both the complementary DNA and protein annotations (eg, c. and p.). Notably, when 2 variants are instead separated by 1 or more nucleotides, they should preferably be described individually in HGVS c. nomenclature and not as a delins (unless they together affect 1 amino acid).

In the somatic NGSST surveys, between 11.9% (5 of 42) and 37.5% (12 of 32) of laboratories did not correctly report variants detected as delins. Of note, the variants in the surveys were engineered at a VAF equal to or above 10% (between 10% and 45%; most laboratories having a limit of detection of 5%–10%). The most challenging dinucleotide to correctly identify was *CDKN2A* NM_000077.4:c.171_172delCCinsTT; p.Arg58*, while the least challenging was *KRAS* NM_004985.3:c.180_181delTCinsAA; p.Gln61Lys (Table 2 and Figure 5). Laboratories either identified only 1 of 2 nucleotide changes (eg, the second change was categorized as synonymous and therefore was not reported) or they reported the dinucleotide variant as 2 single nucleotide substitutions in *cis*. These are not errors in detection but are errors in annotation that could result concomitantly in a false-negative and a false-positive result.

This concept of dinucleotide annotation is relevant in germline testing as well and may be encountered by laboratories during routine testing. The NGS-Germline 2019-A survey included a variant in *POT1*. Laboratories were able to detect the presence of a SNV at the indicated position (chr7:g.124499003) as NM_015450.2:c.702+8A>T (19 of 23; 82.6%); in addition, a subset of laboratories correctly recognized that a dinucleotide variant with a correct annotation of c.702+8_702+9delinsTG, was present, despite the fact that only 1 genomic coordinate was listed for

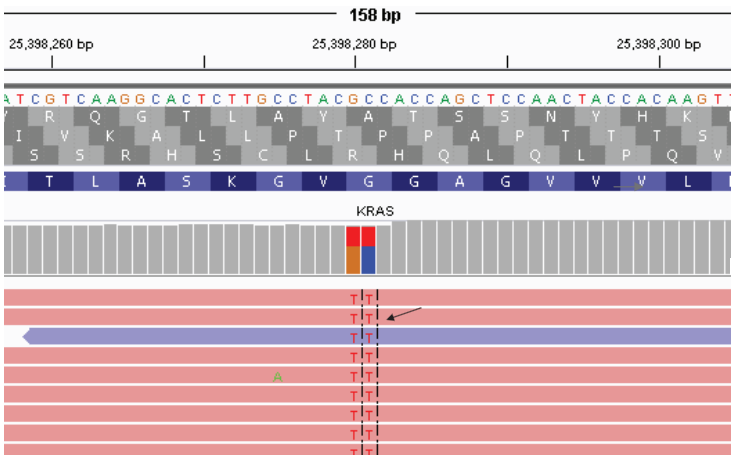

Figure 5. *Dinucleotide changes. Screenshot of the Integrative Genomics Viewer showing sequenced reads containing 2 single base substitutions in cis, indicated by nucleotides in red font and by the arrow, which should be reported as the following deletion–insertion variant:* KRAS *(NM_004985.3) c.38_39delGCinsAA; p.Gly13Glu (p.G13E). The red and blue colors of the sequenced reads indicate reads sequenced using forward or reverse primers, respectively.*

Table 2. Laboratory Performance for Detection of Dinucleotide Variants

Gene	Transcript	Nucleotide Change	Protein Change	Chromosomal Position
CDKN2A	NM_000077.4	c.171_172delCCinsTT	p.Arg58*	chr9:21971186_21971187delGGinsAA
HRAS	NM_005343.2	c.37_38delGGinsAA	p.Gly13Asn	chr11:534285_534286delCCinsTT
HRAS	NM_005343.2	c.181_182delCAinsTT	p.Gln61Leu	chr11:533874_533875delTGinsAA
KRAS	NM_004985.3	c.38_39delGCinsAA	p.Gly13Glu	chr12:25398280_25398281delGCinsTT
KRAS	NM_004985.3	c.180_181delTCinsAA	p.Gln61Lys	chr12:25380277_25380278delGAinsTT
NRAS	NM_002524.4	c.182_183delTTinsTG	p.Gln61Pro	chr1:115256528_115256529delAAinsTG
NRAS	NM_002524.4	c.180_181delACinsTA	p.Gln61Lys	chr1:115256530_115256531delGTinsTA

Abbreviation: VAF, variant allele fraction.

laboratories to query. These laboratories reported this variant as a delins (2 of 23; 8.7%).

