Consistency and Reproducibility of Next-Generation Sequencing and Other Multigene Mutational Assays: A Worldwide Ring Trial Study on Quantitative Cytological Molecular Reference Specimens

Umberto Malapelle, PhD[1]; Clara Mayo-de-Las-Casas, PhD[2]; Miguel A. Molina-Vila, PhD[2]; Rafael Rosell , MD, PhD[3,4]; Spasenija Savic, MD[5]; Michel Bihl, PhD[5]; Lukas Bubendorf, MD[5]; Manuel Salto-Tellez, MD[6]; Dario de Biase, PhD[7]; Giovanni Tallini, MD[8]; David H. Hwang, MD[9]; Lynette M. Sholl, MD[9]; Rajyalakshmi Luthra, PhD[10]; Birgit Weynand, MD[11]; Sara Vander Borght, Msc, PhD[11]; Edoardo Missiaglia, PhD[12]; Massimo Bongiovanni, MD [12]; Daniel Stieber, PhD[13]; Philippe Vielh, MD[13]; Fernando Schmitt, MD[13]; Alessandra Rappa, PhD[14]; Massimo Barberis, MD[14]; Francesco Pepe, PhD[1]; Pasquale Pisapia, MD[1]; Nicola Serra, PhD[1]; Elena Vigliar, MD[1]; Claudio Bellevicine, MD, PhD [1]; Matteo Fassan, MD, PhD[15]; Massimo Rugge, MD, PhD[15]; Carlos E. de Andrea, PhD[16]; Maria D. Lozano, MD[16]; Fulvio Basolo, MD[17]; Gabriella Fontanini, MD[17]; Yuri E. Nikiforov, MD, PhD[18]; Suzanne Kamel-Reid, PhD[19]; Gilda da Cunha Santos, MD, PhD [19]; Marina N. Nikiforova, MD[18]; Sinchita Roy-Chowdhuri, MD, PhD[10]; and Giancarlo Troncone, MD, PhD [1]; The Molecular Cytopathology Meeting Group

BACKGROUND: Molecular testing of cytological lung cancer specimens includes, beyond epidermal growth factor receptor (*EGFR*), emerging predictive/prognostic genomic biomarkers such as Kirsten rat sarcoma viral oncogene homolog (*KRAS*), neuroblastoma RAS viral [v-ras] oncogene homolog (*NRAS*), B-Raf proto-oncogene, serine/threonine kinase (*BRAF*), and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α (*PIK3CA*). Next-generation sequencing (NGS) and other multigene mutational assays are suitable for cytological specimens, including smears. However, the current literature reflects single-institution studies rather than multicenter experiences. **METHODS:** Quantitative cytological molecular reference slides were produced with cell lines designed to harbor concurrent mutations in the *EGFR, KRAS, NRAS, BRAF*, and *PIK3CA* genes at various allelic ratios, including low allele frequencies (AFs; 1%). This interlaboratory

Corresponding author: Sinchita Roy-Chowdhuri, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030; sroy2@mdanderson.org Corresponding author: Marina N. Nikiforova, Department of Pathology and Laboratory Medicine, University of Pittsburgh Medical Center, 200 Lothrop Street, Room C-606, Pittsburgh, PA 15213; nikiforovamn@upmc.edu. Corresponding author: Giancarlo Troncone, Department of Public Health, University of Naples Federico II, Via Sergio Pansini 5, 80131 Naples, Italy; Fax: (011) 390817463679; giancarlo.troncone@unina.it

[1]Department of Public Health, University of Naples Federico II, Naples, Italy; [2]Laboratory of Oncology, Pangaea Biotech, Barcelona, Spain; [3]Catalan Institute of Oncology, Badalona, Spain; [4]Instituto Oncológico Dr Rosell (IOR), Quirón-Dexeus University Institute, Barcelona, Spain; [5]Institute of Pathology, University Hospital Basel, Basel, Switzerland; [6]Molecular Pathology Programme, Centre for Cancer Research and Cell Biology Queen's University Belfast, Belfast, United Kingdom; [7]Department of Pharmacy and Biotechnology - University of Bologna, Bologna, Italy; [8]Department of Medicine University of Bologna, Bologna, Italy; [9]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts; [10]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas; [11]Department of Pathology, Universitair Ziekenhuis Leuven, Leuven, Belgium; [12]Institute of Pathology, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland; [13]National Health Laboratory, Dudelange, Luxembourg; [14]Division of Pathology, European Institute of Oncology, Milan, Italy; [15]Department of Medicine (DIMED), Surgical Pathology Unit University of Padua, Padua, Italy; [16]Department of Pathology, University Clinic of Navarra, Pamplona, Spain; [17]Department of Surgical, Medical and Molecular Pathology and Critical Area, University of Pisa, Pisa, Italy; [18]Department of Pathology and Laboratory Medicine, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; [19]Department of Laboratory Medicine and Pathobiology, University Health Network, Toronto, Ontario, Canada

We thank Jonathan Frampton (Horizon Diagnostics) for his contribution to the design of the panel of clinically relevant mutations with the University of Naples Federico II.

Additional supporting information may be found in the online version of this article.

Received: January 17, 2017; **Revised:** February 23, 2017; **Accepted:** March 10, 2017

Published online May 5, 2017 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cncy.21868, wileyonlinelibrary.com

ring trial study included 14 institutions across the world that performed multigene mutational assays, from tissue extraction to data analysis, on these reference slides, with each laboratory using its own mutation analysis platform and methodology. **RESULTS:** All laboratories using NGS (n = 11) successfully detected the study's set of mutations with minimal variations in the means and standard errors of variant fractions at dilution points of 10% ($P = .171$) and 5% ($P = .063$) despite the use of different sequencing platforms (Illumina, Ion Torrent/Proton, and Roche). However, when mutations at a low AF of 1% were analyzed, the concordance of the NGS results was low, and this reflected the use of different thresholds for variant calling among the institutions. In contrast, laboratories using matrix-assisted laser desorption/ionization–time of flight (n = 2) showed lower concordance in terms of mutation detection and mutant AF quantification. **CONCLUSIONS:** Quantitative molecular reference slides are a useful tool for monitoring the performance of different multigene mutational assays, and this could lead to better standardization of molecular cytopathology procedures. ***Cancer Cytopathol* 2017;125:615-26.** *© 2017 American Cancer Society.*

KEY WORDS: cytological molecular reference; cytology; lung cancer; molecular cytopathology; multigene mutational assay; next-generation sequencing.

