Clinical Chemistry 62:5
743–754 (2016)

Optimization and Standardization of Circulating MicroRNA Detection for Clinical Application: The miR-Test Case

Matteo Jacopo Marzi,[1†] Francesca Montani,[2†] Rose Mary Carletti,[2,3] Fabio Dezi,[2] Elisa Dama,[2,4] Giuseppina Bonizzi,[2] Maria Teresa Sandri,[5] Cristiano Rampinelli,[6] Massimo Bellomi,[6] Patrick Maisonneuve,[4] Lorenzo Spaggiari,[7] Giulia Veronesi,[7,8] Fabrizio Bianchi,[2] Pier Paolo Di Fiore,[2,3,9*‡] and Francesco Nicassio[1,2*‡]

BACKGROUND: The identification of circulating microRNAs (miRNAs) in the blood has been recently exploited for the development of minimally invasive tests for the early detection of cancer. Nevertheless, the clinical transferability of such tests is uncertain due to still-insufficient standardization and optimization of methods to detect circulating miRNAs in the clinical setting.

METHODS: We performed a series of tests to optimize the quantification of serum miRNAs that compose the miR-Test, a signature for lung cancer early detection, and systematically analyzed variables that could affect the performance of the test. We took advantage of a large-scale (>1000 samples) validation study of the miR-Test that we recently published, to evaluate, in clinical samples, the effects of analytical and preanalytical variables on the quantification of circulating miRNAs and the clinical output of the signature (risk score).

RESULTS: We developed a streamlined and standardized pipeline for the processing of clinical serum samples that allows the isolation and analysis of circulating miRNAs by quantitative reverse-transcription PCR, with a throughput compatible with screening trials. The major source of analytical variation came from RNA isolation from serum, which could be corrected by use of external (spike-in) or endogenous miRNAs as a reference for normalization. We also introduced standard operating procedures and QC steps to check for unspecific fluctuations that arise from the lack of standardized criteria in the collection or handling of the samples (preanalytical factors).

CONCLUSIONS: We propose our methodology as a reference for the development of clinical-grade blood tests on the basis of miRNA detection.

MicroRNAs (miRNAs)[10] are a class of small noncoding RNAs (18–24 nt long) that function in posttranscriptional regulation of gene expression *(1, 2)*. The observation that miRNA expression is frequently deregulated in diseases points to their potential utility in diagnostic and prognostic applications *(3)*. miRNAs are abundant in serum and plasma *(4)*, where they are shielded from harsh conditions and endogenous ribonucleases (RNases) by various means *(5, 6, 7)*. Increasing evidence has confirmed that miRNAs exist in almost every biological fluid *(8)* and that signatures of circulating miRNAs with diagnostic potential can be identified for many diseases, including cancer *(8, 9)*. Therefore, circulating miRNAs have been proposed as useful biomarkers to improve risk assessment, diagnosis, prognosis, and monitoring of therapy response.

Although many circulating miRNA signatures have been reported, their clinical transferability remains uncertain because of insufficient standardization of detection methods in the clinical setting *(10–12)*. The issue is exacerbated by the lack of systematic analyses of the possible confounding variables that might affect the fluctuations of circulating miRNAs and, hence, performance of the signatures. We previously described a 34-miRNA sig-

[1] Center for Genomic Science of IIT@SEMM, Istituto Italiano di Tecnologia, Milan, Italy; [2] Molecular Medicine Program, European Institute of Oncology, Milan, Italy; [3] The FIRC Institute for Molecular Oncology Foundation (IFOM), Milan, Italy; [4] Division of Epidemiology and Biostatistics, [5] Division of Laboratory Medicine, [6] Division of Radiology, and [7] Division of Thoracic Surgery, European Institute of Oncology, Milan, Italy; [8] Humanitas Research Hospital, Thoracic Surgery, Rozzano, Milan, Italy; [9] Dipartimento di Scienze della Salute, University of Milan, Milan, Italy.

* Address correspondence to: F.N. at Campus IFOM-IEO, Via Adamello 16, 20139, Milan, Italy. Fax +39-0294375991; e-mail francesco.nicassio@iit.it. P.P.DiF. at European Institute of Oncology, Molecular Medicine Program, Via Ripamonti, 435, 20141 Milan, Italy. Fax +39-0294375991; e-mail pierpaolo.difiore@ieo.eu.

† M.J. Marzi and F. Montani contributed equally to this work.

‡ P.P. Di Fiore and F. Nicassio contributed equally to this work.

Received November 25, 2015; accepted February 19, 2016.
Previously published online at DOI: 10.1373/clinchem.2015.251942

[10] Nonstandard abbreviations: miRNA, microRNA; RNase, ribonuclease; COSMOS, Continuous Observation of Smoking Subjects; SOP, standard operating procedure; RBC, red blood cell; *A*, absorbance (*A* 414, absorbance 414 nm); Cq, quantification cycle; RT-qPCR, reverse transcription quantitative PCR; cTLDA, custom Taqman low-density array; HK, housekeeping; RT-AMP, reverse transcription with amplification; LDCT, low-dose computed tomography.

nature capable of detecting asymptomatic non–small cell lung cancer in a screening cohort of high-risk individuals from the Continuous Observation of Smoking Subjects (COSMOS) screening program for lung cancer *(13)*. Recently, we performed a large-scale validation study of this signature, reducing the number of biomarkers to 13 miRNAs plus 6 reference miRNAs, which compose the miR-Test, and validated the performance of the reduced signature in 1115 individuals *(14)*.

Here, we provide a detailed account of the pipeline, tests, and improvements that we adopted to optimize and standardize the detection of miRNAs composing the miR-Test in view of clinical applications. We propose our methodology as a reference for the development of clinical-grade blood tests on the basis of miRNA detection.

Materials and Methods

Patient selection criteria and procedures used for miR-Test data analysis were previously described *(14)*. Full details of the standard operating procedures (SOPs), including blood collection, RNA isolation, and miRNA detection, are provided in the Supplemental File, which accompanies the online version of this article at http://www.clinchem.org/content/vol62/issue5.

QUANTITATIVE REAL-TIME PCR EXPERIMENTS

Raw and normalized data of clinical samples used in this study are directly available at NCBI GEO (GSE76462). The information provided in this paper and previous publications *(13, 14)* provide the essential requirements of the minimum information for publication of quantitative real-time PCR experiments, compatibly with the nature of the samples analyzed and the commercial probe used, allowing complete reproducibility of this study.

PREPARATION OF HEMOLYZED SAMPLES

Two tubes of blood from the same healthy individual were collected by standard phlebotomy. One blood tube was left to clot and was centrifuged to prepare the serum. The other blood tube was used to prepare red blood cell (RBC) lysate. RBCs were separated from plasma by centrifugation (500*g*, 10 min at 4 °C, according to Hanson et al. *(15)*), resuspended in nonhemolyzed serum from the same healthy individual, and artificially hemolyzed by pipetting 5 times with a syringe (21G needle, 0.8 × 20 mm). The 2 resulting sera (nonhemolyzed and RBC lysate enriched) were mixed at different ratios to produce different hemolytic conditions. Samples were left for 3 h at room temperature to simulate the same conditions of clotting used above. Hemolysis was measured by use of different spectrophotometric methods: lipemia-independent measurement of hemolysis [hemolysis index *(16)*], a method based on the absorption spectrum of oxyhemoglobin [absorbance (*A*) 414 nm] and clinical measurement by a laboratory analyzer (Abbott c8000, Abbott Diagnostics).