Duplication Variants.—According to HGVS, duplications are sequence changes where, compared with a reference sequence, a copy of 1 or more nucleotides is inserted directly 3′ of the original copy of that sequence. Insertions that duplicate the immediately preceding nucleotide or sequence should be described as duplications, not as insertions. As for other variants, the most 3′ position possible is arbitrarily assigned to be where the duplication occurs, the so-called "3′ rule," which is particularly important when the duplication involves stretches of tandem repeats.

In the somatic PT surveys, just over one quarter (25.8%; 8 of 31) of laboratories missed an *ERBB2* duplication (NM_004448.2:c.2313_2324dupATACGTGATGGC; p.Tyr772_A775dupTyrValMetAla) in 2018 (NGSB 1/2B), engineered at a VAF (25.0%), and 12.4% (21 of 170) missed the same duplication at 38.9% VAF in 2019 (NGSST A) (Table 3).

The laboratories missing the duplications either detected them but did not apply the 3′ rule or reported it as an insertion with or without applying the 3′ rule (Figure 6, A and B). Per HGVS, indel variants are right-aligned, while most variant callers left-align them. These are likely not errors in detection but rather errors in annotation. On the somatic NGSHM survey, fewer than 8.7% of laboratories missed a 4-bp duplication in *NPM1* (NM_002520.6:c.860_863dupTCTG; p.Trp288fs*12) engineered at 11.8%, 26.4%, and 45.0% VAF (Table 3). As this is a critical variant in hematologic malignancies, laboratories may have optimized their pipelines for the correct annotation of this specific variant.

In the NGS-Germline 2018-A survey, a variant in *PRKAR1A* highlights similar challenges. This variant results in an intronic single base duplication. Of 93.2% of laboratories (82 of 88) that detected a variant for a query on chromosome 17 (g.66519855-66519864; NM_002734.4), 53 (64.6%) correctly described the variant as c.349-5dupT or c.349-5dup. Other laboratories reported this variant as an indel or deletion, used "ins" instead of "dup," or used a variety of other incorrect nomenclature, including c.349-5_349-4insT, c.349-9_349-8insT, c.349-8_349-9insT, c.349-8-349-9 insT, and c.-5_-4insT (Table 4). Many of these errors demonstrate a failure to apply the 3′ rule.

Transcript Annotation.—While a recommended transcript including the version is provided in the germline survey for each genomic position tested, laboratories are allowed to use an alternate transcript or version, but they must indicate which transcript and version was used. In some cases, laboratories received an unacceptable grade due to failure to list the alternate transcript used. An example of a significant difference in interpretation owing to an alternate transcript version involves *COL5A2* in the NGS-Germline 2018-B survey. The transcript indicated in the survey instructions (NM_000393.3) would result in a "variant not detected" call, while using the transcript NM_000393.4 would result in an SNV call (c.3411T>C). This highlights the importance of correctly reporting the transcript and version used.

Nonanalytic Errors

Preanalytic and Postanalytic Errors.—*Specimen swaps and transcription errors.*—Preanalytic and postanalytic clerical errors are a relatively uncommon yet recurring cause of discordant findings in PT surveys. Specimen swaps and/or transcription errors were seen in at least half of the somatic NGS PT mailings in 2017 through 2019. Specimen swaps