INTRODUCTION

Molecular cytopathology features an increasing number and variety of tests with increasing breadth and depth, and this underscores the effective interplay between genomics and cytology.[1] Molecular cytopathology is especially relevant in lung cancer because most patients are diagnosed with cytological specimens.[2] Aspirated material is typically smeared directly onto glass slides, whereas the residual needle rinse is usually processed as a formalin-fixed, paraffin-embedded (FFPE) cell block preparation. In the absence of an adequate cell block preparation, direct smears are the only source of tumor cells for mutation analysis, and they usually provide excellent-quality DNA.[2] Epidermal growth factor receptor (*EGFR*) testing can be performed successfully with both Diff-Quik–stained smears and Papanicolaou-stained smears.[2] Beyond *EGFR*, emerging predictive/prognostic genomic biomarkers, including Kirsten rat sarcoma viral oncogene homolog (*KRAS*), neuroblastoma RAS viral [v-ras] oncogene homolog (*NRAS*), B-Raf proto-oncogene, serine/threonine kinase (*BRAF*), and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α (*PIK3CA*), are currently being tested for lung cancer patients.[3] Unfortunately, because of the increasing number of biomarkers that are needed for a comprehensive molecular characterization of lung cancer patients, this characterization is frequently difficult to perform with sequential single-gene mutational assays using the small amounts of DNA extracted from a cytological smear. Therefore, it is imperative that multigene mutational assays such as next-generation sequencing (NGS), which can simultaneously screen multiple genes with small amounts of DNA, be able to be performed on cytological smears.[4] Several recent studies have highlighted the use of NGS with cytological smears[2]; however, the available data in the literature reflect only single-institution studies rather than multicenter experiences. Performing interlaboratory ring trials with cytological smears is difficult because smear slides are not reproducible or replaceable. Therefore, a wide interlaboratory comparison of mutation analysis protocols using cytological smears across institutions has not been performed. To fill this knowledge gap, we set up an international multicenter study involving 16 institutions in an interlaboratory ring trial. Each participating laboratory performed a multigene mutational assay that is typically used by the laboratory for routine clinical mutation testing. The mutational assays were performed in each of the clinical laboratories with the usual workflow, from tissue extraction to data analysis, for a set of quantitative cytological molecular reference slides. The reference slides were specifically designed as a surrogate for routine smears to harbor concurrent mutations in the *EGFR, KRAS, NRAS, BRAF*, and *PIK3CA* genes at various allele frequencies (AFs). Results were analyzed with respect to different testing methodologies, platforms, and assay validation parameters to assess the degree of reproducibility of multigene mutational assays.

MATERIALS AND METHODS

Experimental Design Overview

This international ring trial study was designed and coordinated by the University of Naples Federico II (Fed II; Fig. 1). Briefly, Fed II contacted a commercial company (Horizon Diagnostics [HDx], Cambridge, United Kingdom) to produce quantitative molecular reference slides harboring multigene mutations. In a preliminary meeting, Fed II and HDx designed a panel of clinically relevant

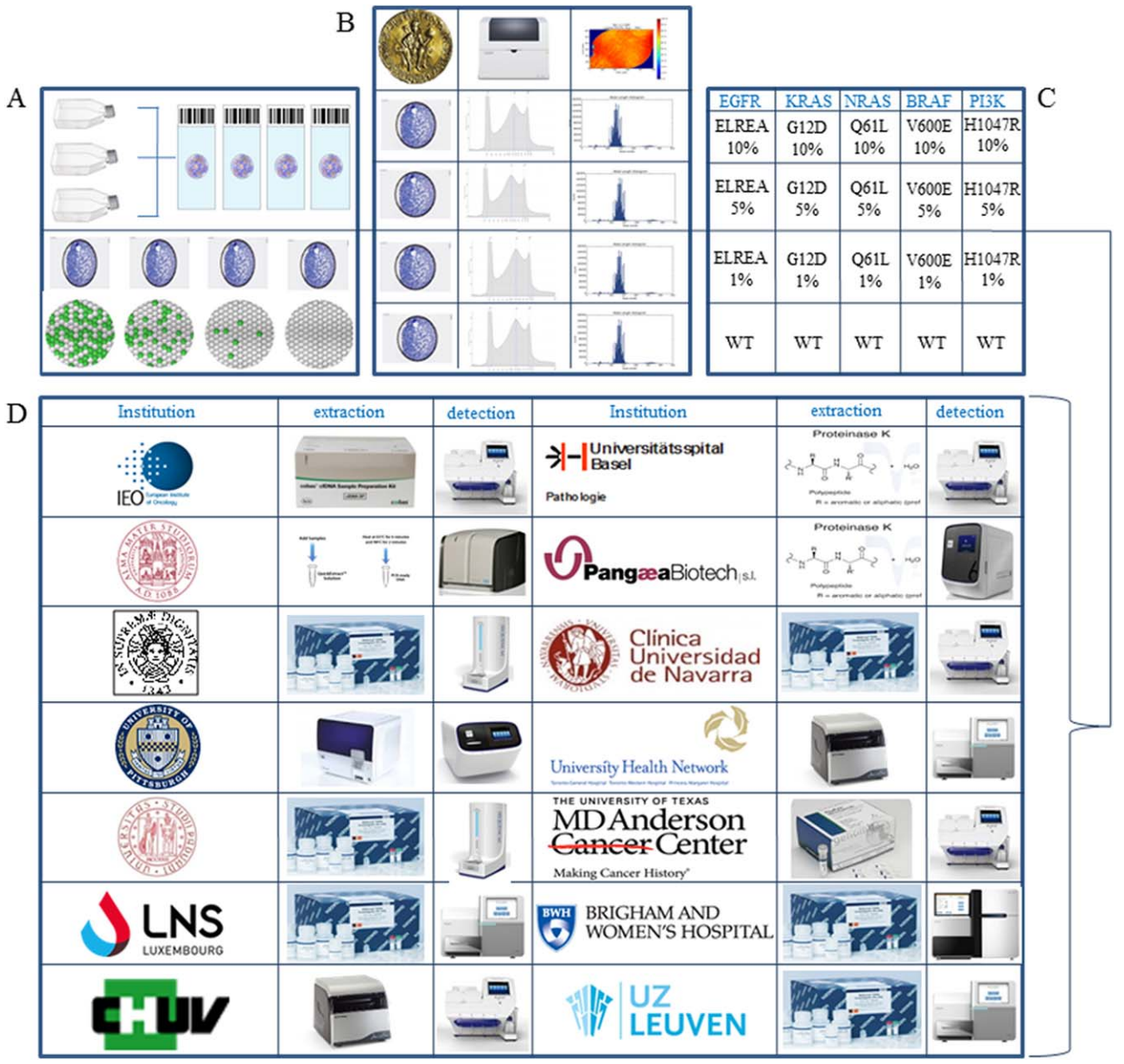

Figure 1. Study design. (A) MCF10A, SW48, and RKO cell lines, genetically modified by Horizon Diagnostics to harbor mutations in *EGFR, KRAS, NRAS, BRAF*, and *PIK3CA*, were used to generate slides harboring mutant alleles at different dilution points (10%, 5%, 1%, and 0%) validated by digital polymerase chain reaction. The study coordinator center at the University of Naples Federico II (B) evaluated the DNA quantity and quality with the 4200 TapeStation system (Agilent) and (C) validated the mutant allele frequency by next-generation sequencing. (D) After validation, sets of unstained slides were distributed to 16 different laboratories, and complete results were obtained from 14 institutions, whose logos, extraction modalities, and platforms are shown. BRAF indicates B-Raf proto-oncogene, serine/threonine kinase; EGFR, epidermal growth factor receptor; KRAS, Kirsten rat sarcoma viral oncogene homolog; NRAS, neuroblastoma RAS viral [v-ras] oncogene homolog; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α; WT, wild type.

