



The expression of circulating miR-504 in plasma is associated with *EGFR* mutation status in non-small-cell lung carcinoma patients

Adam Szpechcinski¹ · Mateusz Florczuk¹ · Katarzyna Duk¹ · Aneta Zdral¹ · Stefan Rudzinski¹ · Maciej Bryl² · Grzegorz Czyzewicz³ · Piotr Rudzinski⁴ · Włodzimierz Kupis⁴ · Emil Wojda⁵ · Dorota Giedronowicz⁶ · Renata Langfort⁶ · Aleksander Barinow-Wojewodzki² · Tadeusz Orłowski⁴ · Joanna Chorostowska-Wynimko¹

Received: 4 October 2018 / Revised: 6 March 2019 / Accepted: 28 March 2019 / Published online: 5 April 2019
© The Author(s) 2019

Abstract

MicroRNAs (miRNAs), key regulators of gene expression at the post-transcriptional level, are grossly misregulated in some human cancers, including non-small-cell lung carcinoma (NSCLC). The aberrant expression of specific miRNAs results in the abnormal regulation of key components of signalling pathways in tumour cells. MiRNA levels and the activity of the gene targets, including oncogenes and tumour suppressors, produce feedback that changes miRNA expression levels and indicates the cell's genetic activity. In this study, we measured the expression of five circulating miRNAs (miR-195, miR-504, miR-122, miR-10b and miR-21) and evaluated their association with *EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR)* mutation status in 66 NSCLC patients. Moreover, we examined the discriminative power of circulating miRNAs for *EGFR* mutant-positive and -negative NSCLC patients using two different data normalisation approaches. We extracted total RNA from the plasma of 66 non-squamous NSCLC patients (31 of whom had tumours with *EGFR* mutations) and measured circulating miRNA levels using quantitative reverse transcription polymerase chain reaction (RT-qPCR). The miRNA expression levels were normalised using two endogenous controls: miR-191 and miR-16. We found significant associations between the expression of circulating miR-504 and *EGFR*-activating mutations in NSCLC patients regardless of the normalisation approach used ($p = 0.0072$ and 0.0236 for miR-16 and miR-191 normalisation, respectively). The greatest discriminative power of circulating miR-504 was observed in patients with *EGFR* exon 19 deletions versus wild-type *EGFR* normalised to miR-191 (area under the curve (AUC) = 0.81, $p < 0.0001$). Interestingly, circulating miR-504 levels were significantly reduced in the v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*)-mutated subgroup compared to *EGFR*-mutated patients ($p < 0.0030$) and those with *EGFR/KRAS* wild-type tumours ($p < 0.0359$). Our study demonstrated the feasibility and potential diagnostic value of plasma miR-504 expression analysis to distinguish between *EGFR*-mutated and wild-type NSCLC patients. However, quality control and normalisation strategies are very important and have a major impact on the outcomes of circulating miRNA analyses.

Keywords Epigenetics · Molecular diagnosis · Liquid biopsy · Predictive biomarker · Targeted therapy

Adam Szpechcinski and Mateusz Florczuk contributed equally to this work.

✉ Adam Szpechcinski
a.szpechcinski@igichp.edu.pl

¹ Department of Genetics and Clinical Immunology, National Research Institute of Tuberculosis and Lung Diseases, 26 Plocka St., 01-138 Warsaw, Poland

² Department of Oncology, E.J. Zeyland Wielkopolska Center of Pulmonology and Thoracic Surgery, Poznan, Poland

³ Department of Oncology, The John Paul II Specialist Hospital, Kraków, Poland

⁴ Department of Surgery, National Research Institute of Tuberculosis and Lung Diseases, Warsaw, Poland

⁵ II Department of Lung Diseases, National Research Institute of Tuberculosis and Lung Diseases, Warsaw, Poland

⁶ Department of Pathomorphology, National Research Institute of Tuberculosis and Lung Diseases, Warsaw, Poland

Introduction

Targeted therapy with tyrosine kinase inhibitors (TKIs) against the human epidermal growth factor receptor (EGFR) is currently the most common form of personalised treatment for non-small-cell lung carcinoma (NSCLC). Somatic mutations in exons 18–21 of the *EGFR* gene are the most important predictive molecular markers, and their clinical value in diagnosing NSCLC has been confirmed [1]. In-frame deletions in exon 19 account for 44% of EGFR TK-activating mutations: these involve the deletion of amino acid residues leucine-747 to glutamic acid-749. The predominant single point mutation in exon 21, L858R, accounts for 41% of EGFR TK-activating mutations. Changing glycine-719 (G719) to serine, alanine or cysteine accounts for 10%, whereas a duplication and/or insertion in exon 20 accounts for the remaining 5% of EGFR TK-activating mutations [2]. Deletions in exon 19 and the point mutation L858R are termed ‘classical’ activating mutations. These mutations result in increased EGFR kinase activity, leading to enhanced downstream events. Another point mutation in codon 790 [threonine-790 to methionine (T790M)] of exon 20 of the *EGFR* gene was reported in approximately 50–60% of all patients who acquired resistance to EGFR TKI therapy. This mutation is probably acquired during treatment, because it is rarely detected de novo in tumours from untreated patients. *EGFR* mutations occur in approximately 10–20% of Caucasian and 30–50% of Asian NSCLC patients. In contrast to standard chemotherapy, EGFR-TKIs selectively inhibit tumour cell growth by targeting the intracellular EGFR domain. Targeted therapy using TKIs, which bind to EGFR reversibly (e.g., erlotinib, gefitinib) or irreversibly (e.g., afatinib, osimertinib), is the standard personalised treatment for NSCLC in eligible patients [3]. However, most patients develop some resistance to EGFR-TKIs after 9–13 months of treatment.