Table 3. Laboratory Performance for Detection of Duplication Variants

Gene	Transcript	Nucleotide Change	Protein Change	Chromosomal Position
ERBB2	NM_004448.2	c.2313_2324dupATACGT GATGGC	p.Tyr772_A775dupTyrValMetAla	chr17:37880984_37880995dupATAC GTGATGGC
ERBB2	NM_004448.2	c.2313_2324dupATACGT GATGGC	p.Tyr772_A775dupTyrValMetAla	chr17:37880984_37880995dupATACGT GATGGC
KIT	NM_000222.2	c.1504_1509dupGCCTAT	p.Ala502_Tyr503dup	chr4:55592180_55592185dupGCCTAT
NPM1	NM_002520.6	c.860_863dupTCTG	p.Trp288fs*12	chr5:170837544_170837547dupTCTG
NPM1	NM_002520.6	c.860_863dupTCTG	p.Trp288fs*12	chr5:170837544_170837547dupTCTG
NPM1	NM_002520.6	c.860_863dupTCTG	p.Trp288fs*12	chr5:170837544_170837547dupTCTG

Abbreviation: VAF, variant allele fraction.

Table 2. Extended

Labs That Tested for Variant, N	No. (%) of Labs That Detected Variant	No. (%) of Labs That Missed the Variant	Engineered VAF, %	Average Reported VAF, %	Median Coverage	Mailing
32	20 (62.5)	12 (37.5)	20.0	16.2	1997.0	NGSB1/2 A 2018
33	25 (75.8)	8 (24.2)	25.0	20.6	1120.0	NGSB1/2 A 2019
32	28 (87.5)	4 (12.5)	30.0	26.8	1996.0	NGSB1/2 B 2018
111	85 (76.6)	26 (23.4)	14.2	13.7	1985.0	NGSHM A 2019
42	37 (88.1)	5 (11.9)	10.0	8.5	2955.0	NGSB1/2 A 2018
31	26 (83.9)	5 (16.1)	45.0	43.1	6888.6	NGSB1/2 A 2016
39	31 (79.5)	8 (20.5)	20.0	16.8	1987.0	NGSB1/2 A 2019

were presumed when 2 of the 3 PT specimens or their results appeared transposed on the PT survey result form.

The handful of variants that were reported and were very similar to those expected, but with slight nomenclature differences, were presumed to be transcription errors. There were also presumed nonanalytic errors that consisted of submitting results of specimens tested in prior mailings or reporting the same results for more than 1 PT specimen. For 2017 through 2019, 6 NGSST participants and 7 NGSHM participants had unacceptable results due to specimen swaps and/or transcription errors (Table 5).

For the germline survey, a single specimen is included in each mailing; therefore, specimen swaps are not relevant (aside from the laboratory swapping the PT specimen with another clinical sample, which has not been observed). Transcription errors likely occur in the germline survey, but cannot be readily quantified, in part because all variants are reported manually on the result form. As a result, it is not always clear whether an incorrect response is due to a misinterpretation or a transcription error.

DISCUSSION

The overall NGS assay performance of the laboratories was excellent, with 95.9% and 97.8% accurate detection of all examined variants across 4 different somatic and germline PT surveys, respectively. Despite this superb accuracy, we sought to identify and categorize the underlying causes of unacceptable results on somatic and germline NGS PT and to provide a guide to help laboratories avoid these errors (summarized in Table 6). For all types of NGS PT, the most common causes of unacceptable results were annotation errors rather than sequencing errors. In addition, for certain somatic NGS PT surveys (ie, NGSHM and NGSST), occasional causes of errors included specimen swaps and transcription errors. These errors do not reflect the ability of NGS assays to accurately detect variants. A minority of unacceptable PT results are due to sequencing challenges pertaining to the detection of variants in regions with high GC content, variants in HP regions, and pseudogene interference.