STATISTICAL ANALYSIS

Tabular data and statistical analyses (1-way, scatterplot) were produced by use of JMP 10 (SAS) software. Microsoft Excel was used to generate bar and linear graphs of results plotted as the mean and SD from repeated experiments. The number of replicates and the statistical tests used are indicated in the figure legends. Heat maps were generated by Java TreeView software (http://jtreeview.sourceforge.net).

Results

To develop a miRNA-based test (such as the miR-Test) applicable to the clinical setting, we implemented a pipeline for the analysis of circulating miRNAs (Fig. 1A), which includes an optimized and semiautomated RNA extraction protocol to improve sensitivity and reproducibility while reducing the amount of manual operations. We investigated several parameters, such as column type, volume of serum, and volume of aqueous phase, in repeated extractions by assessing the recovery of circulating miRNAs from the same volume of input RNA. To this end, we quantified 2 endogenous circulating miRNAs (miR-16 and miR-19b) and 1 exogenous miRNA (hsa-miR-34a, barely detectable in the human sera used for this study), which had been spiked in at a fixed amount during the extraction protocol.

Initially, we tested the performance of binding columns for miRNA isolation, comparing 2 major providers [miRVANA (Ambion) and miRNeasy (Qiagen)]. The miRNeasy column displayed better performance, with higher yields [lower quantification cycle (Cq)] and reproducibility (less variation between repeated extractions) (Fig. 1B). Moreover, RNA extraction with miRNeasy could also be successfully semiautomated with the QIAcube (Qiagen), a robotic workstation that eliminates most of the manual steps (Fig. 1B). Therefore, we chose the miRNeasy-QIAcube platform for the miRNA isolation step and proceeded with the protocol optimization by establishing the appropriate volume of serum or aqueous phase to be used. By measuring the recovery of the spiked-in miR-34a, we observed a constant performance up to 0.4 mL of serum, which decreased thereafter (Fig. 1C). Similarly, the recovery of endogenous miRNAs increased almost linearly with increasing serum volume up to 0.4 mL, whereas no further increase was observed at 0.5 mL (Fig. 1C), suggesting that excessive contaminants introduced by increasing serum volumes interfere with the purification process. Because the maximal volume of serum that can be processed in a single 2-mL tube is 0.3 mL, we selected this amount of serum for subsequent

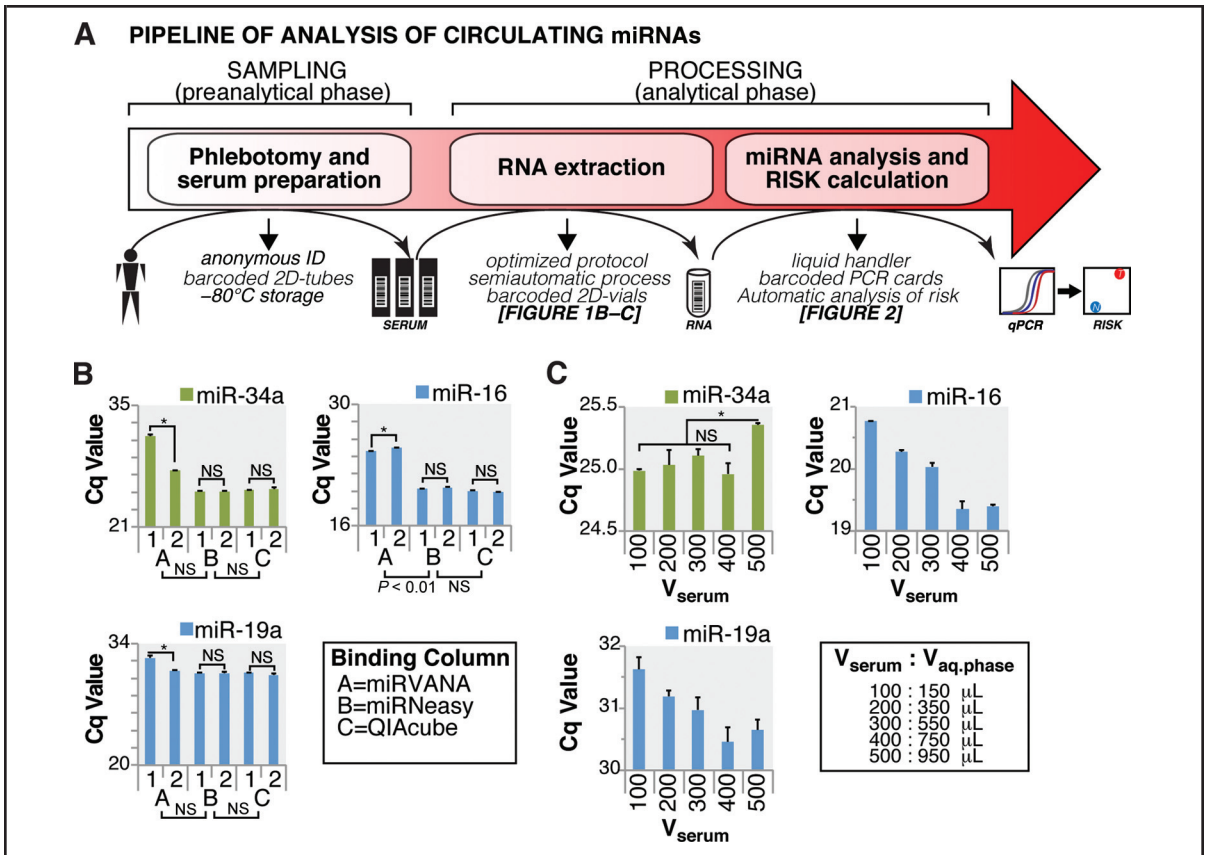

Fig. 1. Optimization of the isolation of circulating miRNAs.

(A), Pipeline of analysis of circulating miRNAs, showing the optimizations introduced toward clinical application of the miR-Test. A streamlined, standardized, and traceable processing of samples has been introduced, ensuring anonymity, through automated sample handling with barcoded tubes or plates. (B and C), A series of tests were performed to compare different extraction protocols (B) and serum starting volumes (C). Extraction efficiency was monitored by RT-qPCR quantification of 2 endogenous miRNAs (miR-19a and miR-16, in blue) and an exogenous miRNA (miR-34a, in green) that was spiked in during the extraction. n = 2 replicates. *P* values were calculated with the Student *t* test. *Significant at $P < 0.05$.

extractions. From 0.3 mL serum, we could generate 0.55 mL aqueous phase. However, to avoid contamination from the interphase/organic phase and allow rapid collection of a fixed volume from each processed sample, we chose to collect only a fraction of the aqueous phase (0.35 mL, 64%). This choice reduced the yield of the synthetic spike-in miRNA to 62% (as expected) but had no effect on the reproducibility of the process (data not shown).