mutations, including deletions (eg, *EGFR* c.2235_2249del and p.E746_A750del) and point mutations (eg, *KRAS* c.35G>A p.G12D, *NRAS* c.182A>T p.Q61L, *BRAF* c.1799T>A p.V600E, and *PIK3CA* c.3140A>G p.H1047R). Although HDx had previously developed multigene reference standards in an FFPE format,[5] standard quantitative cytological molecular reference slides were produced for the first time specifically for this study. As a proof of principle, HDx produced prototype slides simultaneously harboring the *BRAF* c1799T>A (p.V600E)

TABLE 1. Multigene Reference Standards (Slides A-C) Harboring Engineered Mutations and Relative AFs Validated by Digital PCR at Horizon Diagnostics and by Ion Torrent NGS at the University of Naples Federico II

Slide	Report	Gene	HGVS[a]	Exon	Type	AF, %			
						Digital PCR	NGS, Unstained	NGS, Stained and Not Coverslipped	NGS, Stained and Coverslipped
A	Mutation	*EGFR*	NM_005228.3 (*EGFR*):c.2235_2249del p.E746_A750del	19	Indel	9.9	12.6	14.9	12.6
A	Mutation	*KRAS*	NM_004985.3 (*KRAS*):c.35G>A p.G12D	2	SNV	10	10.6	10	10.6
A	Mutation	*BRAF*	NM_004333.4 (*BRAF*):c.1799T>A p.V600E	15	SNV	10.38	10.2	11.9	10.2
A	Mutation	*NRAS*	NM_002524.4 (*NRAS*):c.182A>T p.Q61L	3	SNV	9.1	9.7	10.1	9.7
A	Mutation	*PIK3CA*	NM_006218.2 (*PIK3CA*):c.3140A>G p.H1047R	21	SNV	10.3	13.3	10.5	13.3
B	Mutation	*EGFR*	NM_005228.3 (*EGFR*):c.2235_2249del p.E746_A750del	19	Indel	5.2	8.1	5.9	8.1
B	Mutation	*KRAS*	NM_004985.3 (*KRAS*):c.35G>A p.G12D	2	SNV	5.3	5.5	3.8	5.5
B	Mutation	*BRAF*	NM_004333.4 (*BRAF*):c.1799T>A p.V600E	15	SNV	5.5	6.8	5.5	6.8
B	Mutation	*NRAS*	NM_002524.4 (*NRAS*):c.182A>T p.Q61L	3	SNV	4.9	5	4.3	5
B	Mutation	*PIK3CA*	NM_006218.2 (*PIK3CA*):c.3140A>G p.H1047R	21	SNV	5.2	6.7	6.2	6.7
C	Mutation	*EGFR*	NM_005228.3 (*EGFR*):c.2235_2249del p.E746_A750del	19	Indel	1.08	1.5	1.7	1.6
C	Mutation	*KRAS*	NM_004985.3 (*KRAS*):c.35G>A p.G12D	2	SNV	1.07	0.7	1.4	1.3
C	Mutation	*BRAF*	NM_004333.4 (*BRAF*):c.1799T>A p.V600E	15	SNV	1.07	1.4	1	0.8
C	Mutation	*NRAS*	NM_002524.4 (*NRAS*):c.182A>T p.Q61L	3	SNV	1.05	0.8	1.2	1.7
C	Mutation	*PIK3CA*	NM_006218.2 (*PIK3CA*):c.3140A>G p.H1047R	21	SNV	0.88	0.8	1	0.1

Abbreviations: AF, allele frequency; BRAF, B-Raf proto-oncogene, serine/threonine kinase; EGFR, epidermal growth factor receptor; HGVS, Human Genome Variation Society; KRAS, Kirsten rat sarcoma viral oncogene homolog; NGS, next-generation sequencing; NRAS, neuroblastoma RAS viral [v-ras] oncogene homolog; PCR, polymerase chain reaction; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α; SNV, single-nucleotide variant.

[a] All variants are reported according to HGVS guidelines.

TABLE 2. Overview of the Location, Extraction Method, Quantitation Method, and Multigene Mutation Detection Methods for Each Laboratory

Laboratory	Country	Extraction Method	Quantification Method	Platform	Panel
1	Italy	QIAamp Mini kit (Qiagen)	NanoDrop	MALDI-TOF (Sequenom)	Myriapod lung status and colon status panels (Diatech)
2	Belgium	AllPrep DNA/RNA FFPE kit (Qiagen)	Qubit	Illumina MiSeq	TruSight Tumor 26 kit (Illumina)
3	Italy	Cobas DNA sample preparation kit (Roche)	Qubit	Ion Torrent PGM	Oncomine solid tumor panel (Thermo Fisher Scientific)
4	Switzerland	Maxwell 16 FFPE Plus LEV DNA purification kit (Promega)	Qubit	Ion Torrent PGM	AmpliSeq Custom Cancer Hotspot panel (Thermo Fisher Scientific)
5	Luxemburg	QIAamp Mini kit (Qiagen)	Qubit	Illumina MiSeq	TruSight Tumor 15 kit (Illumina)
6	United States	PicoPure	Qubit	Ion Torrent PGM	AmpliSeq Custom Cancer Hotspot panel (Thermo Fisher Scientific)
7	Spain	QIAamp Mini kit (Qiagen)	Qubit	Ion Torrent PGM	AmpliSeq Custom Cancer Hotspot panel (Thermo Fisher Scientific)
8	Canada	Maxwell 16 FFPE Plus LEV DNA purification kit (AS1135; Promega)	Qubit	Illumina MiSeq	TruSight Tumor 26 kit (Illumina)
9	Switzerland	QIAamp Mini kit (Qiagen)	Qubit and NanoDrop	Ion Torrent PGM	Oncomine solid tumor panel (Thermo Fisher Scientific)
10	Italy	QuickExtract (Epicentre)	Qubit	454 GS Junior	Custom assay
11	Italy	QIAamp Mini kit (Qiagen)	NanoDrop 1000	MALDI-TOF (Sequenom)	Myriapod lung status and colon status panels (Diatech)
12	United States	QIAcube (Qiagen)	Qubit	Ion Torrent/Proton	AmpliSeq Custom Cancer Hotspot panel (Thermo Fisher Scientific)
13	Spain	Only digestion without purification		Real-time PCR	Custom assay
14	United States	QIAamp Mini kit (Qiagen)	Qubit	HiSeq 2500	Custom assay (OncoPanel [447-gene panel])