Sequencing Challenges in Regions With High GC Content

GC-rich DNA sequences are more thermostable and can form secondary structures (hairpin loops) and consequently are more difficult to amplify by PCR. A template (or at least a 100 to 150 base-long part) with greater than 60.0% to 65.0% GC content could reasonably be considered difficult to sequence.[12,13] Library construction protocols are generally recognized to be biased toward fragments of intermediate GC content, the most GC-rich fraction of the target DNA being underrepresented.[14] Most often, the solution of choice is to add dimethyl sulfoxide to a final concentration of 2.5% to 5.0% (it seems to be effective in templates with up to 60.0% to 72.0% GC content), a 5-minute heat-denaturation step or 1 molar betaine.[13] *CEBPA* mutations are the perfect example of the challenges in detecting variants with high GC content, with a coding sequence that is over 75.0% GC rich, a trinucleotide repeat region, and complex mutations that frequently occur in mononucleotide repeats. Laboratories should be aware that many NGS library preparation methods are optimized for an intermediate GC content, and this will result in drops in coverage or overall limited coverage with a high error rate for high GC content regions. Therefore, laboratories should consider excluding from the list of covered targets those with limited coverage due to

Table 3. Extended

Labs That Tested for Variant, N	No. (%) of Labs That Detected Variant	No. (%) of Labs That Missed the Variant	Engineered VAF, %	Average Reported VAF, %	Median Coverage	Mailing
31	23 (74.2)	8 (25.8)	25.0	24.7	2874.0	NGSB1/2 B 2018
170	149 (87.6)	21 (12.4)	38.9	33.3	1990.0	NGSST A 2019
167	157 (94.0)	10 (6.0)	48.7	44.7	1990.0	NGSST B 2019
48	47 (97.9)	1 (2.1)	45.0	40.4	5795.0	NGSHM A 2016
69	63 (91.3)	6 (8.7)	11.8	12.3	3070.1	NGSHM A 2017
89	85 (95.5)	4 (4.5)	26.4	21.5	1496.0	NGSHM A 2018