Next, we optimized the detection of miRNAs by reverse-transcription quantitative PCR (RT-qPCR) with the intent of increasing sensitivity and reducing variability. We designed a customized low-density array of TaqMan® qPCR probes (cTLDA) that permits the simultaneous analysis of 48 miRNAs in up to 8 biological samples (miRNA array). The miRNA array is composed of the 13 biomarkers and the 6 housekeeping (HK) miRNAs that constitute the miR-Test *(14)*, the remaining 21 miRNAs of the original 34-miRNA signature *(13)*, and additional controls (spike-in miRNA and others) (see online Supplemental Table 1A). A series of tests (summarized in Fig. 2A) were performed to verify the reproducibility of circulating miRNA detection and discard any assays that were not compliant with a standard performance. We introduced a preamplification step after reverse transcription, which considerably improved the detection of circulating miRNAs (mean of approximately 23 vs approximately 31 Cq), while maintaining linear detection in an 8-point serial dilution experiment (see online Supplemental Table 1A). All the miRNAs that compose the miR-Test (13 biomarkers and 6 HK) were quantified with high sensitivity (mean approximately 21.9 Cq) and with high correlation coefficients (mean 0.96) in the serial dilution experiment. Thus, these miRNAs passed this validation step. Conversely, 10

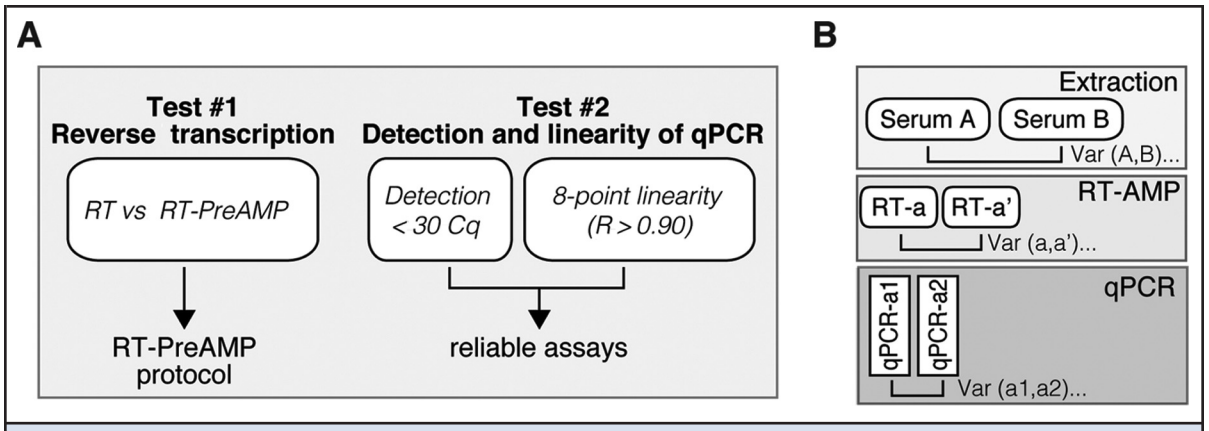

Fig. 2. Optimization of the detection of circulating miRNAs by RT-qPCR.

(A), Outline of the preliminary tests used to assess the detection of serum miRNAs by qRT-PCR. miRNA assays included in the cTLDA (miRNA array) were tested for optimal detection (<30 Cq) and linearity ($R > 0.9$, as measured by an 8-point serial dilution experiment). Further details on these analyses are provided in online Supplemental Table 1. (B and C), Variability in detection of the 13 miRNA markers and 6 HK endogenous reference miRNAs of the miR-Test was evaluated by performing replicas of RT-qPCR, RT preamplification (PreAMP), and serum extraction starting from different aliquots of the same serum (pool of healthy donors). (C), Mean variances of each of the 19 miRNAs. Extraction replicates were performed on the same day (group A) or different days (group B). n, number of replicates. (D), Variability in miRNA detection was measured on multiple extraction replicates (n = 8) performed with different protocols (gray dots, 13 miRNA markers; blue dots, 6 HK). The semiautomatic (QIAcube) protocol was repeated with 2 different workstations (group A vs group B). Aqueous phase (A.Ph.) refers to samples that were mixed together before loading onto different (n = 8) columns. (E), Variances of the 13 miRNA markers were calculated with raw data from D or the same data normalized with the endogenous references (6 HK) or the miRNA spiked in during extraction (Spike). (F), Variability in miRNA detection was measured (as in C) in different clinical serum samples from 12 healthy individuals (12N) or 12 lung cancer patients (12T). n = 2 replicates. (G and H), Effect of data normalization on assay precision. Variances of the 13 miRNA markers were calculated before and after normalization of data shown in F, as described in E. (I), Distribution of the risk score calculated with the 13 miRNA model (miR-Test) and 6-HK normalization on independent aliquots of the same serum (n = 32). Data were fitted with a normal distribution ($P = 0.17$ by Shapiro–Wilk normality test, where significant P rejects the hypothesis) and 95% CIs [mean (2σ)]. P values (C, D, E, G, and H) were calculated with the nonparametric (Wilcoxon) test.

Continued on page 747

miRNAs of the original 34-miRNA signature, which were not included in the final miR-Test *(14)*, failed to pass and were excluded from further analyses (see online Supplemental Table 1A).

We next measured intra- and interassay variability of miRNA detection through nested precision experiments by use of different aliquots of a serum pool from healthy donors and evaluated the specific contribution of imprecision in miRNA extraction, reverse transcription, and preamplification (RT-AMP) and qPCR analysis (Fig. 2B). For most miRNAs (18/19 of the miR-Test), intraassay variance of the entire procedure was low (<0.2 Cq), mainly because of the extraction step rather than the RT-AMP and qPCR (Fig. 2C; online Supplemental Table 1B), in line with previous observations *(11)*. Importantly, when we repeated the extraction of the same serum over different days (interassay variability), we observed comparable variances ($P = 0.51$) (Fig. 2C), suggesting high reproducibility of the entire procedure.

We analyzed further the variability due to the extraction step. No significant differences were observed between the manual and semiautomated protocols ($P = 0.16$) or between 2 different workstations ($P = 0.45$) (Fig. 2D). Next, we investigated variability specifically due to miRNA binding and elution from the column. We repeated the entire procedure combining the aqueous phase of 8 samples and subsequently loading this mix onto 8 different columns, thus removing the variability due to the organic extraction step. As expected, this aqueous-phase group displayed a reduced variance (0.088 vs 0.145 or 0.151 for QIAcube-A and QIAcube-B, respectively; $P < 0.05$) (Fig. 2D), corresponding to 64% of the variance for the entire procedure. Thus, the RNA extraction step is the major source of technical variability. We reasoned that the variability of the extraction procedure could have resulted from differences in the recovery of total RNA, which should be corrected by normalizing data *(4, 10)*. Indeed, when raw data were normalized by use of the endogenous reference (6 HK) or synthetic spiked-in miRNA (miR-34a), the variance of the 13 diagnostic miRNAs was significantly reduced (approximately 50%) (Fig. 2E).