NSCLC is a challenging diagnostic target due to the substantial diversity of neoplastic clones within tumours, and the difficulties associated with obtaining sufficient quantities of diagnostic material suitable for pathological and molecular evaluation [4]. Moreover, genetic alterations within the tumours change the molecular marker profile during the course of the disease. Currently, most advanced NSCLC patients are diagnosed using biopsies or cytology specimens, due to a lack of availability of resected tumour tissue [5]. However, in many patients, adequate diagnostic sampling is problematic, both de novo and post-EGFR therapy. This is reflected in published data from major clinical trials on EGFR-TKI efficacy, which show that tissue specimens suitable for molecular testing were obtained from fewer than 50% of patients [6–8]. Plasma

cell-free DNA (cfDNA) as a liquid biopsy might be a valuable surrogate for detecting tumour-specific alterations in NSCLC patients [9]. The advantages of using cfDNA for tumour mutation detection include: (1) non-invasive specimen collection; (2) availability throughout the course of the disease; (3) real-time detection and monitoring of biomarker status and dynamics; and (4) potentially fewer heterogeneity issues compared to tumour tissue testing [10]. In fact, cfDNA-based assessment of *EGFR*-activating mutation status has proven reliable in identifying NSCLC patients for targeted therapies [11]. However, the rate of *EGFR* mutation detection in cfDNA varies, depending on the method used. A large meta-analysis that included 1591 cases tested using different methods reported a pooled sensitivity rate of 64.5% and results were significantly affected by the assay type ($p=0.04$) [12].

MicroRNAs (miRNAs) are a class of noncoding, endogenous, single-stranded and relatively small RNA molecules (19–22 nucleotides) that play a key role in the regulation of basic cellular processes including differentiation, proliferation and apoptosis [13]. MiRNA molecules can regulate genes at the post-transcriptional level by specifically recognising messenger RNA (mRNA) molecules, based on the degree of complementarity to the targeted sequence [14]. The aberrant expression of particular miRNAs or miRNA families can result in the abnormal regulation of key signalling pathway components in tumour cells. The oncogenic or suppressor activity of miRNAs in carcinogenesis appears to affect the efficacy of targeted therapies, including EGFR-TKIs in lung cancer [15]. The abnormal regulation of gene expression by miRNAs can trigger alternative signalling pathways or activate downstream signalling mediators, bypassing pathways blocked by EGFR-TKIs. MiRNA expression may change dynamically in response to feedback from miRNA levels and the activity of genes targeted by miRNAs, including oncogenes and tumour suppressors. Furthermore, stable forms of tumour-related miRNAs, called circulating miRNAs, can be detected in the peripheral blood of NSCLC patients. Recently, several circulating miRNA signatures were linked to the mutation status of the *EGFR* gene and/or the response to EGFR-TKIs in NSCLC patients [16–18]. The identification of miRNA profiles linked to altered EGFR signalling and the cellular response to EGFR-TKIs may create new opportunities to develop more effective personalised therapy for patients with NSCLC. These miRNA profiles may also significantly increase the efficacy of molecular diagnostics and targeted treatment monitoring.

However, one major challenge is identifying reliable methods for analysing miRNA expression, particularly that of circulating miRNAs in the blood. MiRNA molecules are small, and those from the same family are highly homologous. They occur at low concentrations in body fluids, and standardised methodologies to detect them are lacking.

The various commercial reagent kits and systems available differ in detection method and sensitivity [19]. It has also proved difficult to identify the optimal miRNAs to use as NSCLC biomarkers, and errors have frequently been made during the planning stage of experiments. These errors have included poor patient selection and representation of patient groups, incorrect processing and storage of specimens, and no controls for haemolysis in plasma samples. There is also no consensus on the best method to normalise the resulting data, which is why all steps of an analytical procedure must be validated and standardised before any miRNA-based diagnostic method can be implemented in clinical practice.

In this study, we measured the expression of five circulating miRNAs (miR-195, miR-504, miR-122, miR-10b and miR-21) in plasma taken from NSCLC patients and evaluated their association with *EGFR* mutation status in tumour tissue and cfDNA compared to previously reported results. Moreover, we examined the discriminative power of circulating miRNAs to identify *EGFR* mutant-positive and -negative NSCLC patients, using two different data normalisation approaches.

Materials and methods

Blood collection

Blood was collected from consecutive patients with histologically confirmed primary NSCLC (non-squamous subtype), whose tumour tissue specimens were tested for the *EGFR* mutation. The study was reviewed and approved by the local ethics committee. All patients provided written informed consent.

Peripheral blood was collected from each patient before treatment. Blood was collected in 9-mL tubes containing EDTA as an anticoagulant and processed within 1 h. Plasma was separated from the cellular fraction by centrifugation twice at $1000\times g$ for 10 min at 4 °C, and stored at -80 °C for up to 3 months.

Initial assessment of haemolysis

The level of haemolysis in the plasma samples was provisionally assessed visually using a haemolysis reference chart [20]. Plasma samples showing moderate or gross haemolysis (i.e., those that were pink to dark red in colour, corresponding to haemoglobin concentrations of approximately > 200 mg/dL) were not used.

DNA extraction and *EGFR*/*KRAS* mutation analysis

Genomic DNA was isolated from deparaffinised 5- μ m sections of formalin-fixed, paraffin-embedded (FFPE) tissue

specimens using the Cobas DNA Sample Preparation Kit (Roche Diagnostics GmbH, Mannheim, Germany). CfDNA was extracted from 2 mL plasma aliquots using the Cobas cfDNA Sample Preparation Kit (Roche Diagnostics GmbH), in accordance with the manufacturer's instructions. The mutation status of *EGFR* and v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*) in cfDNA was analysed by quantitative polymerase chain reaction (qPCR) using the Cobas *EGFR* Mutation Test v2 (IVD), the Cobas *KRAS* Mutation Test v2 (LSR) and the Cobas z480 Instrument (Roche Diagnostics GmbH). The Cobas platform can identify 42 different mutations that may be present in exons 18–21 of the *EGFR* gene and 28 mutations that may be present in exons 2–4 of the *KRAS* gene. The Cobas *EGFR* test can detect *EGFR* mutations at a mutation level of at least 5% in DNA isolated from FFPE tissue, whereas in plasma, the limit of detection ranges from 10 to 100 copies of mutated cfDNA per millilitre. The Cobas *KRAS* test can reliably detect *KRAS* variants at $\geq 1\%$ mutation in FFPE and ≤ 75 copies of the mutant allele in a wild-type background of 64,000 copies per millilitre of plasma.

Immunohistochemical analysis of ALK gene rearrangements

Glass slides with 4- μ m FFPE sections were stained immunohistochemically using the D5F3 rabbit monoclonal antibody against human anaplastic lymphoma kinase (ALK) protein, in accordance with the manufacturer's instructions, and visualised using the OptiView enhanced detection and amplification system and the automated Ventana BenchMark GX Slide Staining System (Roche Diagnostics International Ltd., Basel, Switzerland). Sections were counterstained with haematoxylin and eosin. Human ganglion cells were used as the internal positive control for all specimens. Staining results were evaluated independently by two trained pathologists using the Ventana binary system, which classified strong granular cytoplasmic staining in any tumour cells as a positive result, and the absence of strong granular cytoplasmic staining as a negative result (Ventana Medical Systems, Tucson, AZ, USA).