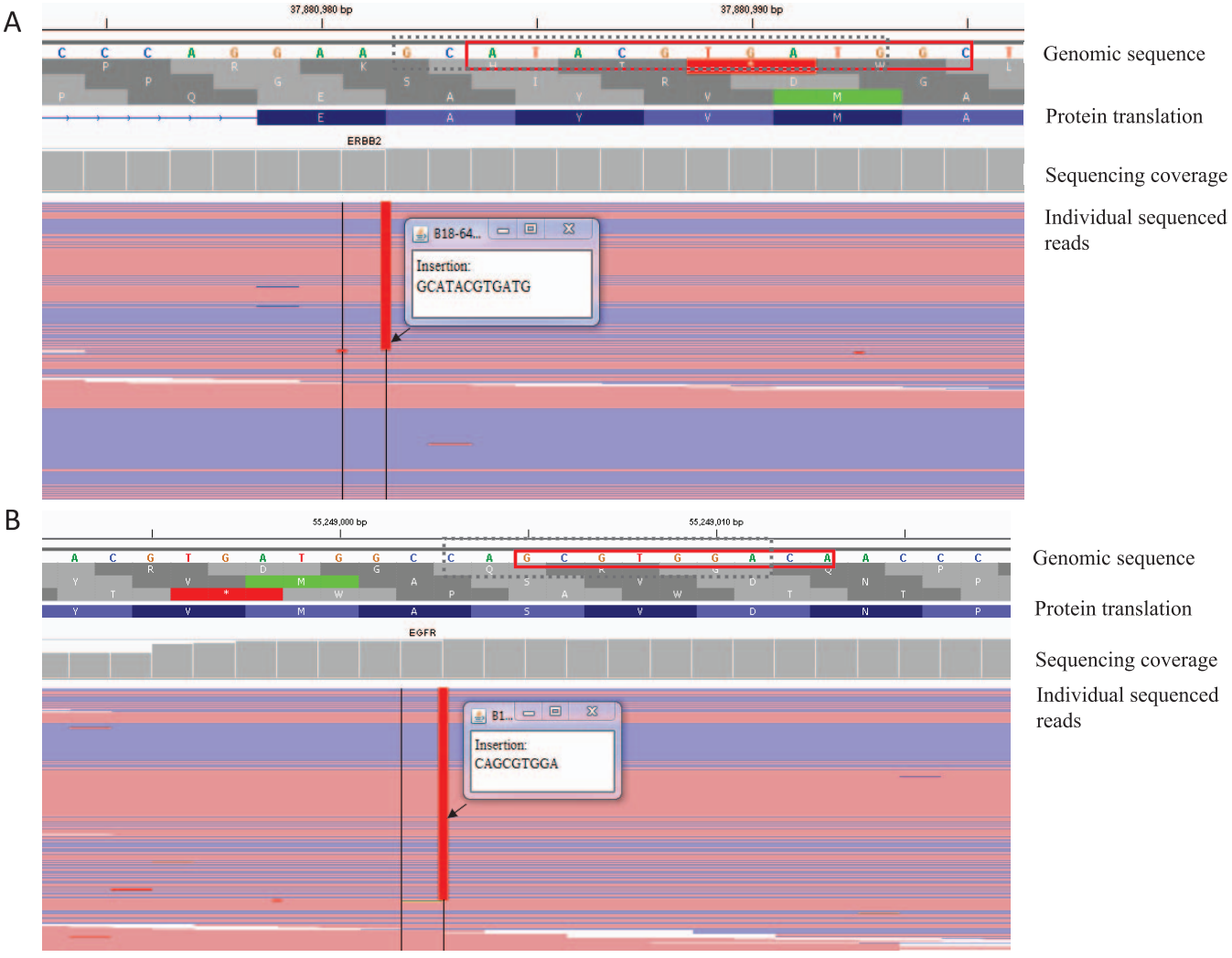

Figure 6. *Duplicating insertions (duplications) and 3′ alignment. Screenshot of the Integrative Genomics Viewer showing sequenced reads containing a duplicating insertion indicated by a vertical red bar (arrow) for (A)* ERBB2 *(NM_004448.2) c.2313_2324dup ATACGTGATGGC; p.Tyr772_Ala775dupTyrValMetAla and (B)* EGFR *(NM_005228.3) c.2303_2311dupGCGTGGACA; p.Ser768_Asp770dup. The dotted black box of a portion of the reference sequence indicates the inserted sequence, while the red box indicates the sequence that should be reported based on the 3′ alignment rule. The red and blue colors of the sequenced reads indicate reads sequenced using forward or reverse primers, respectively.*

high GC content or consider using an orthogonal method (namely, Sanger sequencing or dedicated NGS assay) to supplement panel testing with limited or no coverage for genes like *CEBPA*.

Sequencing Challenges in HP Regions

Laboratories should note recurrent technical challenges, such as variants encountered in many samples across a plate (plate-wide variants), that may result in over-calling errors. Plate-wide variants can present as recurrent deletions or duplications and are most likely to occur in HP regions. While the use of high-fidelity DNA polymerase can limit the rate of these false-positive results, the use of variant calling parameters needs to be optimized to distinguish artifacts from real pathogenic variants. In most cases, variation in HP regions in germline testing is located in intronic regions and is not clinically significant. In those cases, whether a false negative or a false positive were to occur, it would likely be classified as benign or likely benign and not clinically significant.

In somatic testing, the *ASXL1* c.1934dupG variant is an example of a variant that can be detected at very low levels in most specimens. When found at high VAF, it is a true biologic and pathogenic mutation and is the most common *ASXL1* mutation in myelodysplastic syndromes[15] and acute myeloid leukemia. As a general rule, *ASXL1* c.1934dupG can be called with confidence at higher VAFs (>10.0%–15.0%), while it cannot easily be distinguished from background noise at low VAFs (<5.0%). *ASXL1* c.1934delG is even more challenging to detect as a true mutation at low levels, because the background noise for the mononucleotide deletion can exceed 5.0% VAF. Across all specimens tested, the distribution of VAFs for variants detected in HP regions tends to be bimodal, with VAFs 5.0% or less representing slipped strand mispairing artifact and VAFs greater than 10.0% to 15.0% representing real pathogenic mutation events (Figure 3, B). Additional support specifically for a true duplication event includes the marked excess of duplications over deletions or the identification of a triplication of the G (owing to artifactual duplication of a variant sequence now containing an extra G).