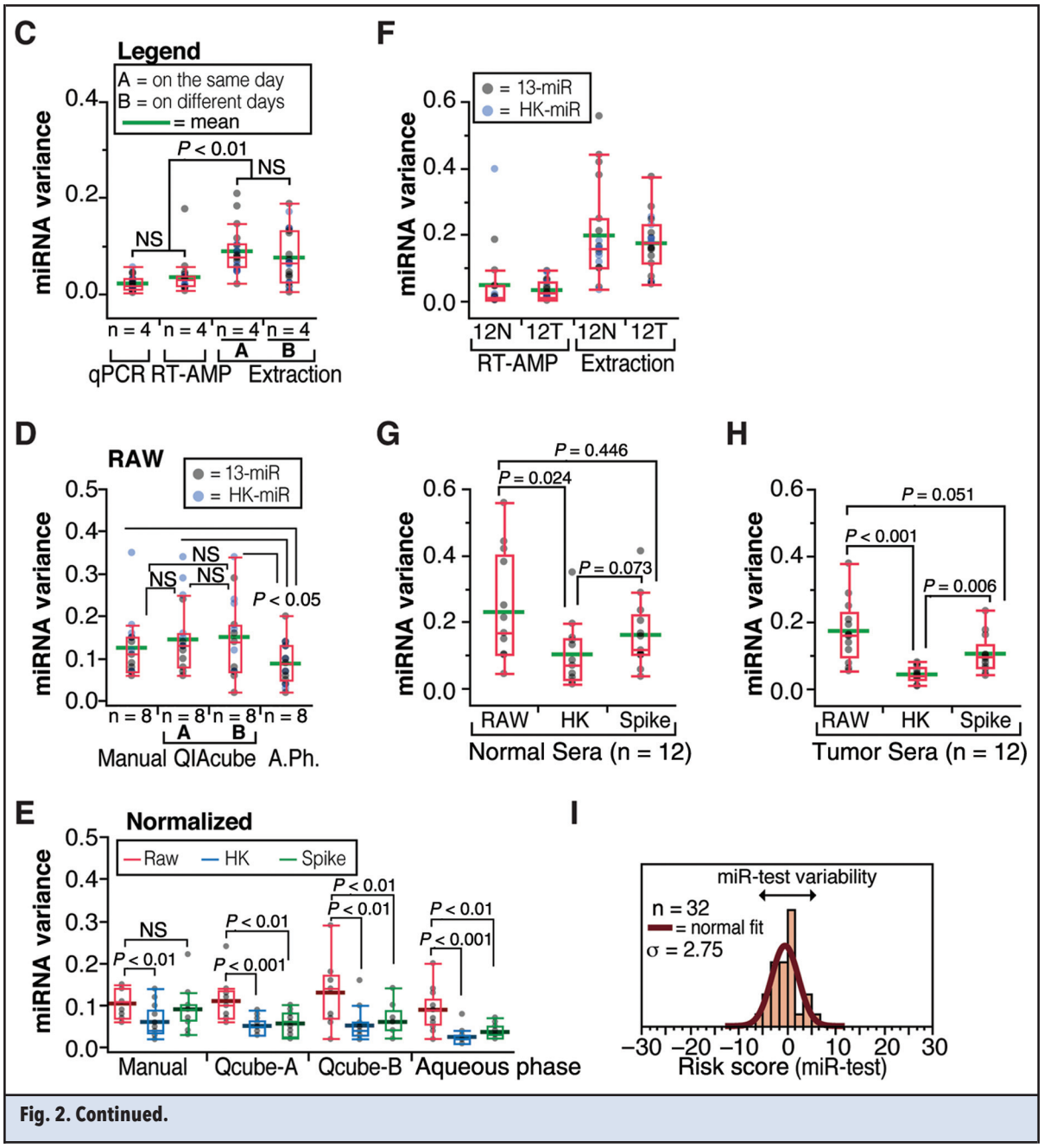

Fig. 2. Continued.

We also evaluated the fluctuations of miRNA detection by use of different clinical samples. We measured assay variability by repeating (n = 2) the extraction and analysis on a set of samples from the COSMOS screening cohort, composed of 12 healthy individuals [low-dose computed tomography (LDCT) negative] and 12 asymptomatic lung cancer patients found by LDCT. The RT-AMP replicates displayed a variance similar to intraassay variability (Fig. 2F; online Supplemental Table 1B). In contrast, serum miRNA extraction displayed a higher variability (>0.2 Cq) (Fig. 2F), likely because of sample-related differences in recovery of total RNA. Therefore, we tested the effect of normalization by using the endogenous reference (6 HK) or the synthetic spiked-in miRNA. Both approaches significantly reduced the interassay variability, with the 6-HK method outperforming the synthetic spiked-in method (Fig. 2, G and H).

Overall, these results show that the main source of analytical variability comes from the extraction of RNA

and is likely due to highly variable serum matrices (i.e., lipid or serum protein concentration), a frequent occurrence in clinical samples. Normalization on the basis of multiple endogenous miRNAs (6 HK) was able to reduce this analytical variability and should be the preferred choice over the use of synthetic spikes.

Finally, to assess the impact of assay variability on miR-Test performance, we calculated the risk score in the 32 sera aliquots previously analyzed (Fig. 2, D and E). We observed a normal distribution of the risk scores (Fig. 2I), with a mean of 0.25 and SD (σ) of 2.75, meaning that 95% of the miR-Test scores, when considering the variability due to analytical factors, are within the 5.5 range (2σ) (Fig. 2I).

PREANALYTICAL CONFOUNDING VARIABLES

A critical issue concerns the identification of variability caused by lack of standardized criteria in the collection or handling of the samples (preanalytical factors). For instance, a marked variability in miRNA levels has been reported from storage conditions of serum/plasma samples or the presence of hemolysis *(11, 17, 18)*. Aware of these critical aspects, we decided to perform a series of experiments to identify possible sources of circulating miRNA variability and analyze their impact on miR-Test performance (Fig. 3A). For each experimental condition, we measured the levels of circulating miRNAs by the miR-Test (13-miRNA and 6-HK models).

We initially focused on the nutritional status of individuals at phlebotomy. Blood samples were collected at fasting and 1, 2, or 3 h after food intake. The amount of miRNAs varied substantially in serum from fasting (0 h) and nonfasting individuals, and this variation clearly influenced the miR-Test results (Fig. 3, B and C). In particular, the risk score decreased immediately after food consumption, with a minimal recovery in the observational time window (Fig. 3B). Thus, we established that blood collection should occur under strictly controlled dietary conditions (i.e., fasting).

We also assessed whether differences in sample handling might affect miRNA detection. We used different aliquots of blood samples from healthy individuals to prepare sera, varying the clotting time (1, 3, or 6 h) and storage conditions [4 °C or −80 °C]. Clotting time only slightly affected miRNA levels, which, consequently, translated into minimal changes in the miR-Test risk score that were within the normal range of test variability, at least for samples clotted for up to 3 h (Fig. 3, D and E). Conversely, storage conditions greatly affected miRNA levels and miR-Test risk score (Fig. 3, D–F). Almost all tested miRNAs decreased in quantity after storage at 4 °C compared with serum samples that were immediately frozen and stored at −80 °C (Fig. 3D). On the basis of these results, we defined a routine protocol with a fixed clotting time (3 h) and immediate storage at −80 °C.

Next, we evaluated whether low-level hemolysis, which could occur during serum collection, affects the performance of the miR-Test. Low-level hemolysis causes an increase of some miRNA species, such as miR-16 and miR-451, which are present at high levels in RBCs *(11, 17, 18)*. We generated a hemolytic curve (8 points, Fig. 4A) by serially diluting erythrocyte lysate in nonhemolyzed serum of a healthy individual. With this series of samples, we evaluated the level of hemolysis directly by spectrophotometry (Fig. 4, B and C). We used a lipemia-independent measurement of hemolysis, the hemolysis index *(16)*, as a reference to compare against other methods based either on the absorption spectrum of oxyhemoglobin (*A* 414 nm) or the clinical measurement of free hemoglobin (in milligrams per deciliter) by a laboratory analyzer (Abbott c8000, Abbott Diagnostics). All methods produced comparable and highly correlated results ($R^2 > 0.99$) (Fig. 4B). In the absence of data normalization, we observed a general increase in serum miRNA concentrations, proportional to the hemolytic levels (Fig. 4C), in line with other recent reports *(17)*. The expression of half of the 6 HK also significantly correlated with the percentage of hemolysis (miR-15b, -19a, and -19b), whereas others (e.g., miR-146a) were not influenced (Fig. 4, C and D; online Supplemental Table 2). Nevertheless, we did not observe a proportional increase or decrease of the risk score of these samples with increasing hemolysis (Fig. 4E), suggesting that data normalization by 6 HK compensates for variations due to hemolysis. Hence, we concluded that the performance of the miR-Test is not affected by low-level hemolysis.