MiRNA extraction and spike-in controls

Total circulating RNA was extracted from 100 μ L of plasma using QIAzol Lysis Reagent (Qiagen GmbH, Hilden, Germany) and the miRNeasy Mini Kit (Qiagen GmbH), in accordance with the manufacturer's instructions. A total of 1 μ L of synthetic UniSp6 RNA spike-in template, provided with the miRCURY LNA Universal cDNA synthesis kit (Exiqon A/S, Vedbaek, Denmark), was added to each plasma sample treated with QIAzol to monitor the quality of miRNA extraction, reverse transcription and qPCR.

MiRNA expression measurement using RT-qPCR

The circulating miRNA expression was measured by RT-qPCR. The extracted total circulating RNA was reverse-transcribed into complementary DNA (cDNA) using the miRCURY LNA Universal cDNA Synthesis Kit (Exiqon A/S). Tubes containing 6 μ L of RNA template, reaction buffer, primers and reverse transcriptase were incubated for 60 min at 42 °C, followed by 5 min at 95 °C, and cooled to 4 °C in a thermal cycler. The cDNA was diluted 20-fold with water before qPCR amplification.

The qPCR amplification was performed using miRNA sequence-specific miRCURY LNA PCR primers and Exi-LENT SYBR Green MasterMix Kit (Exiqon A/S). The qPCR was carried out in 96-microwell plates using the LightCycler 480 II instrument (Roche Diagnostics GmbH). The qPCR cycling conditions were: 95 °C for 10 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 1 min. The threshold for all samples was set at the same value. Moreover, a melting curve was analysed for each sample to verify the specificity of the amplification products. Each sample was tested in duplicate. In blank samples, the cDNA was replaced by an equivalent volume of water. The targeted miRNA sequences are shown in Table 1.

Final assessment of haemolysis

Each cDNA sample was again tested for haemolysis using a qPCR assay [21]. The level of the erythrocyte-specific miRNA (miR-23a-3p) was related to a miRNA known to be unaffected by haemolysis (miR-451a), allowing outlier samples to be identified and eliminated using the following formulae: $\Delta C_t(\text{miR-23a-miR-451}) > 5$, which suggests possible erythrocyte miRNA contamination; and $\Delta C_t(\text{miR-23a-miR-451}) > 7-8$, which suggests a high likelihood that haemolysis will affect the data. All samples with $\Delta C_t(\text{miR-23a-miR-451}) > 5$ were excluded from the analysis.

Statistical analyses

The raw data from the RT-qPCR analysis are expressed as mean \pm standard deviation (SD) of the qPCR C_t values for each sample duplicate. The stability of miRNA expression was evaluated using RefFinder, a web-based comprehensive tool integrating the computational programs geNorm [22], NormFinder [23], BestKeeper [24], and the comparative $\Delta\Delta C_t$ method [25] to compare and rank the tested genes according to their overall variability among the samples. Based on the rankings from each program, an appropriate weight was assigned to each miRNA and the geometric mean of the weights was calculated to obtain the overall final rankings [26]. The expression levels of the miRNAs tested (miR-122, miR-195, miR-504, miR-10b, and miR-21) were normalised relative to the expression level of each of two endogenous miRNA controls (miR-16 and miR-191). The normalised expression levels of the five miRNAs tested were calculated as ΔC_t s, where $\Delta C_t = C_t(\text{miRNA of interest}) - C_t(\text{endogenous control miRNA})$, and characterised by their mean \pm SD and medians (range), and the coefficient of variance (%CV). The normality of the data was assessed using the Shapiro–Wilk test, and the nonparametric Mann–Whitney U test was used to compare the means. Receiver operating characteristic (ROC) curves were constructed to evaluate the diagnostic capacity of circulating miRNAs to predict the *EGFR* mutation status of NSCLC patients [27, 28]. A p value < 0.05 was considered statistically significant. All statistical analyses were performed using Statistica (ver. 11; StatSoft, Inc., Tulsa, OK, USA) and MedCalc (Ostend, Belgium) software.

Results

Study population

In total, 66 blood samples from NSCLC patients met the quality criteria regarding lack of haemolysis, known *EGFR*

Table 1 Characteristics of the microRNA (miRNA) sequences analysed in this study

Mature miRNA ID	Gene family	Sequence of mature miRNA within the stem loop (5' \rightarrow 3')	miRBase accession number
hsa-miR-195-5p	mir-15	UAGCAGCACAGAAAUUUGGC	MIMAT0000461
hsa-miR-504-5p	mir-504	AGACCCUGGUCUGCACUCUAUC	MIMAT0002875
hsa-miR-122-5p	mir-122	UGGAGUGUGACAAUGGUGUUUG	MIMAT0000421
hsa-miR-10b-5p	mir-10	UACCCUGUAGAACCGAAUUUGUG	MIMAT0000254
hsa-miR-21-5p	mir-21	UAGCUUAUCAGACUGAUGUUGA	MIMAT0000076
hsa-miR-16-5p	mir-15	UAGCAGCACGUAUUAUUGGCG	MIMAT0000069
hsa-miR-191-5p	mir-191	CAACGGAAUCCCAAAGCAGCUG	MIMAT0000440
hsa-miR-451a	mir-451	AAACCGUUACCAUACUGAGUU	MIMAT0001631
hsa-miR-23a-3p	mir-23	AUCACAUUGCCAGGGAUUUCC	MIMAT0000078

mutation in tumour tissue and paired cfDNA. NSCLC histological subtypes were determined according to the World Health Organization's classification [29]. A total of 56 of the 66 NSCLC patients (85%) had been diagnosed with adenocarcinoma (ADC), 1 patient (2%) had been diagnosed with large cell carcinoma (LCC), 2 patients (3%) with a mixed histology pattern, and 7 patients (10%) with a 'not otherwise specified' subtype. NSCLC pathological stages were defined according to the tumour node metastasis (TNM) International Staging System [30]. There were 38 patients (58%) with resectable NSCLC (stages I–IIIA) and 28 patients (42%) with advanced metastatic disease (stages IIIB–IV). No patients had received preoperative chemotherapy or radiotherapy. In total, 31 of the 66 patients (47%) had *EGFR*-activating mutations in their tumour tissue and the remaining 35 patients had *EGFR* wild-type tumours. Among the *EGFR* wild-type group, 12 patients had *KRAS* mutations, 2 had *ALK*-rearrangements and 21 patients had *EGFR/KRAS/ALK* wild-type tumours. All the patients were Caucasians. The demographic and clinical characteristics of the patients are summarised in Table 2.