Table 4. Variability in Nomenclature Used by Laboratories to Describe a Duplication in *PRKAR1A* (NM_002734.4) in the Next-Generation Sequencing-Germline 2018–A Survey

Gene	Transcript	Description	Grade	No. (%) of Labs
PRKAR1A	NM_002734.4	c.349-5dupT	Preferred	42 (50.0)
		c.349-5dupT, p.?	Preferred	2 (2.4)
		c.349-5dup	Acceptable	8 (9.5)
		c.349-5dup, p.?	Acceptable	1 (1.2)
		c.349-5_349-4insT	Unacceptable	10 (11.9)
		c.349-9_349-8insT	Unacceptable	6 (7.2)
		c.349-9insT	Unacceptable	3 (3.6)
		c.349-5_349-4insT, intronic	Unacceptable	1 (1.2)
		c.349-5insT	Unacceptable	1 (1.2)
		c.349-8_349-7insT	Unacceptable	1 (1.2)
		c.349-8_349-9insT	Unacceptable	1 (1.2)
		c.349-8-349-9 insT	Unacceptable	1 (1.2)
		c.349-8insT	Unacceptable	1 (1.2)
		c.-5_-4insT	Unacceptable	1 (1.2)
		c.349-5dupT	Unacceptable	2 (2.4)
		c.349-5dupT	Unacceptable	1 (1.2)

Table 5. Percentage of Proficiency Testing Participants With Specimen Swaps and/or Transcription Errors by Survey and Series

	Mailing							
Proficiency Test	2016-A	2016-B	2017-A	2017-B	2018-A	2018-B	2019-A	2019-B
NGSHM	0.0% (0/57)	0.0% (0/59)	1.2% (1/81)	0.0% (0/87)	2.0% (2/99)	0.0 % (0/101)	1.6% (2/122)	1.4% (2/141)
NGSST	0.0% (0/116)	0.0% (0/120)	0.6% (1/154)	1.2% (2/171)	0.0% (0/188)	0.5% (1/197)	1.0% (2/195)	0.0% (0/205)

Abbreviations: NGSHM, next-generation sequencing hematologic malignancies; NGSST, next-generation sequencing solid tumor.
Numbers are presented in parentheses.

Table 6. Recurrent Proficiency Testing Challenges and Possible Remedies

Challenge	Remedy
Detection of variants in genomic region with high GC content, difficult to sequence	Add 2.5%–5.0% DMSO Add 1.0 M betain Add a heat-denaturation step Consider use of an orthogonal method
Detection of variants from homopolymer regions	Use of a high-fidelity DNA polymerase Optimization of variant calling parameters to distinguish artifacts Use of error correction methods, such as unique molecular identifiers
Missing variants or false positives from pseudogene interference	Align to the hs37d5 reference genome Identify region of homology requiring specific attention Consider long-range PCR and Sanger sequencing
Errors in reporting dinucleotide variants	Manual review of variants Appropriate use of current HGVS nomenclature
Errors in reporting duplication variants as insertions	Manual or bioinformatic review of the raw data Appropriate use of current HGVS nomenclature
Errors due to use of different transcript	Report the transcript and version used
Postanalytic errors due to specimen swaps or transcription errors	Conduct a critical analysis of potential steps that could lead to nonanalytic errors Have a second person check every entry before submission Avoid multiple patient specimens in the active work area at the same time Label only 1 specimen at a time before proceeding to the next specimen Have a second person check the labeling of tubes Consider investigating potential sample swaps with molecular methods

Abbreviations: DMSO, dimethyl sulfoxide; HGVS, Human Genome Variation Society; PCR, polymerase chain reaction.

Sequencing Challenges Due to Pseudogenes

Laboratories may be able to avoid this pitfall by aligning to the hs37d5 reference genome rather than the standard hg19/GRCh37 genome and by identifying regions of homology and critically evaluating variants identified within these regions (Figure 4, B, and Supplemental Table 3). When pseudogene interference is present, germline variants may not have the expected VAF of approximately 50.0% for heterozygous variants or 100.0% for homozygous variants due to loss of reads that were aligned to the pseudogene. In some cases, it may not be possible to evaluate variants in regions of high homology by NGS, and a supplemental or confirmatory method, such as long-range PCR followed by Sanger sequencing, may be required.[16] Error correction methods, such as unique molecular identifiers, may also mitigate these PCR errors.