Notwithstanding the above result, we decided to introduce in the routine protocol 2 QC steps to monitor low-level hemolysis. The first is based on measurement of the hemolysis index by spectrophotometry on freshly prepared samples, with samples flagged as hemolytic when values are >0.2 (>5% of hemolysis). The second step, miRNA hemosensor, is based on the mean Cq ratio between HK miRNAs whose levels depend on hemolysis (miR-19a, -19b, and -15b) and a circulating HK-miRNA unaffected by hemolysis (miR-146a) *(11, 17)*. The miRNA hemosensor highly correlates with hemolysis (Pearson $r = 0.91$) (Fig. 4F) and can be used to stratify samples even from retrospective cohorts, when sera have already been processed and the measurement of absorbance is not possible.

TOWARD CLINICAL APPLICATION: EVALUATION OF TEST RELIABILITY IN A LARGE CONSECUTIVE COHORT

Translation into clinical application of signatures on the basis of circulating miRNAs requires extensive validation in independent and large cohorts of patients. Aware of

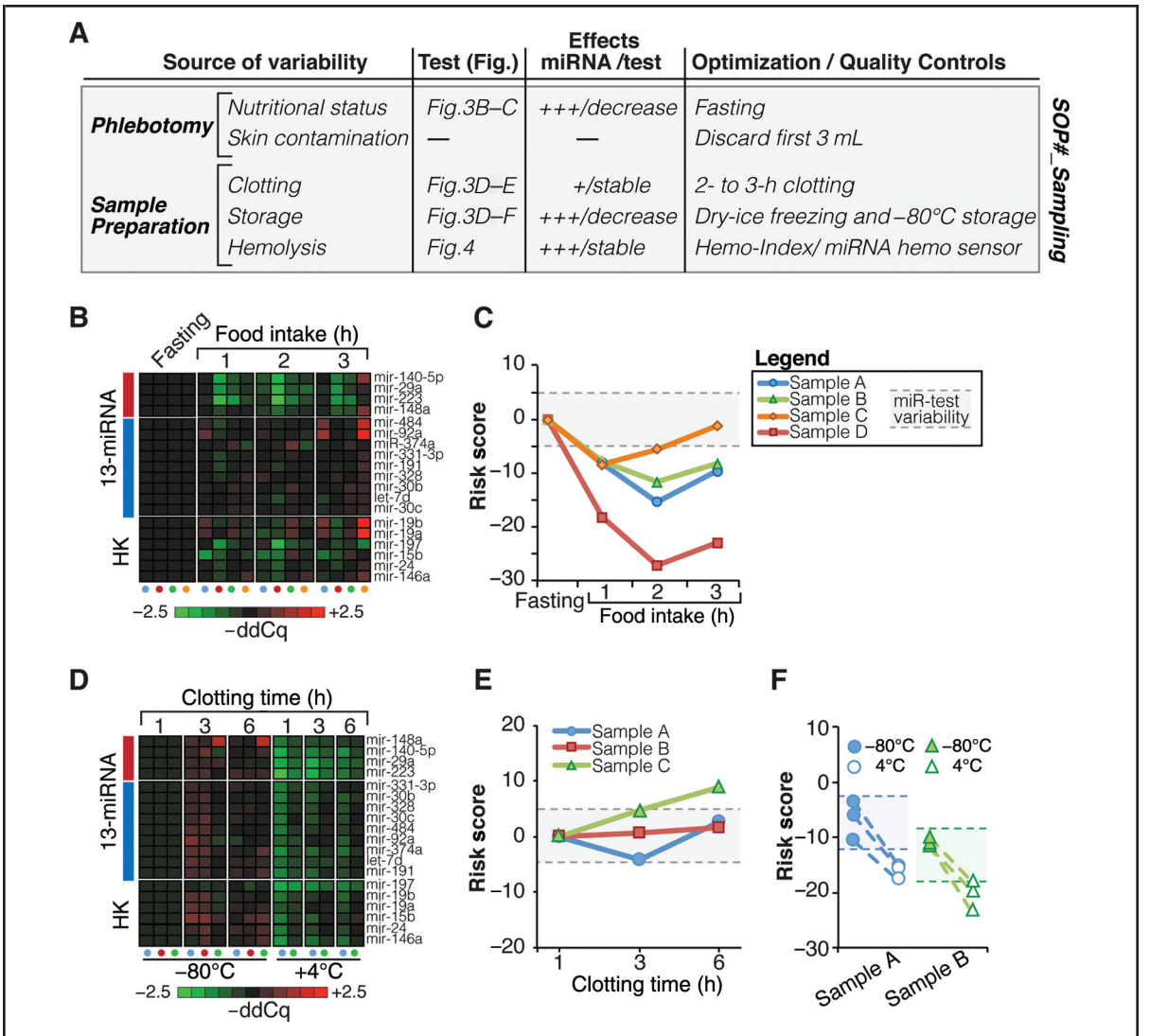

Fig. 3. Effects of preanalytical factors on circulating miRNA levels and the miR-Test risk score.

(A), Summary of the preanalytical factors considered as a potential source of variability for the quantification of circulating miRNAs. For each condition, a test was performed to evaluate the impact on miRNA levels (+++, great; +, mild) and the effect on the miR-Test risk score. Also shown are the optimizations and QC steps included in SOPs (see online Supplemental File, SOPs) for blood sampling and circulating miRNA profiling. (B), Differences in circulating miRNA expression in 4 healthy individuals upon collection of blood under different dietary conditions: fasting or 1, 2, or 3 hours after food intake. Rows represent $\log_2$ ratios (−ddCq) of expression of each miRNA in each individual (red, increased expression; green, decreased expression). −ddCq data were normalized for the spiked-in control and centered on miRNA expression at fasting. (C), Fluctuations of the miR-Test risk scores in different samples from each individual taken under different dietary conditions. The 2 dashed lines highlight boundaries of technical variability of the miR-Test score (within 2σ, as in Fig. 2I). (D), Apparent differences in circulating miRNA expression in healthy individuals due to different sample preparation conditions; i.e., clotting time (1, 3, and 6 h) and storage conditions (samples snap-frozen in dry ice and stored at −80 °C or 4 °C overnight before the analysis). Data were normalized as in B and centered on the miRNA expression levels at 1-h clotting. (E and F), Fluctuations of the miR-Test risk score within each individual due to different clotting times (E) or storage conditions (F).

this issue, we performed a large-scale validation study of miR-Test in a cohort of high-risk individuals (>1000 samples) enrolled in the COSMOS lung cancer screening trial, whose results have been published separately *(14)*. We took advantage of this large-scale study to analyze the effects of technical and nontechnical variables on the quantification of circulating miRNAs and to monitor the QC measurements introduced.