Table 2 Demographic and clinical characteristics of the study population

Characteristics	<i>N</i>	Percent (%)
NSCLC patients	66	100
Median age (years), range	66 (49–88)	
Sex		
Male	37	56
Female	29	44
Histology (WHO)		
ADC	56	85
NOS/other	10	15
Stage (TNM)		
Localized (I–IIIA)	38	58
Metastatic (IIIB–IV)	28	42
EGFR status in tumour		
Mutated	31	47 (100)
Exon 19 deletions	16	(52)
Exon 21 L858R	13	(42)
Other	2	(6)
EGFR wild-type	35	53 (100)
<i>KRAS</i> -mutated	12	(34)
Exon 2 substitutions	10	
Other	2	
<i>KRAS</i> wild-type	23	(66)
ALK-positive (IHC)	2	
EGFR/ <i>KRAS</i> /ALK wild-type	21	

EGFR and *KRAS* mutation status of cfDNA

The *EGFR* and *KRAS* mutation status of cfDNA from NSCLC patients was evaluated by qPCR assay using the Cobas platform (Roche Diagnostics GmbH). In total, 14 of the 31 patients (45%) with *EGFR*-activating mutations in their tumour tissue also had mutations in their plasma cfDNA. No *EGFR* mutations were found in plasma cfDNA from patients with localised NSCLC. In the advanced NSCLC subgroup, 14 of the 17 patients (82%) had the same *EGFR*-activating mutation in both tumour tissue and paired plasma. No *EGFR* mutations were detected in cfDNA from plasma taken from the 35 NSCLC patients with *EGFR* wild-type tumours, regardless of the clinical stage of the disease. The 5 out of 12 patients (42%) with *KRAS*-mutated tumours had the same variant in their plasma cfDNA (two patients had localised NSCLC and three patients had advanced disease). No *KRAS* mutations were found in plasma from patients with *KRAS* wild-type tumours.

Quality control of plasma samples and miRNA expression analysis

The level of haemolysis in plasma samples was assessed by RT-qPCR using the delta C_t (miR-23a–miR-451) formula. The 66 plasma samples included in the study had a mean \pm SD delta C_t (miR-23a–miR-451) of 1.15 ± 1.94 and a median of 1.05 (range – 6.41 to 4.95), which was below the cut-off value of 5.0. The synthetic UniSp6 RNA spike-in template was added to each plasma sample to monitor the quality of miRNA extraction, reverse transcription and qPCR. The samples had highly uniform levels of UniSp6 expression (mean \pm SD, $C_t = 19.96 \pm 0.87$) and relatively low variability among the C_t values (%CV = 4.36%) indicating good repeatability of sample processing and analysis. Using RT-qPCR with SYBR Green detection to analyse miRNA expression had one major advantage over other techniques: for each sample, amplification product specificity was confirmed by the melting curve analysis. In our study, all the samples had a single melting curve peak at 74 °C, corresponding to a specific qPCR amplification product (Fig. 1). This confirmed that the LNA primers were highly specific and that there were no nonspecific amplification products. The C_t values of the duplicate samples had very low SDs, ranging from 0.00 to 1.03. None of the blank samples showed any amplification.

Expression levels of circulating miRNA across all samples tested

The RT-qPCR results showed that the most abundant miRNA species were miR-16 (mean \pm SD $C_t = 18.67 \pm 1.47$) and miR-21 (20.01 ± 1.61), whereas