Annotation Challenges of Dinucleotide Variants

Unfortunately, many variant calling algorithms used in NGS data analyses will detect dinucleotide or trinucleotide variants as multiple individual substitution variants, leading to inaccurate variant representation and reporting, although some bioinformatic solutions that group variants in *cis* are becoming available.[17] This bioinformatic limitation can be circumvented by manual review of the raw data followed by the appropriate use of current HGVS nomenclature before reporting the variant(s) detected.

Annotation Challenges of Duplication Variants

Many laboratories report duplications as insertions, which results in an annotation error, not a detection error. HGVS recommends distinguishing between insertions and duplications with the intention to keep the description simpler, shorter, and unequivocal; this avoids confusion regarding the exact position of the variant. The HGVS recommendation also helps avoid confusion about the origin of duplicating insertions, which is likely DNA polymerase slippage with duplication of a local sequence. Most current variant calling algorithms are designed to detect only nonalignment and, therefore, do not distinguish duplications from insertions. As a result of this failure to appropriately identify duplications, the algorithms correspondingly do not apply the 3′ rule (also known as right versus left alignment, the latter of which is typically used by most variant callers). The combination of these 2 effects makes these types of algorithms noncompliant with current HGVS guidelines. This limitation can be circumvented by manual or bioinformatic review of the raw data followed by verification that the algorithms have followed and applied the current HGVS nomenclature. Tools like Variant Effect Predictor or Mutalyzer can be used to manually annotate variants.[18]

Other Annotation Challenges

To assist laboratories in correct application of nomenclature, a table has been included in the PT kit instructions for the germline NGS survey (Supplemental Table 4). Laboratories should also review the HGVS nomenclature website (https://varnomen.hgvs.org/) and ensure that their pipelines and processes use the most recent recommendations. Also, laboratories should ensure that they are reporting complementary DNA and protein changes along with the version of the transcript used.

Nonanalytic Errors.—In this study, we identified specimen swaps and reporting errors as infrequent but recurring challenges in NGS PT surveys. It is possible, though, that the number of specimen swaps and transcription errors we found does not reflect actual clinical practice. This is because some manual steps are required to report PT results, and this may differ from the laboratory's normal workflow. Human errors can occur because of lack of attention, not following the standard operating procedure, rushing, or performing an infrequent task. Laboratories should conduct a critical analysis of potential steps that could lead to these errors. Having a second person check every PT survey entry before submission could reduce or eliminate transcription errors. Avoiding multiple patient specimens in the active work area at the same time, labeling only one specimen at a time before proceeding to the next specimen, and having a second person check the labeling of tubes are all measures that can prevent specimen swap errors.

It is important to note that although somatic and germline testing surveys had different issues, the approach to PT of these surveys has been different by design. The germline survey was originally developed to test laboratories' overall ability to detect and identify variants in general by NGS, and therefore, in some instances, may include some technically difficult regions that do not have known clinical significance. Conversely, somatic testing surveys are focused on ensuring the ability to detect known, clinically important variants, whether technically challenging or not. Therefore, further technical development of these surveys will likely also reflect this difference, with somatic surveys adding elements to address more technically challenging variants and germline surveys, increasing focus on known clinically relevant variants.

CONCLUSIONS

In conclusion, this study provides a detailed categorization and discussion of recurring challenges found in somatic and germline NGS PT. This study also highlights the importance of PT to identify these challenges so that laboratories can iteratively address and improve their performance. Of note, the overall performance of somatic and germline laboratories on NGS PT surveys was excellent, with the majority of errors related to annotation. With the issues described in this study and the remedies mentioned, laboratories should be able to overcome any annotation and nonanalytic errors to rapidly improve performance. Only a minority of incorrect responses on the surveys were due to actual failures of the sequencing to provide a clear result. These sequencing challenges included known issues with regions of high GC content, HPs, and pseudogenes.

The authors wish to thank Ellen Lazarus, MD, for editorial support.