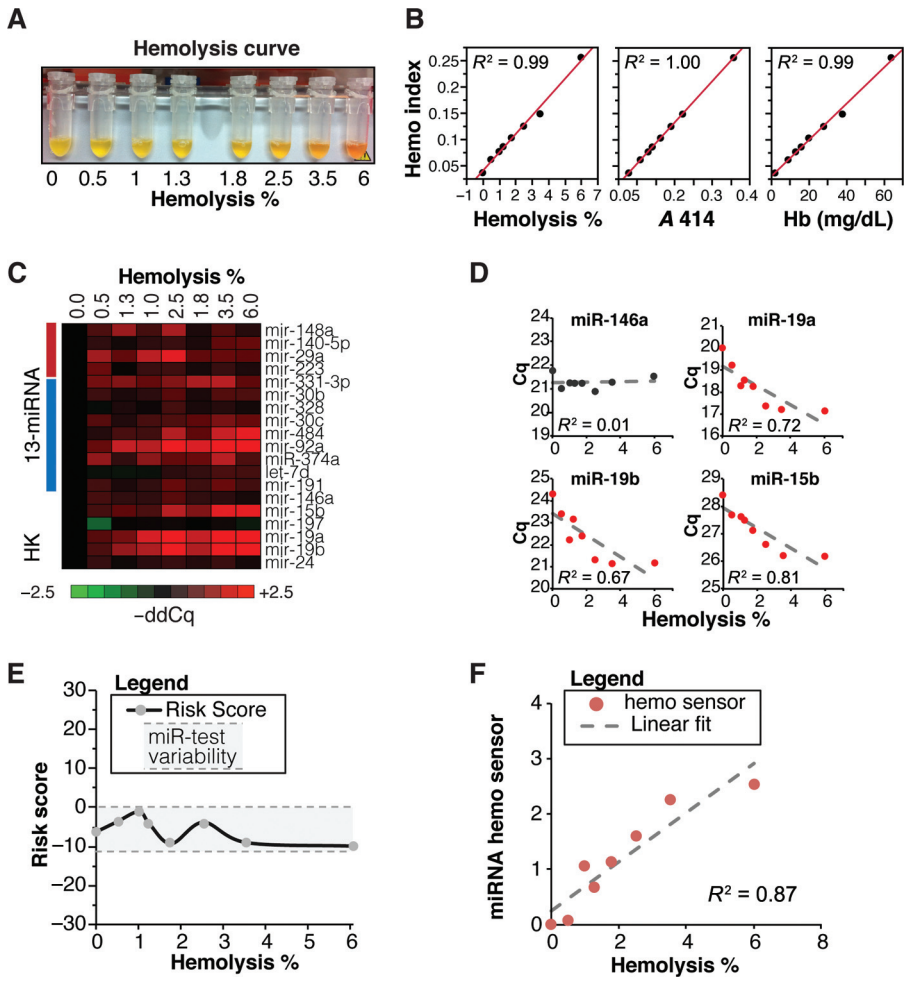

Fig. 4. Effects of hemolysis on circulating miRNAs and the miR-Test risk score.

(A), Picture of samples used for the analyses related to hemolysis. Values below indicate the percentage of erythrocyte lysate that was added to a nonhemolyzed serum sample (0% hemolysis) from a healthy individual to generate the 8-point curve. (B), Hemolysis in samples described in A was measured directly with different spectrophotometric methods [hemo index, *A* 414 nm, and hemoglobin (Hb) in milligrams per deciliter; see Methods for details), which were compared by bivariate analysis. Coefficients of determination (R^2) are shown. (C), Differences in miR-Test miRNA (13 miRNA and 6 HK) quantities in serum samples with different percentage hemolysis as described in A. Rows represent $\log_2$ ratios (−ddCq) of quantities of miRNAs in each sample relative to the 0% hemolysis sample used as a reference (red, increased expression; green, decreased expression). (D), Correlation between miRNA expression levels (shown in C) and percentage hemolysis shown by bivariate analysis. Four miRNAs are reported: miR-146a, with no correlation, and miR-19a, miR-19b, and miR-15b, showing high correlation. Coefficients of determination (R^2) are also reported. (E), miR-Test risk scores of the samples described in A. Gray dashed lines mark boundaries for technical variability of miR-Test score. (F), Correlation between the miRNA hemo sensor (mean of the ratios between miR-146a Cq value vs miR-19a/miR-19b/miR-15b Cq values) and percentage of hemolysis for samples described in A shown by bivariate analysis. Coefficient of determination (R^2) is reported.

To keep track of the analytical variables (RNA extraction plus RT-qPCR), we exploited a negative control (cel-miR-39) that always displayed levels below the detection limit ("no amp" or Cq >30; 971/972, 99.87%) and 2 synthetic spiked-in miRNAs (hsa-miR-34a and ath-miR-159) used to monitor RNA extraction (Fig. 5). In a few cases (17/972, 1.75%), the spike-in miRNAs displayed very different results (ΔCq >3). In those cases, the spike-in with the higher Cq value was not concordant with the mean of the endogenous reference (6 HK), suggesting not that the higher Cq values were reflecting low extraction efficiency but that a technical problem was

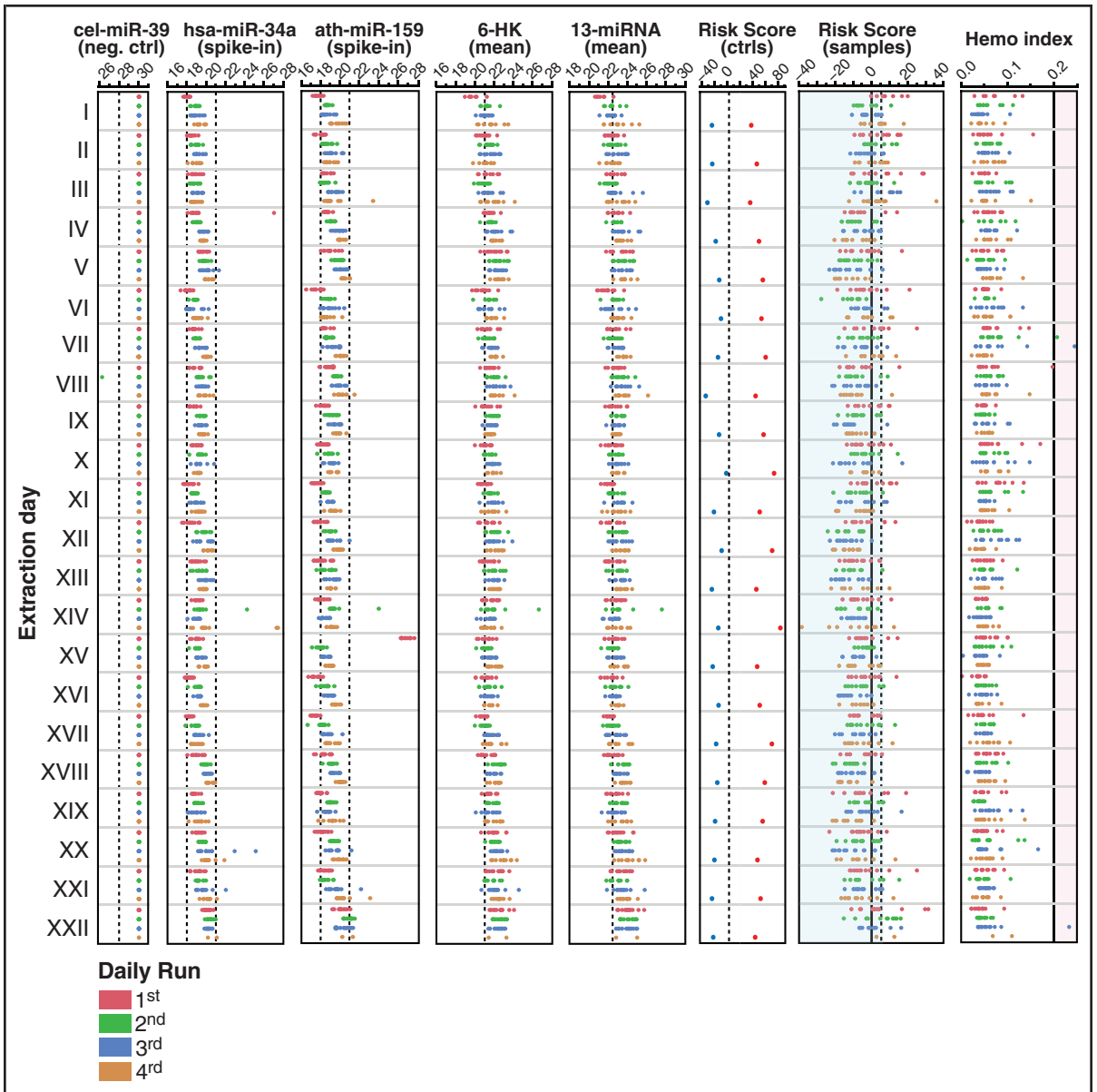

Fig. 5. Quantification of circulating miRNAs in a clinical set (COSMOS cohort).