Abbreviations: FFPE, formalin-fixed, paraffin-embedded; LEV, low-elution volume; MALDI-TOF, matrix-assisted laser desorption/ionization–time of flight; PCR, polymerase chain reaction; PGM, Personal Genome Machine.

mutation and the *PIK3CA* c.3140A>G (p.H1047R) mutation. These mutations were correctly detected by Fed II (Supporting File 1 and Supporting Table 1 [see online supporting information]). On the basis of these satisfactory preliminary results, HDx proceeded to manufacture multiple sets of the study slides, as described next. The MCF10A, SW48, and RKO cell lines were genetically modified through adenovirus vectors to harbor mutations in *EGFR, KRAS, NRAS, BRAF*, and *PIK3CA*. The modified cell lines were titrated against a wild-type cell line (MCF10A) to generate mixtures containing quantifiable amounts of mutations. Cells were grown under standard tissue culture conditions; trypsin was used to release the cells, which were counted with the NC100 NucleoCounter (ChemoMetec, Lillerød, Denmark). The cells were then used to generate slides with a validated proprietary method to achieve a fixed number of cells (2×10^6) per slide. Slides A, B, and C were prepared to harbor the previously described mutations at different dilution points (slide A, 10%; slide B, 5%; and slide C, 1%), whereas slide D represented parental wild-type cell lines. To ensure consistency for each slide produced, multigene reference slides were Diff-Quik–stained before scanning with NanoZoomer 3.0 (Hamamatsu, Hamamatsu City, Japan), which indicated a consistent and homogeneous distribution of cells between slides. The mutant AF was validated by HDx with digital polymerase chain reaction (dPCR) and by Fed II with NGS on an Ion Torrent platform (Thermo Fisher Scientific Inc, Waltham, Mass), as described in Supporting File 1 (see online supporting information). Interestingly, the mutation detection and the relative mutant AFs were similar among unstained reference slides, noncoverslipped, Diff-Quik–stained reference slides, and coverslipped, Diff-Quik–stained reference slides (Table 1).

After validation, sets of unstained and air-dried slides (A-D) were distributed to 16 different laboratories; those especially experienced in multigene testing of cytological samples[4,6-19] joined the working group of the annual meeting on molecular cytopathology held in Naples, Italy. Each laboratory conducted mutation testing on the reference slides with the usual multigene assay that it routinely used for testing clinical samples (Table 2). All laboratory

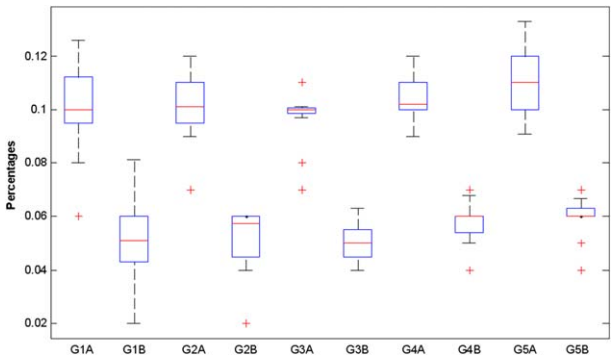

Figure 2. Box plots summarizing allele frequencies observed across laboratories taking part in the ring trial. The bottom and top of each box represent the first and third quartiles of the distribution, respectively, whereas the red line between them indicates the median value. Whiskers include values within 1.5 times the interquartile range. All other values are reported as outliers. Each gene (G) and each slide dilution (A and B) are reported separately: G1 (*EGFR* c.2235_2249del p.E746_A750del), G2 (*KRAS* c.35G>A p.G12D), G3 (*NRAS* c.182A>T p.Q61L), G4 (*BRAF* c.1799T>A p.V600E), and G5 (*PIK3CA* c.3140A>G p.H1047R). BRAF indicates B-Raf proto-oncogene, serine/threonine kinase; EGFR, epidermal growth factor receptor; KRAS, Kirsten rat sarcoma viral oncogene homolog; NRAS, neuroblastoma RAS viral [v-ras] oncogene homolog; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α.

procedures, including internal quality control, were performed for the test slides in a fashion similar to that for any routine clinical samples. For the purpose of this multi-institutional study, the analysis was focused on known hotspot mutations in the *EGFR, KRAS, NRAS, BRAF*, and *PIK3CA* genes. Additional genes harboring endogenous alterations that were detected as a result of using a variety of gene panels interrogating different genes by the various institutions were excluded from this study. Participating institutions provided in 10 days the results for each single slide together with information on the DNA extraction procedure and genotyping methods. The results and the mutant AF data were unblinded and confirmed against HDx specifications.

Statistical Analysis

The statistical analysis was performed with the Matlab statistics toolbox (version 2008; MathWorks, Natick, Mass) for Windows (32-bit). The data are presented as numbers and percentages for categorical variables, and continuous data are expressed as means and standard deviations unless otherwise specified (Supporting Table 2 [see online supporting information]). Levene's test for equality of variances was performed with consideration of laboratory percentages among gene mutations for every set of mutations.

To assess the variations in mutant AF data, we performed the heterogeneity test while taking into account only the mutant allele percentages provided by laboratories using NGS. For this purpose, the I^2 statistic was considered. It represents the percentage of observed total variation across percentages that is due to real heterogeneity rather than chance. It is calculated as follows:

$$I^2 = 100\% \times (Q - df)/Q$$

where Q is Cochran's heterogeneity statistic and df is the degrees of freedom. Particularly negative values of I^2 are put equal to zero so that I^2 lies between 0% and 100%. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity. In addition, box plots were used to show the variability of laboratory percentages for every gene mutation in slides A, B, and C (Fig. 2). Finally, we considered whether there were significant differences between the platforms (Supporting Table 3 [see online supporting information]). There were 2 platforms: Illumina, which was adopted by 4 laboratories (group 1), and the Ion Torrent Personal Genome Machine System, which was adopted by 7 laboratories (group 2). Fed II preliminary validation data (Table 1) were included in terms of mean percentages.