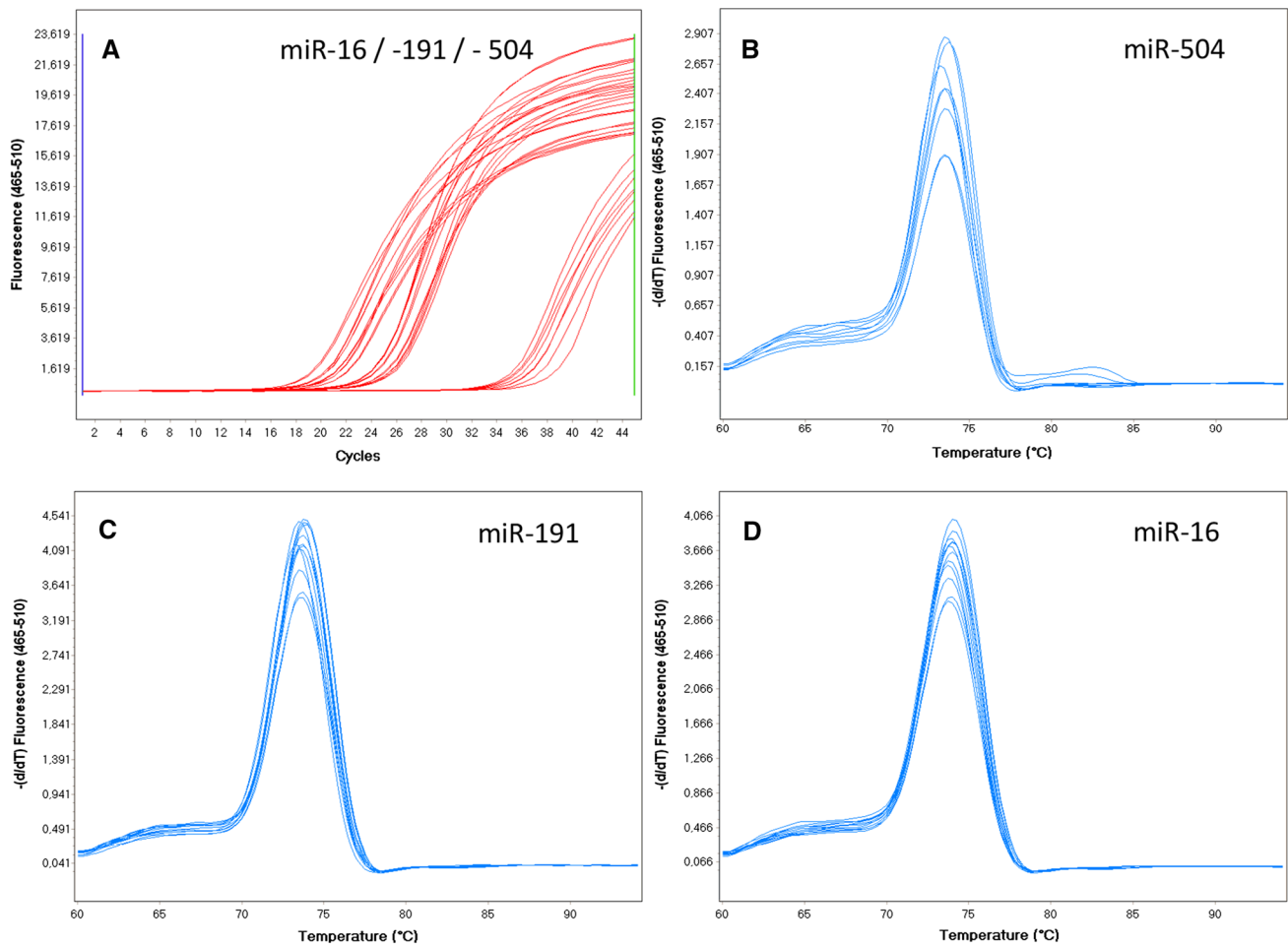


Fig. 1 The quantitative polymerase chain reaction (qPCR) amplification curves (a) and melting curves (b–d) for miR-504, miR-16 and miR-191 are shown. The peak at 74 °C in the melting curve corresponds to a particular qPCR amplification product

miR-504 was the least abundant miRNA species (35.02 ± 1.57) in plasma samples. The expression stability of the miRNA species was evaluated using RefFinder. This comprehensive algorithm, described above, selected miR-191 as the most stable gene (comprehensive ranking value of 1.41) and miR-122, miR-10b, and miR-21 as the least stable genes (comprehensive ranking values of 5.18, 5.48 and 7.00, respectively). MiR-16, miR-195 and miR-504 showed moderate variability of expression across the plasma samples (comprehensive ranking values of 2.06, 2.45 and 2.99, respectively; Fig. 2). The expression levels of five tested miRNAs, miR-195, miR-504, miR-122, miR-10b and miR-21, were then normalised using the two endogenous miRNA controls: miR-191 and miR-16. The mean ΔC_1 values for the miRNAs tested are shown in Table 3. A Shapiro–Wilk analysis confirmed that the variables deviated significantly from a normal distribution ($p < 0.05$). Therefore, the miRNA expression results were further compared using nonparametric Mann–Whitney U and Kruskal–Wallis tests.

Expression levels of circulating miRNA relative to EGFR mutation status

NSCLC patients with *EGFR*-activating mutations in their tumours had significantly higher plasma levels of miR-504 compared to those with wild-type *EGFR*, regardless of the normalisation approach used ($p = 0.0072$ for miR-191-normalised data; $p = 0.0236$ for miR-16-normalised data; Fig. 3). When the NSCLC patients were grouped according to the presence of the *EGFR* mutation in cfDNA, only the miR-191-normalised data showed a significant difference ($p = 0.0122$), whereas there was a clear trend approaching significance for the results normalised using miR-16 ($p = 0.0688$). None of the other miRNAs tested showed significant expression level differences when the two NSCLC patient subgroups were compared.

Fig. 2 The stability of micro-RNA (miRNA) expression across the plasma samples was evaluated by RefFinder. The rankings from the major computational programs (geNorm, NormFinder, BestKeeper, and the comparative $\Delta\Delta C_t$ method) were used to assign an appropriate weight to each individual miRNA and the geometric means of the weights were calculated to generate overall final rankings. The most stable gene was miR-191

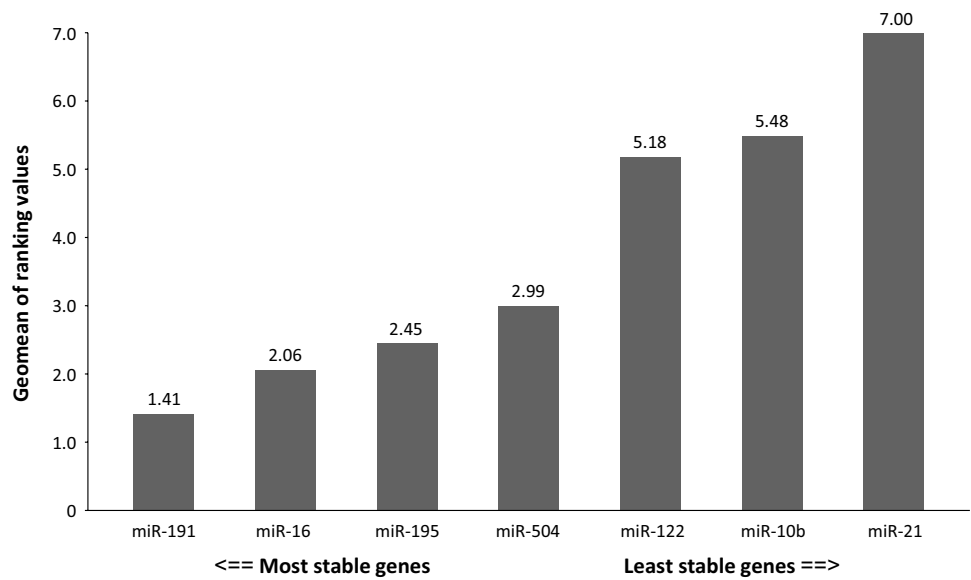


Table 3 The normalised expression levels of circulating microRNAs (miRNAs) in plasma relative to the EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR) and v-Ki-ras2 Kirsten rat sarcoma viral

oncogene homolog (KRAS) mutation status of non-small-cell lung carcinoma (NSCLC) patients and the endogenous control miRNAs used

Candidate biomarker miRNA	Mutant EGFR		Wild-type EGFR		Group comparison (p values)	
	Normalized expression level (mean ± SD ΔC_t)		Normalized expression level (mean ± SD ΔC_t)			
	miR-191	miR-16	miR-191	miR-16	miR-191	miR-16
miR-195	7.55 ± 0.84	11.36 ± 1.13	7.38 ± 0.87	11.30 ± 1.06	0.1448	0.9386
miR-504	13.24 ± 1.40	17.01 ± 1.54	12.15 ± 1.22	16.10 ± 1.13	0.0072	0.0236
miR-122	3.57 ± 2.48	7.36 ± 2.38	2.96 ± 2.08	6.88 ± 2.40	0.1857	0.3965
miR-10b	7.96 ± 1.62	11.63 ± 1.46	7.29 ± 2.06	11.05 ± 2.05	0.1553	0.3887
miR-21	-2.19 ± 0.43	1.49 ± 0.84	-2.33 ± 0.73	1.43 ± 1.03	0.6882	0.7224