Measurement of circulating miRNAs and controls over a large cohort of individuals from the COSMOS study (n = 972) *(14)*. Shown are the raw Cq values of the negative control (cel-miR-39; neg. ctrl), 2 spiked-in miRNAs (hsa-miR-34a and ath-miR-159a), the mean of the endogenous reference (6 HK), and the mean of the 13 diagnostic miRNAs. The miR-Test risk scores of the control samples (ctrls: negative and positive samples, blue and red dots) and the clinical samples (samples) are also shown. The blue area marks samples that gave a negative result in the miR-Test. The measurement of hemolysis by spectrophotometry (hemo index) is shown on the far right. The light pink area marks samples with a hemo index >0.2.

likely (i.e., imprecision in pipetting or degradation) (Fig. 6A). Overall, the endogenous reference was less variable than the spiked-in miRNAs, with a mean centered on 21.79 Cq and SD of 0.78 Cq, much lower than that of the 2 spike-in miRNAs (1.33 and 0.89), supporting our previous conclusion on normalization (Fig. 2).

We also introduced control samples, which were obtained by mixing different amounts of synthetic versions of the 13 miR-Test miRNAs to produce positive and negative scores in the miR-Test (Fig. 5). These controls were processed together with clinical samples to monitor possible systematic biases in each batch (1 batch = 46

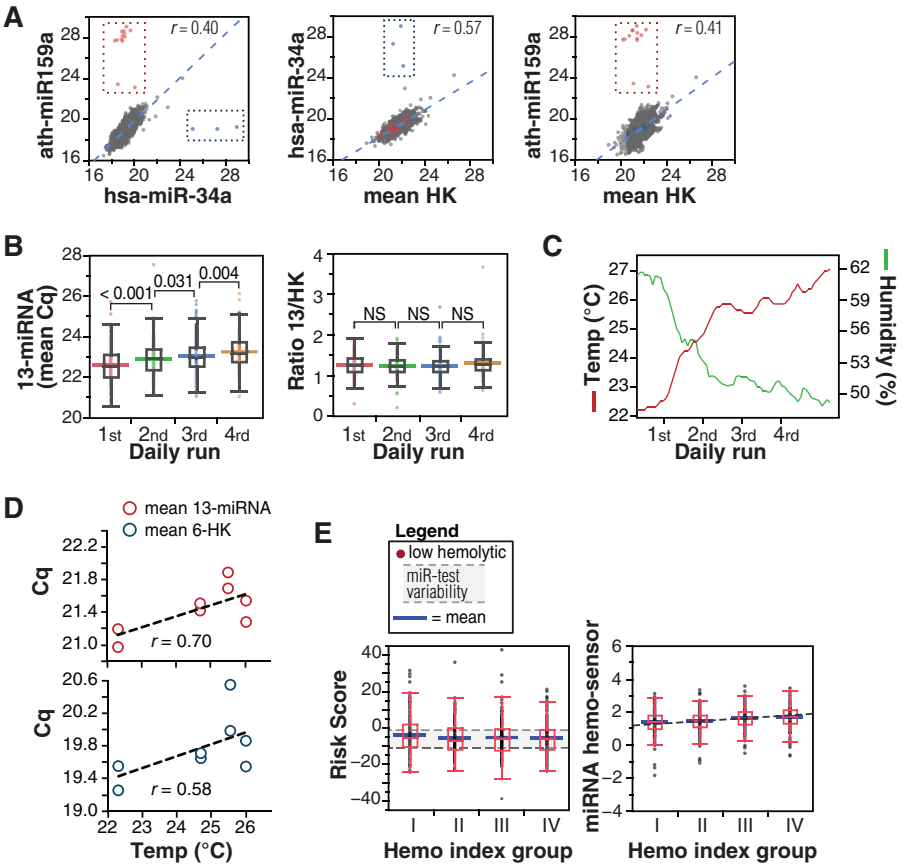

Fig. 6. Analysis of circulating miRNAs in a clinical set (COSMOS cohort).

(A), Expression levels (Cq) of the 2 spiked-in synthetic miRNAs (hsa-miR-34a and ath-miR-159a) and the mean of the endogenous reference (6 HK) in the 972 samples of the COSMOS study were compared by bivariate analysis. Samples that displayed very different levels of spike-in miRNAs are highlighted and boxed. Pearson correlation (*r*) is shown. (B), Left, distribution of the mean of the 13-miRNAs in the COSMOS cohort, grouping samples according to the round of extraction per day. Right, the 13 miRNAs normalized to the 6 HK. *P* values were calculated by nonparametric Wilcoxon test. (C and D), Temperature (red line) and humidity (green line) in the QIAcube workstation measured during 4 consecutive rounds of RNA extraction from 8 aliquots of the same sample (2 samples per run). (D), Expression of 13 miRNAs (red) or 6 HK (blue) in these samples was correlated with the temperature measured during the rounds of extraction. (E), Left, clinical samples from the COSMOS cohort were divided into quartiles on the basis of the hemolysis score (hemo index from I to IV). The plot shows the mean risk score measured by the miR-Test for each group of samples. The samples with hemolysis (hemolytic index >0.2) are marked in red. Right, values of the miRNA hemolytic sensor in clinical samples.

samples plus 2 controls). The control samples always fell in the expected range of risk (22/22, 100%) and maintained a similar gap between replicates (Fig. 5). When we stratified miRNA expression (mean of 13 miRNAs) according to the extraction run, we noticed a change in the extraction efficiency between the first run of the day and the other rounds (second, third, and fourth) (Fig. 6B), suggesting a systematic bias affecting miRNA recovery. Nonetheless, 6-HK normalization by and large compensated for this effect (ratio 13/HK) (Fig. 6B). We hypothesized that the variable efficiency in RNA isolation might be caused by different working temperatures during the extraction, which has no temperature control. Indeed, we observed a shift in the temperature (from 22 °C to 26 °C) and humidity (from 60% to 50%) inside the instrument in 4 consecutive rounds of extraction from different aliquots of the same serum pool (Fig. 6C), which correlated with a decrease in miRNA recovery (Fig. 6D). As this bias equally affects diagnostic (13 miRNA) and endogenous reference (6 HK) miRNAs, normalization compensates for such an effect and, hence, the risk score of the miR-Test was not affected.

Finally, we monitored low-level hemolysis and the possible related biases introduced in the miR-Test. Almost all samples (969/972, 99.7%) had a hemo index below the threshold (<0.2) (Fig. 5). Three samples were flagged as hemolytic but had Cq values and risk scores in the same range as the other samples (see online Supplemental Table 3). To assess the impact (if any) of low-level hemolysis in the entire cohort, we subdivided all samples into quartiles according to the hemo index and correlated these groups with miR-Test scores. In line with results described above (Fig. 3), we did not observe significant differences between groups, confirming that the miR-Test is not affected by low-level hemolysis (Fig. 6E).

Discussion

Here we describe the results of an extensive analysis performed on a circulating miRNA signature *(13, 14)*, with the aim of optimizing the performance and reproducibility of miRNA detection from serum samples toward clinical application. This study has led to the development of a pipeline for processing of clinical serum samples that minimizes nonspecific fluctuations at analytical and preanalytical levels.