RESULTS

Multigene Sequencing Method Approaches

Mutation analysis results were received from 15 participating laboratories before the cutoff date. One laboratory produced an incomplete data set and was excluded from the analysis. Among the remaining 14 institutions, the most used platform was NGS (n = 11 [79%]), with a majority of these laboratories (n = 6) adopting a semiconductor sequencing–based assay. Five laboratories used the Ion Torrent Personal Genome Machine System (Thermo Fisher Scientific), whereas 1 laboratory used the Ion Proton System (Thermo Fisher Scientific; Fig. 1). All 6 laboratories using Ion Torrent platforms adopted commercial panels validated by Thermo Fisher Scientific, including the Ion AmpliSeq Cancer Hotspot panel (n = 4) and the Oncomine solid tumor panel (n = 2; Table 3). Four laboratories used Illumina platforms (Illumina, Inc, San Diego, Calif): the MiSeq platform was used by 3 institutions in

TABLE 3. Overall Results Obtained for Slides A, B, and C by the 14 Laboratories Joining the Study With Reference to the Platforms and Gene Panels

			Slide A					Slide B					Slide C				
Laboratory	Platform	Panel	*EGFR* c.2235_2249del p.E746_A750del	*KRAS* c.35G>A p.G12D	*NRAS* c.182A>T p.Q61L	*BRAF* c.1799T>A p.V600E	*PIK3CA* c.3140A>G p.H1047R	*EGFR* c.2235_2249del p.E746_A750del	*KRAS* c.35G>A p.G12D	*NRAS* c.182A>T p.Q61L	*BRAF* c.1799T>A p.V600E	*PIK3CA* c.3140A>G p.H1047R	*EGFR* c.2235_2249del p.E746_A750del	*KRAS* c.35G>A p.G12D	*NRAS* c.182A>T p.Q61L	*BRAF* c.1799T>A p.V600E	*PIK3CA* c.3140A>G p.H1047R
1	Sequenom	Myriapod lung and colon panels	20%	20%	15%	8%	13%	12%	12%	ND	ND	ND	ND	ND	ND	ND	ND
2	MiSeq	Illumina TruSight Tumor 26 kit	11%	11%	11%	11%	11%	6%	6%	6%	6%	6%	ND	2%	1%	1%	1%
3	PGM	Oncomine solid tumor panel	9%	12%	10%	10%	11%	4%	6%	4%	5%	6%	1.2%	1.2%	1.3%	1%	ND
4	PGM	AmpliSeq Custom Cancer Hotspot panel	10%	10%	10%	11%	12%	5%	6%	5%	6%	6%	1%	1%	1%	2%	1%
5	MiSeq	Illumina TruSight Tumor 15 kit	11%	10%	10%	10%	10%	4%	5%	5%	5%	5%	ND	1%	1%	2%	ND
6	PGM	AmpliSeq Custom Cancer Hotspot panel	11%	10%	10%	10%	9%	5%	6%	6%	5%	6%	1%	1%	1%	1%	1%
7	PGM	AmpliSeq Custom Cancer Hotspot panel	11%	9%	10%	10%	12%	4%	4%	5%	6%	6%	ND	ND	ND	ND	ND
8	MiSeq	Illumina TruSight Tumor 26 kit	10%	11%	11%	11%	11%	6%	5%	5%	7%	6%	ND	ND	ND	ND	ND
9	PGM	Oncomine solid tumor panel	10%	11%	10%	11%	11%	6%	6%	5%	6%	6%	1%	1%	1%	1%	1%
10	454 GS Junior	Custom assay	10%	9%	7%	9%	10%	8%	4%	4%	4%	7%	2%	1%	2%	ND	ND
11	Sequenom	Myriapod lung and colon panels	20%	22%	20%	10%	20%	20%	10%	ND	10%	ND	ND	ND	ND	ND	ND
12	Proton	AmpliSeq Custom Cancer Hotspot panel	8%	10%	10%	12%	13%	5%	6%	6%	6%	6%	ND	ND	ND	ND	ND
13	RT-PCR	Custom assay	<1%	3%	12%	2%	NC	<1%	2%	4%	1%	NC	<1%	<1%	4%	<1%	NC
14	HiSeq 2500	OncoPanel (447-gene panel)	6%	7%	8%	9%	10%	2%	2%	4%	6%	4%	1%	1%	1%	ND	ND

Abbreviations: BRAF, B-Raf proto-oncogene, serine/threonine kinase; EGFR, epidermal growth factor receptor; KRAS, Kirsten rat sarcoma viral oncogene homolog; NC, not covered; ND, not detected; NRAS, neuroblastoma RAS viral [v-ras] oncogene homolog; PGM, Personal Genome Machine; RT-PCR, reverse transcription–polymerase chain reaction.

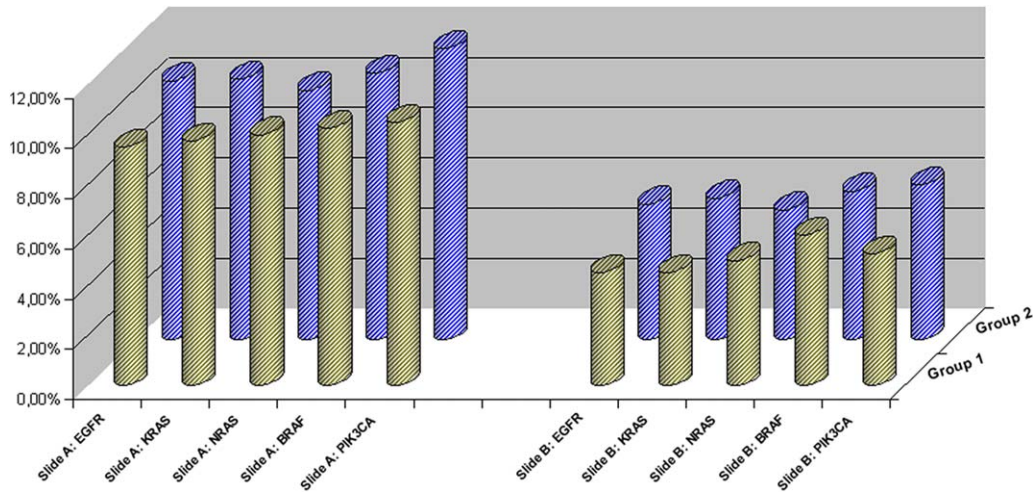

Figure 3. The mean percentage of each mutated gene is graphically reported for slides A and B with respect to the platforms used: Illumina (adopted by 4 laboratories [group 1]) and the Ion Torrent Personal Genome Machine System (adopted by 7 laboratories [group 2]). BRAF indicates B-Raf proto-oncogene, serine/threonine kinase; EGFR, epidermal growth factor receptor; KRAS, Kirsten rat sarcoma viral oncogene homolog; NRAS, neuroblastoma RAS viral [v-ras] oncogene homolog; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit α.

combination with the manufacturer-validated commercial panels, the TruSight Tumor 26 kit (n = 2) and the TruSight Tumor 15 kit (n = 1), whereas 1 institution used the Illumina HiSeq 2500 sequencer in association with the OncoPanel (Table 3).[10] Only 1 institution performed NGS with the 454 GS Junior platform (Roche Diagnostics, Mannheim, Germany; Table 3). Two of the 3 institutions that did not use NGS used matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) on the MassARRAY system (Agena Bioscience, Hamburg, Germany), and both used the Myriapod colon status and lung status panels (Diatech Pharmacogenetics, Jesi, Italy). Finally, 1 laboratory used an in-house–developed multiplex TaqMan reverse transcription–polymerase chain reaction (RT-PCR) assay in the presence of a peptide nucleic acid. This assay, however, covered the 4 hotspot mutations in *EGFR, KRAS, NRAS*, and *BRAF* but did not cover the *PIK3CA* c.3140A>G p.H1047R mutation.