Candidate biomarker miRNA	Mutant KRAS		Wild-type EGFR/KRAS		Group comparison (p values)	
	Normalized expression level (mean ± SD ΔC_t)		Normalized expression level (mean ± SD ΔC_t)			
	miR-191	miR-16	miR-191	miR-16	miR-191	miR-16
miR-195	7.68 ± 1.11	11.58 ± 0.76	7.23 ± 0.70	11.15 ± 1.17	0.0854	0.0630
miR-504	11.35 ± 1.18	15.19 ± 1.00	12.43 ± 1.13	16.44 ± 1.00	0.0359	0.0054
miR-122	3.04 ± 0.94	6.94 ± 1.46	2.92 ± 2.50	6.84 ± 2.80	0.4762	0.5901
miR-10b	7.33 ± 1.70	10.95 ± 1.64	7.27 ± 2.26	11.10 ± 2.26	0.8213	0.8757
miR-21	-2.53 ± 0.97	1.08 ± 1.37	-2.22 ± 0.57	1.62 ± 0.78	0.6390	0.4141

A nonparametric Mann–Whitney *U* test was used to compare the mean expression levels among groups. A *p* value < 0.05 was considered statistically significant

Expression levels of circulating miRNA relative to KRAS mutation status

Among the EGFR wild-type patients, 12 had KRAS mutations in their tumours. The same variants were detected in the plasma cfDNA of five of these patients. Plasma miR-504 levels were significantly lower in the KRAS-mutated

group than in patients with EGFR mutations (*p* = 0.0030 for miR-191-normalised data; *p* = 0.0023 for miR-16-normalised data; Fig. 4) and patients with EGFR/KRAS wild-type tumours (*p* = 0.0359 for miR-191-normalised data; *p* = 0.0054 for miR-16-normalised data; Table 3), regardless of the normalisation approach used. Significant differences in miR-504 plasma levels were not attributable

Fig. 3 Circulating miR-504 expression levels in epidermal growth factor receptor (EGFR)-mutated and wild-type EGFR patients with miR-191-based and miR-16-based normalisation (ΔC_t values). The middle mark, box, and whiskers in each plot represent the mean, standard error of the mean, and standard deviation (SD), respectively

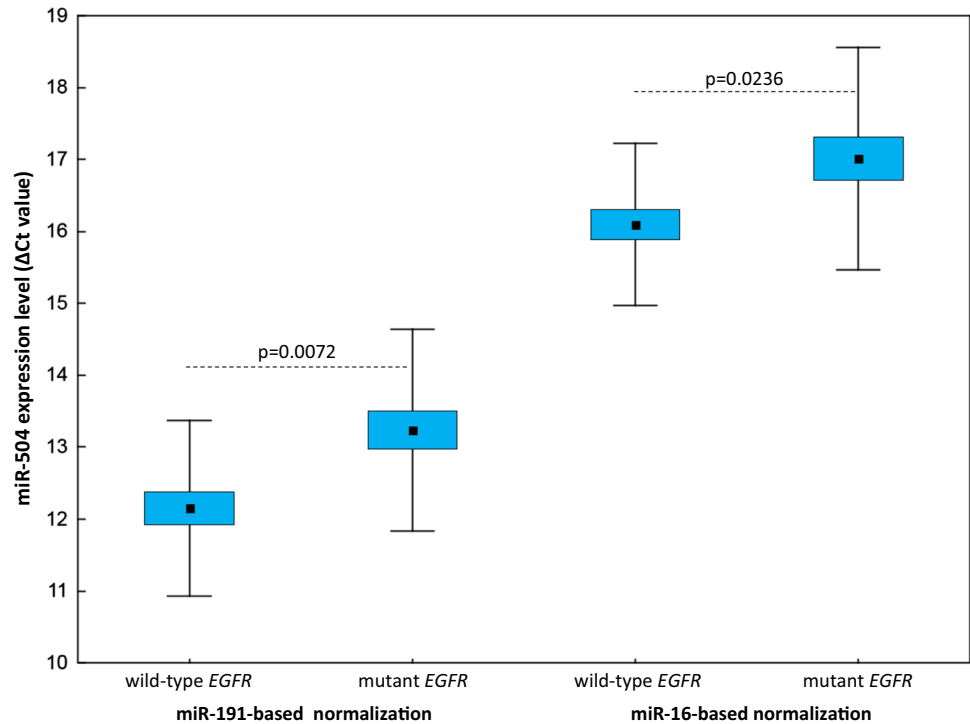
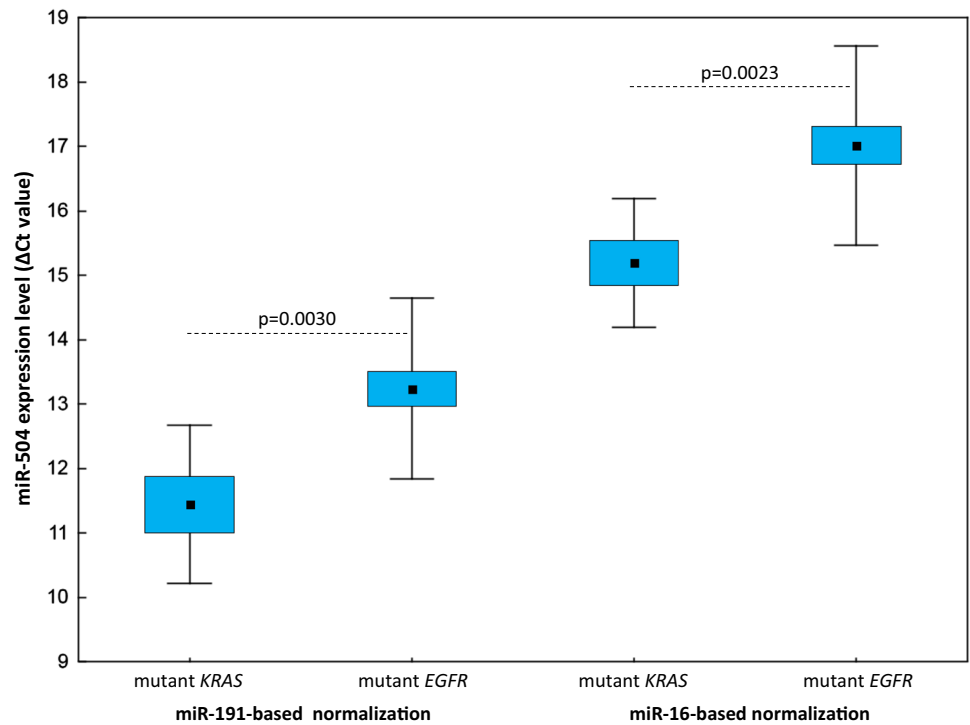