At the analytical level, the major source of imprecision came from the RNA extraction rather than RT-qPCR, in agreement with previous reports *(11)*. To limit analytical imprecision, we developed a tailored protocol that combines liquid–liquid extraction with column-based purification. This protocol yielded results with a very high level of reproducibility because of a decrease in manual operations and optimization of volumes used during RNA extraction. Overall, variability in miRNA levels was low (<0.2 Cq), in particular compared with the variability observed with other protocols *(11)*, although a direct comparison was possible for only a few miRNAs (miR-24, variance 0.06 vs 0.131; miR-15b, variance 0.09 vs 0.224). Furthermore, the variability introduced in the RNA extraction step could be corrected by normalizing the amount of input RNA used in the quantification. In the absence of any reliable methodology for quantification of RNA isolated from serum *(19)*, we preferred to use fixed volumes of input RNA and exogenous (such as synthetic spiked-in miRNAs) or endogenous (6 HK miRNAs) references for data normalization. Although both strategies were effective, we consistently achieved better results by use of a combination of 6 endogenous miRNAs (6 HK, described in Bianchi et al. *(13)*).

When the impact of preanalytical variables was systematically analyzed, we found that the levels of circulating miRNAs were often heavily influenced. One major confounding variable was the nutritional status of individuals at phlebotomy. Similarly, serum contamination by hemolysis caused a marked variability in miRNA levels, as previously reported *(11, 17, 18)*. We evaluated these effects specifically on the miR-Test by comparing the intrinsic variability of the test due to technical imprecision (within 5.5 points) with the variability introduced by preanalytical variables. Although the fluctuations introduced by food intake strongly altered the risk score, the effects of low-level hemolysis could be mostly compensated for by sample normalization. On the basis of these findings, we recommend as general rule to use strictly controlled procedures in the way samples are collected and prepared. Consequently, we developed SOPs for sampling (blood withdrawal and serum preparation) and processing (RNA extraction and analysis) to limit and keep track of possible sources of unspecific fluctuations (see online Supplemental File).

In view of the clinical application of the miR-Test, we decided to introduce QC steps to monitor the performance of the signature and ensure the reliability of the test. These steps included measurement of hemolysis, multiple spike-in miRNAs to monitor RNA extraction from serum, and internal controls (positive and negative) for assessing miR-Test performance in each batch of analysis. In a large clinical cohort (approximately 1000 samples), we evaluated how each of these controls behaved and checked for unexpected technical/biological issues. For instance, we observed a bias in the yields of extracted miRNA between different runs performed on the same day. This systematic bias was effectively corrected by normalization and did not alter miR-Test performance. Nevertheless, it highlights the importance of standardizing every aspect of circulating miRNA analysis to avoid unspecific variations that could hamper the clinical application of even the most promising miRNA biomarkers.

In conclusion, we have described the optimization and standardization of the detection of circulating miRNAs with a view toward clinical application. Although we focused on a specific set of miRNAs, our results highlight general rules and implementations that need to be taken into account when considering the transfer of tests on the basis of circulating miRNA biomarkers from the wet laboratory to the diagnostic/clinical setting.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

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

Employment or Leadership: L. Spaggiari, European Institute of Oncology; F. Bianchi, European Institute of Oncology.

Consultant or Advisory Role: F. Nicassio, European Institute of Oncology.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: G. Veronesi, Associazione Italiana per la Ricerca sul Cancro IG12818; P.P. di Fiore, Associazione Italiana per la Ricerca sul Cancro MCO (AIRC) 10.000, the Monzino Foundation; F. Nicassio, Associazione Italiana per la Ricerca sul Cancro IG14085, the Umberto Veronesi Foundation.
Expert Testimony: None declared.
Patents: M.J. Marzi, coinventor on WO/2012/089630 (licensed to GENSIGNIA) and coinventor on WO2016038119 (A1) METHODS FOR LUNG CANCER DETECTION; F. Montani, coinventor on WO/2012/089630 (licensed to GENSIGNIA) and coinventor on WO2016038119 (A1) METHODS FOR LUNG CANCER DETECTION; F. Bianchi, coinventor on WO/2012/089630 (licensed to GENSIGNIA) and coinventor on WO2016038119 (A1) METHODS FOR LUNG CANCER DETECTION; P.P. Di Fiore, coinventor on WO/2012/089630 (licensed to GENSIGNIA) and coinventor on WO2016038119 (A1) METHODS FOR LUNG CANCER DETECTION; F. Nicassio, coinventor on WO/2012/089630 (licensed to GENSIGNIA) and coinventor on WO2016038119 (A1) METHODS FOR LUNG CANCER DETECTION.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: We are indebted to Davide Proverbio, Giulio Taliento, Andrea Uggetti, and Maria Lucia Longo for blood collection and sera preparation. We also thank Rosalind Gunby for critically editing the manuscript.

References

1. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009;136:215–33.
2. Kozomara A, Griffiths-Jones S. MiRbase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 2014;42:D68–73.
3. Lujambio A, Lowe SW. The microcosmos of cancer. Nature 2012;482:347–55.
4. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A 2008;105:10513–8.
5. Skog J, Wurdinger T, van Rijn S, Meijer DH, Gainche L, Sena-Esteves M, et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nat Cell Biol 2008;10:1470–6.
6. Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF, et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. Proc Natl Acad Sci U S A 2011;108:5003–8.
7. Vickers KC, Palmisano BT, Shoucri BM, Shamburek RD, Remaley AT. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. Nat Cell Biol 2011;13:423–33.
8. Kosaka N, Iguchi H, Yoshioka Y, Takeshita F, Matsuki Y, Ochiya T. Secretory mechanisms and intercellular transfer of microRNAs in living cells. J Biol Chem 2010;285:17442–52.
9. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. Nature Rev Cancer 2011;11:426–37.
10. Kirschner MB, van Zandwijk N, Reid G. Cell-free microRNAs: potential biomarkers in need of standardized reporting. Front Genet 2013;4:56.
11. McDonald JS, Milosevic D, Reddi HV, Grebe SK, Algeciras-Schimnich A. Analysis of circulating microRNA: preanalytical and analytical challenges. Clin Chem 2011;57:833–40.
12. Witwer KW. Circulating microRNA biomarker studies: pitfalls and potential solutions. Clin Chem 2015;61:56–63.
13. Bianchi F, Nicassio F, Marzi M, Belloni E, Dall'olio V, Bernard L, et al. A serum circulating miRNA diagnostic test to identify asymptomatic high-risk individuals with early stage lung cancer. EMBO Mol Med 2011;3:495–503.
14. Montani F, Marzi MJ, Dezi F, Dama E, Carletti RM, Bonizzi G, et al. MiR-test: a blood test for lung cancer early detection. J Natl Cancer Inst 2015;107:djv063.
15. Hanson MS, Stephenson AH, Bowles EA, Sridharan M, Adderley S, Sprague RS. Phosphodiesterase 3 is present in rabbit and human erythrocytes and its inhibition potentiates iloprost-induced increases in cAMP. Am J Physiol Heart Circ Physiol 2008;295:H786–93.
16. Appierto V, Callari M, Cavadini E, Morelli D, Daidone MG, Tiberio P. A lipemia-independent Nanodrop®-based score to identify hemolysis in plasma and serum samples. Bioanalysis 2014;6:1215–26.
17. Kirschner MB, Edelman JJ, Kao SC, Vallely MP, van Zandwijk N, Reid G. The impact of hemolysis on cell-free microRNA biomarkers. Front Genet 2013;4:94.
18. Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM, et al. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. Cancer Prev Res (Phila) 2012;5:492–7.
19. Kroh EM, Parkin RK, Mitchell PS, Tewari M. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). Methods 2010;50:298–301.