Testing Performance

Slide A

All laboratories successfully detected the 5 engineered mutations on slide A with the exception of 1 laboratory using a multiplex TaqMan RT-PCR assay that did not cover the *PIK3CA* hotspot mutation (laboratory 13; see the Materials and Methods section). Notably, all the laboratories using NGS yielded the expected AF data (10%) in agreement ($P = .171$) with the HDx dPCR assessment (Fig. 2 and Supporting Table 2 [see online supporting information]). In fact, the mean AFs plus or minus the standard deviation for each mutation by NGS were 10.03% ± 1.67% (*EGFR* c.2235_2249del p.E746_A750del), 10.27% ± 1.24% (*KRAS* c.35G>A p.G12D), 9.73% ± 1.09% (*NRAS* c.182A>T p.Q61L), 10.37% ± 0.84% (*BRAF* c.1799T>A p.V600E), and 11.11% ± 1.22% (*PIK3CA* c.3140A>G p.H1047R).

Subsequently, we compared the laboratories using Illumina (group 1) and Ion Torrent methodologies (group 2; Fig. 3 and Supporting Table 3 [see online supporting information]). There were not significant differences in the mean percentages between group 1 and group 2 for all gene mutations. In contrast, although the 2 laboratories using MALDI-TOF (laboratories 1 and 11) were able to detect all 5 mutations, this methodology tended to overestimate the mutant AF in the *EGFR, KRAS, NRAS*, and *PIK3CA* genes (20% [*EGFR* c.2235_2249del p.E746_A750del], 21% [*KRAS* c.35G>A p.G12D], 15% [*NRAS* c.182A>T p.Q61L], and 13% [p*PIK3CA* c.3140A>G p.H1047R]) while underestimating the mutant AF for *BRAF* c.1799T>A p.V600E (8%; Table 3).

Slide B

Although there was a general consensus among the different institutions for slide A, the results for slide B highlight differences resulting from the various analytical sensitivities of the different platforms. All laboratories using NGS detected the 5 engineered mutations on slide B with the

expected mutant AF data (5%) in concordance ($P = .063$) with the HDx dPCR assessment (Fig. 2 and Supporting Table 2 [see online supporting information]). In fact, the mean AFs plus or minus the standard deviation for each mutation by NGS were 5.33% ± 1.62% (*EGFR* c.2235_2249del p.E746_A750del), 5.13% ± 1.19% (*KRAS* c.35G>A p.G12D), 5.02% ± 0.74% (*NRAS* c.182A>T p.Q61L), 5.80% ± 0.77% (*BRAF* c.1799T>A p.V600E), and 5.95% ± 0.76% (*PIK3CA* c.3140A>G p.H1047R). No differences in mutant AFs were observed among laboratories using Illumina (group 1) and Ion Torrent methodologies (group 2; Fig. 3). Only *PIK3CA* c.3140A>G p.H1047R slightly differed ($P = .0383$) between Illumina (5.25% ± 0.83%) and Ion Torrent (6.19% ± 0.30%; Supporting Table 3 [see online supporting information]). In contrast, the 2 laboratories using MALDI-TOF (laboratories 1 and 11) missed the *NRAS* c.182A>T p.Q61L and *PIK3CA* c.3140A>G p.H1047R mutations in slide B. In addition, laboratory 1 also missed *BRAF* c.1799T>A p.V600E. In terms of estimating mutant AFs, MALDI-TOF overestimated the AFs for all genes: *EGFR* c.2235_2249del p.E746_A750delELREA was detected at a 12% AF by laboratory 1 and at a 20% AF by laboratory 11, *KRAS* c.35G>A p.G12D was detected at a 12% AF by laboratory 1 and at a 10% AF by laboratory 11, and *BRAF* c.1799T>A p.V600E was detected at a 10% AF by laboratory 11. The multiplex TaqMan RT-PCR assay (laboratory 13) detected all covered engineered mutations on slide B (Table 3).

Slide C

Although for slide B there was a general agreement between the different institutions using NGS, the results derived for slide C underscored significant differences not only in the platform analytic sensitivity but also in variant calling modalities. Three laboratories using NGS (laboratories 7, 8, and 12) were unable to detect any mutation because their validated bioinformatic pipeline had a variant calling threshold set at 5%. The remaining laboratories were able to detect low-level mutations, although it was below the reportable threshold of variant calls. Three laboratories (laboratories 5, 10, and 14) were able to detect 3 of the 5 mutations, whereas 2 laboratories (laboratories 2 and 3) detected 4 of the 5 mutations (Table 3). Only 3 laboratories (laboratories 4, 6, and 9) successfully detected all 5 engineered mutations (Table 3). The mean AFs plus or minus the standard deviation for each mutation by NGS were 1.24% ± 0.35% (*EGFR* c.2235_2249del p.E746_A750del), 1.10% ± 0.34% (*KRAS* c.35G>A p.G12D), 1.12% ± 0.33% (*NRAS* c.182A>T p.Q61L), 1.34% ± 0.44% (*BRAF* c.1799T>A p.V600E), and 0.96% ± 0.08% (*PIK3CA* c.3140A>G p.H1047R). Thus, all mutations detected by NGS yielded the expected mutant AF data (1%) in agreement with the HDx dPCR assessment. A comparison of group 1 and group 2 was not performed for slide C because of the limited available data (Fig. 3 and Supporting Table 3 [see online supporting information]). The 2 laboratories using MALDI-TOF (laboratories 1 and 11) failed to identify any of the mutations, but they were detected by an orthogonal real-time polymerase chain reaction assay in laboratory 1 (data not shown), and this emphasizes the inherent lower sensitivity of the MALDI-TOF assay in comparison with the NGS platform. In contrast, the multiplex TaqMan RT-PCR assay (laboratory 13) detected all covered engineered mutations on slide C (Table 3). Because AFs were in significant discordance with the HDx dPCR assessment, after the cutoff data, laboratory 13 reanalyzed the residual DNA with respect to slides A to D with the GeneReader NGS platform (Qiagen, Valencia, Calif), and the results were similar to those revealed by the manufacturer and the other NGS platforms (Supporting Table 4 [see online supporting information]).

Slide D

All laboratories correctly genotyped slide D as the wild type, and there were no false-positive results. In addition, 13 of the 14 laboratories reported additional mutations in slides A to C because of the genetic background of the employed cell lines (available on request). For example, *EGFR* c.2155G>A p.G719S was detected with consistency (data not shown); however, because these endogenous mutations had not been validated by HDx, these mutations were not included in the assay comparisons.