Fig. 4 Circulating miR-504 expression levels in EGFR-mutated versus EGFR wild-type/*Ki-ras2 Kirsten rat sarcoma viral oncogene homolog* (*KRAS*)-mutated patients with miR-191-based and miR-16-based normalisation (ΔC_t values). The middle mark, box, and whiskers in each plot represent the mean, standard error of the mean, and SD, respectively



to the presence of *KRAS*-mutated cfDNA in plasma. None of the other miRNAs tested showed significant expression level differences between NSCLC patient groups

separated according to *KRAS* mutation status. Because only two cases with ALK-positive tumours were identified

in the *EGFR/KRAS* wild-type group, these data were not statistically analysed.

Diagnostic power of circulating miR-504 to predict *EGFR* mutation status

A ROC curve analysis was used to evaluate the diagnostic power of circulating miR-504 levels. This showed that NSCLC patients with *EGFR*-mutated tumours could be distinguished accurately from those with wild-type *EGFR*. The area under the curve (AUC) was 0.75 (95% confidence interval (CI) 0.61–0.87; $p=0.0003$) for the miR-191 normalised

data, and 0.70 (95% CI 0.55–0.82; $p=0.0087$) for the miR-16 normalised data. For the miR-191-based normalisation approach, the test showed 70.37% sensitivity (95% CI 49.8–86.2) and 82.61% specificity (95% CI 61.2–95.0) at the mean $\Delta C_t > 12.74$ cut-off point (Figs. 5a, b). When the type of *EGFR* mutation was taken into account, the circulating miR-504 test showed the greatest power to discriminate between the patients with *EGFR* exon 19 deletions and those with wild-type *EGFR* ($p=0.0290$ for miR-191 normalised data; $p=0.0142$ for miR-16-normalised data; Figs. 5c, d). At the optimal cut-off point of mean $\Delta C_t > 12.74$, the test sensitivity was 73.33% (95% CI 44.9–92.2), the specificity

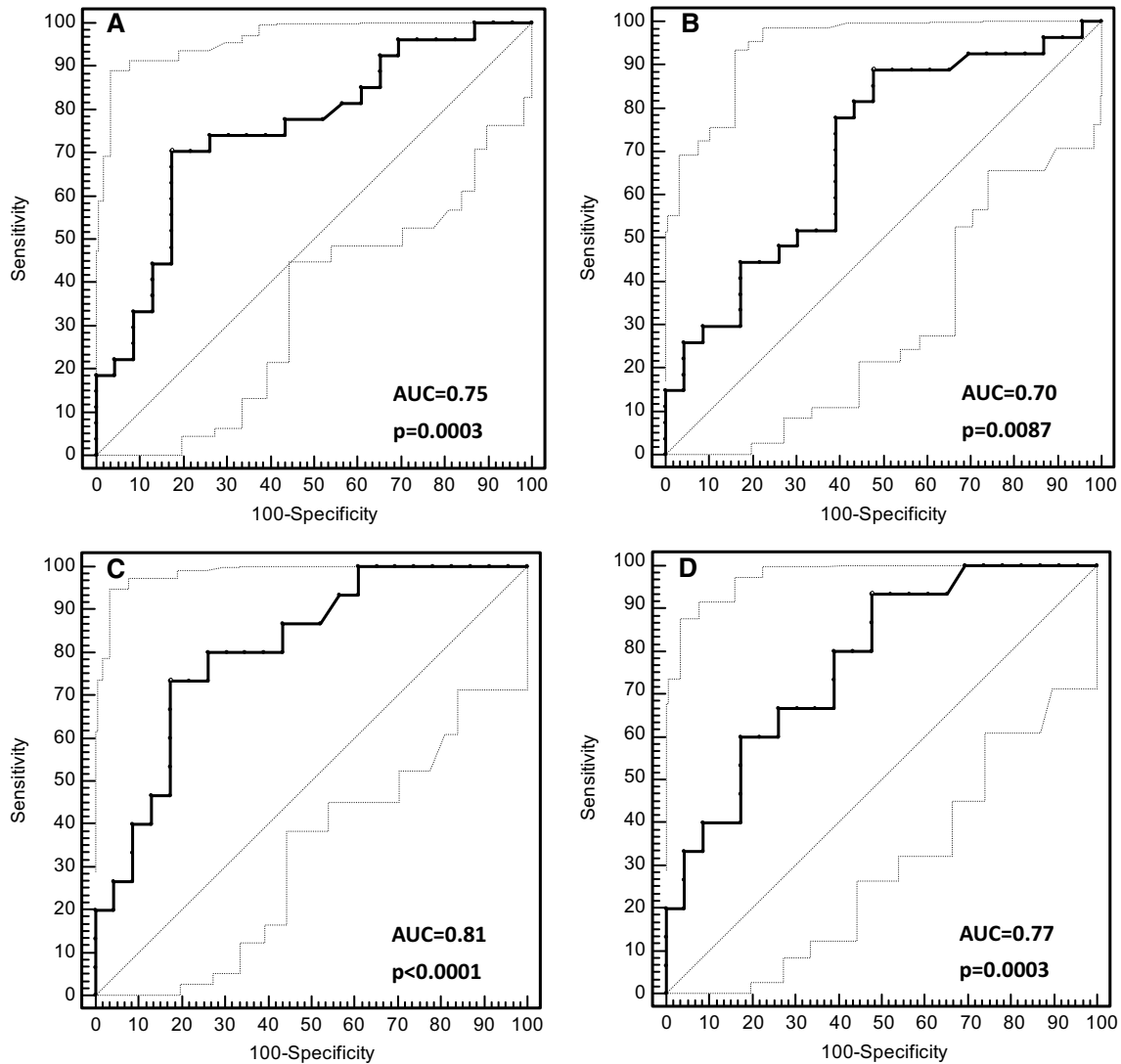


Fig. 5 A receiver operating characteristic (ROC) curve analysis was used to evaluate the diagnostic capacity of circulating miR-504 expression to predict *EGFR* mutation status in non-small-cell lung carcinoma (NSCLC) patients. Plots **a** and **b** show how miR-504 can distinguish NSCLC patients with *EGFR* mutations from those with wild-type *EGFR*, using endogenous control miRNA for data normali-

sation (i.e., miR-191 and miR-16). Plots **c** and **d** show how miR-504 can distinguish NSCLC patients with *EGFR* exon 19 deletions from those with wild-type *EGFR*, using the normalisation approach (i.e., miR-191 and miR-16). The 95% confidence interval for the ROC curve is marked by the dashed grey lines. AUC, area under the curve; P, p value

was 82.61% (95% CI 61.2–95.0) and the AUC was 0.81 ($p < 0.0001$) for the miR-191-normalised data. No significant relationship was observed for the L858R mutation in exon 21 of the *EGFR* gene.