References

1. Merker JD, Devereaux K, Iafrate AJ, et al. Proficiency testing of standardized samples shows very high interlaboratory agreement for clinical next-generation sequencing-based oncology assays. *Arch Pathol Lab Med*. 2019;143(4):463–471.

2. Surrey LF, Oakley FD, Merker JD, et al. Next-generation sequencing (NGS) methods show superior or equivalent performance to non-NGS methods on *BRAF, EGFR*, and *KRAS* proficiency testing samples. *Arch Pathol Lab Med*. 2019; 143(8):980–984.

3. Keegan A, Bridge JA, Lindeman NI, et al. Proficiency testing of standardized samples shows high interlaboratory agreement for clinical next generation sequencing-based hematologic malignancy assays with survey material-specific differences in variant frequencies. [Published online January 27, 2021]. *Arch Pathol Lab Med.* doi:10.5858/arpa.2019-0352-CP

4. Jennings LJ, Arcila ME, Corless C, et al. Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommenda-

tion of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn*. 2017;19(3):341–365.

5. Grossmann V, Schnittger S, Schindela S, et al. Strategy for robust detection of insertions, deletions, and point mutations in *CEBPA*, a GC-rich content gene, using 454 next-generation deep-sequencing technology. *J Mol Diagn*. 2011;13(2): 129–136. doi:10.1016/j.jmoldx.2010.09.001

6. Viguera E, Canceill D, Ehrlich SD. Replication slippage involves DNA polymerase pausing and dissociation. *EMBO J*. 2001;20(10):2587–2595.

7. Fazekas A, Steeves R, Newmaster S. Improving sequencing quality from PCR products containing long mononucleotide repeats. *Biotechniques*. 2010; 48(4):277–285.

8. Levinson G, Gutman GA. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol*. 1987;4(3):203–221. Review.

9. Yannakou CK, Jones K, McBean M, et al. *ASXL1* c.1934dup; p.Gly646Trpfs*12-a true somatic alteration requiring a new approach. *Blood Cancer J*. 2017;7(12):656.

10. Montes-Moreno S, Routbort MJ, Lohman EJ, et al. Clinical molecular testing for *ASXL1* c.1934dupG p.Gly646fs mutation in hematologic neoplasms in the NGS era. *PLoS One*. 2018;13(9):e0204218. doi: 10.1371/journal.pone.0204218

11. Den Dunnen JT, Dalgleish R, Maglott DR, et al. HGVS recommendations for the description of sequence variants: 2016 update. *Hum Mutat*. 2016;37(6): 564–569. doi:10.1002/humu.22981

12. DNA sequencing of a moderately difficult template: evaluation of results from a Thermus thermophilus unknown test sample. Association of Bromolecular [sic] Resource Facilities (ABRF) DNA Sequence Committee. *Biotechniques*. 1996; 21(4):678.

13. Kieleczawa J. Fundamentals of sequencing of difficult templates–an overview. *J Biomol Tech*. 2006;17(3):207–217.

14. van Dijk EL, Jaszczyszyn Y, Thermes C. Library preparation methods for next-generation sequencing: tone down the bias. *Exp Cell Res*. 2014;322(1):12–20. doi:10.1016/j.yexcr.2014.01.008

15. Gelsi-Boyer V, Trouplin V, Adelaide J, et al. Mutations of polycomb-associated gene *ASXL1* in myelodysplastic syndromes and chronic myelomonocytic leukaemia. *Br J Haematol*. 2009;145(6):788–800. doi: 10.1111/j.1365-2141.2009.07697.x

16. Mandelker D, Schmidt RJ, Ankala A, et al. Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing. *Genet Med*. 2016;18(12):1282–1289. doi:10.1038/gim.2016.58

17. Schmidt RJ, Macleay A, Le LP. VarGrouper: a bioinformatic tool for local haplotyping of deletion-insertion variants from next-generation sequencing data after variant calling. *J Mol Diagn*. 2019;21(3):384–389. doi:10.1016/j.jmoldx.2018.12.003

18. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122. doi:10.1186/s13059-016-0974-4