DISCUSSION

In advanced lung cancer, the appropriate administration of targeted treatments or the eligibility for clinical trials requires tumor multigene mutational testing.[14] This should provide timely, accurate, and reliable results, even with small tissue specimens, including challenging cytological samples. Differences in gene panels, target enrichment strategies, and platforms are potential sources of

inconsistent results, and methods to monitor the accuracy of multigene mutational testing are crucial.[20]

To date, only histological samples have been included in proficiency testing panels because the quality of molecular testing with cytology is difficult to assess.[21] Ideally, interlaboratory ring trials should use routine clinical samples and evaluate both the pre-analytical parameters and the genotyping analysis.[22] FFPE sections from histological tissue blocks that can be distributed among laboratories make accurate interlaboratory comparisons possible. In contrast, cytological smears are unique specimen preparations that cannot be reproduced or replaced once they have been sacrificed for molecular testing. This inherent difficulty in using direct smears for testing, together with the need for additional validation, precludes the use of smears for mutation analysis in a large number of laboratories performing molecular assays. Some institutions digitally scan cytological smears before molecular testing to archive the cytomorphology of representative diagnostic microscopic fields to mitigate medicolegal constraints.[15,23,24] Moreover, the variety of cytological specimen preparations (eg, smears, cytospins, liquid-based preparations, and cell block preparations) and the heterogeneous mutational profile of clinical specimens with concurrent mutations present in various AFs and/or in subclonal populations of tumor cells represent significant challenges when it comes to using archival cytological smears for quality assurance purposes.[23]

Conversely, to compare the performances of different methodologies, identical routine pre-analytical and genotyping procedures need to be replicated with standardized artificial reference standards. Certainly, control material should be highly reproducible and mimic clinical specimens as closely as possible.[25] To this end, previous studies have validated engineered cell lines in paraffin-embedded curls and sections.[5,26,27] In this study, to mirror a cytological smear, for the first time, we have developed and validated a cytological molecular reference slide set. A panel of clinically relevant mutations in lung cancer, including both deletion and point mutations (ie, mutations in *EGFR* as well as emerging predictive/prognostic mutational biomarkers such as *KRAS, NRAS, BRAF*, and *PIK3CA*) was engineered into cell lines specifically for this study. As a proof of principle, prototype slides were preliminarily tested; a fixed number of cells (2×10^6) were prepared for each slide, and this led to a homogeneous distribution of cells, as observed on Diff-Quik–stained slides (Fig. 1).

In all institutions joining this study, the multigene mutational methods were robust and accurate. Our results indicate that the performance of the NGS assays across different laboratories was excellent, regardless of the platform or gene panel used, with concordance for all 5 engineered mutations at 10% and 5% dilution points. In addition, NGS assays were able to accurately estimate the mutant AF, and this was consistent with the dPCR assessment performed at the source. In contrast, both laboratories adopting MALDI-TOF reported false-negative results at the 5% dilution point and a less than accurate estimation of the mutant AF. Notably, the performance of NGS among institutions in the low mutant AF setting (1%; slide C) showed lower concordance because of the use of different thresholds for variant calling, with some laboratories relying solely on the variant calls made by the bioinformatic pipelines and the set threshold, whereas others supplemented the variant calling with visual inspections of the sequencing reads. Our data showed that laboratory performance was not influenced by the types of mutations (deletions vs point mutations). The publication of recent guidelines establishing consensus standards for somatic mutation diagnostic testing, specifically for identifying and reporting mutations in solid tumors, may lead to a more uniform approach in reporting mutations below the threshold of 5%.[28,29]

Additional studies are needed by the Molecular Cytopathology Meeting Group to optimize quantitative cytological molecular reference slides that will better represent actual clinical cytological samples. Although Diff-Quik– or Papanicolaou-stained smears better reflect clinical routine samples, the mutation detection and relative mutant AFs were similar among unstained reference slides, noncoverslipped, Diff-Quik–stained reference slides, and coverslipped, Diff-Quik–stained reference slides in our preliminary validation study (Table 1). Finally, a follow-up ring trial study is currently underway by the group to optimize and compare results with less cellular reference slides to simulate real-life molecular testing conditions for paucicellular smears.[30-32]

FUNDING SUPPORT

This study was funded by the Department of Public Health of the University of Naples Federico II.

CONFLICT OF INTEREST DISCLOSURES

Lukas Bubendorf reports personal fees from Roche and AstraZeneca outside the submitted work. Yuri Nikiforov reports personal fees from Quest Diagnostics outside the submitted work.

AUTHOR CONTRIBUTIONS

Umberto Malapelle: Conceptualization, writing–original draft, writing–review and editing, visualization, and supervision. **Clara Mayo-de-Las-Casas**: Methodology, investigation, and resources. **Miguel A. Molina**: Methodology, investigation, and resources. **Rafael Rosell**: Supervision. **Spasenija Savic**: Validation, investigation, resources, and writing–review and editing. **Michel Bihl**: Methodology, validation, and resources. **Lukas Bubendorf**: Validation. **Manuel Salto-Tellez**: Conceptualization and validation. **Dario de Biase**: Methodology, investigation, and resources. **Giovanni Tallini**: Validation. **David H. Hwang**: Validation and writing–review and editing. **Lynette M. Sholl**: Methodology, investigation, and resources. **Rajyalakshmi Luthra**: Methodology, investigation, and resources. **Sinchita Roy-Chowdhuri**: Conceptualization, validation, writing–review and editing, and supervision. **Birgit Weynand**: Methodology and writing–review and editing. **Sara Vander Borght**: Methodology, investigation, and resources. **Edoardo Missiaglia**: Methodology, investigation, and resources. **Massimo Bongiovanni**: Validation. **Daniel Stieber**: Methodology, investigation, and resources. **Philippe Vielh**: Conceptualization. **Fernando Schmitt**: Conceptualization. **Alessandra Rappa**: Methodology, investigation, and resources. **Massimo Barberis**: Conceptualization. **Francesco Pepe**: Methodology, investigation, resources, and data curation. **Pasquale Pisapia**: Methodology, investigation, resources, data curation, writing–original draft, and visualization. **Nicola Serra**: Investigation, resources, software, and formal analysis. **Elena Vigliar**: Methodology, investigation, resources, and data curation. **Claudio Bellevicine**: Methodology, investigation, and resources. **Matteo Fassan**: Methodology, investigation, and resources. **Massimo Rugge**: Supervision. **Carlos E. de Andrea**: Methodology, investigation, and resources. **Maria D. Lozano**: Conceptualization, validation, and writing–review and editing. **Fulvio Basolo**: Methodology, validation, and writing–review and editing. **Gabriella Fontanini**: Conceptualization, validation, and writing–review and editing. **Marina N. Nikiforova**: Conceptualization, validation, writing–review and editing, and supervision. **Yuri E. Nikiforov**: Validation, writing–review and editing, and supervision. **Suzanne Kamel-Reid**: Methodology, investigation, and resources. **Gilda da Cunha Santos**: Conceptualization. **Giancarlo Troncone**: Conceptualization, writing–original draft, writing–review and editing, visualization, project administration, funding acquisition, and supervision.

REFERENCES

1. Bellevicine C, Vita GD, Malapelle U, Troncone G. Applications and limitations of oncogene mutation testing in clinical cytopathology. *Semin Diagn Pathol.* 2013;30:284-297.
2. Roy-Chowdhuri S, Aisner DL, Allen TC, et al. Biomarker Testing in Lung Carcinoma Cytology Specimens: A Perspective From Members of the Pulmonary Pathology Society. *Arch Pathol Lab Med*. 2016. [Epub ahead of print]
3. Sholl LM, Aisner DL, Varella-Garcia M, et al. Multi-institutional oncogenic driver mutation analysis in lung adenocarcinoma: the Lung Cancer Mutation Consortium experience. *J Thorac Oncol.* 2015;10:768-777.
4. Vigliar E, Malapelle U, de Luca C, Bellevicine C, Troncone G. Challenges and opportunities of next-generation sequencing: a cytopathologist's perspective. *Cytopathology.* 2015.26:271-283.
5. Dijkstra JR, Opdam FJ, Boonyaratanakornkit J, et al. Implementation of formalin-fixed, paraffin-embedded cell line pellets as high-quality process controls in quality assessment programs for KRAS mutation analysis. *J Mol Diagn.* 2012;14:187-191.
6. Casadio C, Guarize J, Donghi S, et al. Molecular testing for targeted therapy in advanced non–small cell lung cancer: suitability of endobronchial ultrasound transbronchial needle aspiration. *Am J Clin Pathol.* 2015;144:629-634.
7. Bongiovanni M, Molinari F, Eszlinger M, et al. Laser capture microdissection is a valuable tool in the preoperative molecular screening of follicular lesions of the thyroid: an institutional experience. *Cytopathology.* 2015;26:288-296.
8. Fassina A, Cappellesso R, Olivotto A, Fassan M. MicroRNA analysis as a potential diagnostic tool in the cytological diagnosis of non–small cell lung cancer. *J Clin Oncol.* 2011;29(15 suppl):10605.
9. Giannini R, Lupi C, Sensi E, et al. EGFR and KRAS mutational analysis in a large series of Italian non–small cell lung cancer patients: 2,387 cases from a single center. *Oncol Rep.* 2016;36: 1166-1172.
10. Sholl LM, Do K, Shivdasani P, et al. Institutional implementation of clinical tumor profiling on an unselected cancer population. *JCI Insight.* 2016;1:e87062.
11. Shiau CJ, Babwah JP, da Cunha Santos G, et al. Sample features associated with success rates in population-based EGFR mutation testing. *J Thorac Oncol.* 2014;9:947-956.
12. Molina-Vila MA, Bertran-Alamillo J, Reguart N, et al. A sensitive method for detecting EGFR mutations in non–small cell lung cancer samples with few tumor cells. *J Thorac Oncol.* 2008;3:1224-1235.
13. Lozano MD, Zulueta JJ, Echeveste JI, et al. Assessment of epidermal growth factor receptor and K-ras mutation status in cytological stained smears of non–small cell lung cancer patients: correlation with clinical outcomes. *Oncologist.* 2011;16:877-885.
14. Dacic S, Nikiforova MN. Present and future molecular testing of lung carcinoma. *Adv Anat Pathol.* 2014;21:94-99.
15. Roy-Chowdhuri S, Stewart J. Preanalytic Variables in Cytology: Lessons Learned From Next-Generation Sequencing-The MD

Anderson Experience. *Arch Pathol Lab Med*. 2016. [Epub ahead of print]
16. Savic S, Tapia C, Grilli B, et al. Comprehensive epidermal growth factor receptor gene analysis from cytological specimens of non–small-cell lung cancers. *Br J Cancer.* 2008;98:154-160.
17. Schmitt FC. Molecular cytopathology and flow cytometry: preanalytical procedures matter. *Cytopathology.* 2011;22:355-357.
18. de Biase D, Visani M, Malapelle U, et al. Next-generation sequencing of lung cancer EGFR exons 18-21 allows effective molecular diagnosis of small routine samples (cytology and biopsy). *PLoS One.* 2013;8:e83607.
19. Dooms C, Weynand B, Vander Borght S, et al. Molecular diagnostics on tissue samples obtained through EBUS-TBNA: review on practice guidelines. *BJMO.* 2016;10:15-20.
20. Gargis AS, Kalman L, Berry MW, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat Biotechnol.* 2012;30:1033-1036.
21. Clark DP. Seize the opportunity: underutilization of fine-needle aspiration biopsy to inform targeted cancer therapy decisions. *Cancer Cytopathol.* 2009;117:289-297.
22. Tembuyser L, Dequeker EM. Endorsing good quality assurance practices in molecular pathology: risks and recommendations for diagnostic laboratories and external quality assessment providers. *Virchows Arch.* 2016;468:31-41.
23. Killian JK, Walker RL, Suuriniemi M, et al. Archival fine-needle aspiration cytopathology (FNAC) samples: untapped resource for clinical molecular profiling. *J Mol Diagn.* 2010;12:739-745.
24. van Eijk R, Licht J, Schrumpf M, et al. Rapid KRAS, EGFR, BRAF and PIK3CA mutation analysis of fine needle aspirates from non–small-cell lung cancer using allele-specific qPCR. *PLoS One.* 2011;6:e17791.
25. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. *Clin Biochem Rev.* 2007;28:139-147.
26. Kapp JR, Diss T, Spicer J, et al. Variation in pre-PCR processing of FFPE samples leads to discrepancies in BRAF and EGFR mutation detection: a diagnostic RING trial. *J Clin Pathol.* 2015;68:111-118.
27. Patton S, Normanno N, Blackhall F, et al. Assessing standardization of molecular testing for non–small-cell lung cancer: results of a worldwide external quality assessment (EQA) scheme for EGFR mutation testing. *Br J Cancer.* 2014;111:413-420.
28. Deans ZC, Costa JL, Cree I, et al. Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours; an expert opinion on behalf of IQN Path ASBL. *Virchows Arch.* 2017;470:5-20.
29. Aziz N, Zhao Q, Bry L, et al. College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. *Arch Pathol Lab Med.* 2015;139:481-493.
30. Roy-Chowdhuri S, Goswami RS, Chen H, et al. Factors affecting the success of next-generation sequencing in cytology specimens. *Cancer Cytopathol.* 2015;123:659-668.
31. Bellevicine C, Sgariglia R, Malapelle U, et al. Young investigator challenge: can the Ion AmpliSeq Cancer Hotspot v2 panel be used for next-generation sequencing of thyroid FNA samples? *Cancer.* 2016;124:776-784.
32. Piqueret-Stephan L, Marcaillou C, Reyes C, et al. Massively parallel DNA sequencing from routinely processed cytological smears. *Cancer Cytopathol.* 2016;124:241-253.