Discussion

In this study, we measured the expression levels of five circulating miRNAs (miR-195, miR-504, miR-122, miR-10b and miR-21) in the plasma of NSCLC patients and evaluated their association with *EGFR* mutation status. These miRNAs had been pre-selected by other researchers as candidate biomarkers capable of distinguishing between mutant and wild-type *EGFR* NSCLC patients. To date, these miRNA signatures have not been revalidated in an independent study using a different sample set and a well-established methodology.

To the best of our knowledge, this is the first study to measure the expression of circulating miR-504 in the plasma of NSCLC patients and evaluate its association with *EGFR* mutation status in tumour tissue and cfDNA. Previously, Gasparini et al. used a NanoString nCounter assay to identify miR-1253, miR-504 and miR-26a-5p [31]. These miRNAs were used to distinguish mutation-free (wild-type) NSCLC from translocated *ALK*-, mutant *EGFR*-, or mutant *KRAS*-driven NSCLC, showing an accuracy of 0.79 (95% CI 0.67–0.88) and a multiclass AUC of 0.692 in a set of 67 FFPE tissue samples from NSCLC patients. The *EGFR*-mutated NSCLC patients presented significantly higher (approx. 2.5-fold) normalised expression level of miR-504 (but not miR-1253 and miR-26a; 4.1) than those with wild-type *EGFR* (1.7; $p < 0.0001$). A total of 24 patients had wild-type *EGFR* and 11 had mutated *EGFR*. In total, 10 of the 11 patients with mutated *EGFR* had deletions in exon 19. The association between miR-504 (but not miR-26a) expression and mutant *EGFR* was then confirmed in a validated cohort of 105 lung adenocarcinoma tissue samples from The Cancer Genome Atlas: a high level of miR-504 expression was associated with *EGFR* mutations (odds ratio = 2.86; 95% CI 1.07–7.71; $p = 0.04$). MiR-1253 was not included in the validation set.

In our study, the *EGFR*-mutated NSCLC patients had significantly higher (twofold) circulating miR-504 expression levels than those with wild-type *EGFR*, regardless of the RT-qPCR data normalisation strategy ($p < 0.05$). The discriminative power of circulating miR-504 was particularly high for a subgroup of NSCLC patients with *EGFR* exon 19 deletions in their tumours (AUC = 0.81; $p < 0.0001$). This finding is particularly important because numerous clinical studies have demonstrated that NSCLC patients with exon 19 deletions showed significantly longer median progression-free survival when treated with EGFR-TKIs compared

to patients with L858R mutations and uncommon or dual mutations [32–35].

In the ROC analysis, circulating miR-504 expression levels showed more than 70% sensitivity and 83% specificity in distinguishing *EGFR*-mutated NSCLC patients from those with wild-type *EGFR*. Obviously, these values are too low for diagnostic application of surrogate miRNA biomarkers. To date, somatic mutations in exons 18–21 of the *EGFR* gene are the only predictive biomarkers for EGFR-TKI treatment outcomes. In our laboratory, the Cobas *EGFR* Mutation Test (Roche Diagnostics GmbH) showed more than 83% sensitivity and 100% specificity for detecting *EGFR* mutations in plasma cfDNA from 50 advanced NSCLC patients [36]. In contrast, we found no *EGFR* mutations in plasma cfDNA from 14 NSCLC patients with resectable mutant *EGFR*-positive tumours, or in 3 of the 17 patients (18%) with metastatic disease whose tumours probably did not shed enough DNA for molecular detection. Nonetheless, five of these patients had miR-504 expression levels that were above the cut-off value of > 12.74 , which corresponds to a mutant *EGFR*-positive result. This clearly demonstrates the diagnostic potential of circulating miRNAs in NSCLC patients to supplement the established algorithm based on the liquid biopsy evaluation, although several specific miRNAs would offer better accuracy than a single biomarker [37, 38]. Blood product fractionation to enrich for tumour-specific miRNAs (i.e., those associated with extracellular vesicles) might further increase the diagnostic power of such biomarkers [39].

Approximately 15–25% of patients with lung adenocarcinoma have tumour-associated *KRAS* mutations [40]. In the majority of cases, these mutations are missense mutations, which involve amino acid substitution at positions 12, 13 or 61, resulting in constitutive activation of *KRAS* signalling pathways. The RAS proteins [*KRAS*, v-Ha-ras Harvey rat sarcoma viral oncogene homolog (*HRAS*) and neuroblastoma RAS viral oncogene homolog (*NRAS*)] are central mediators of growth factor receptor signalling and are, therefore, critical for cell proliferation, survival, and differentiation. RAS can activate several downstream effectors, including the PI3K–AKT–mTOR pathway, which is involved in cell survival, and the RAS–RAF–MEK–ERK pathway, which is involved in cell proliferation [41]. Either *KRAS* mutations are not observed if *EGFR*-activating mutations or *ALK* translocations are present in NSCLC patients [42]. Currently, there are no specific anti-*KRAS* therapies available. However, *KRAS* mutations are negative predictors of the radiographic response to EGFR TKIs; therefore, their molecular analysis is included in the diagnostic algorithm by some clinical laboratories [43]. The original study by Gasparini et al. showed that the tissue expression level of miR-504 was significantly reduced in *KRAS*-mutated tumours (normalised expression level

of 1.6) with respect to *EGFR*-mutated tumours (normalised expression level of 4.1; $p < 0.0001$) [31]. It was also reduced in *EGFR/KRAS* wild-type tumours, although not significantly (normalised expression level of 1.7; $p > 0.05$). In our study, we observed a similar association between the levels of circulating miR-504 and *KRAS* mutation status, again suggesting that circulating miR-504 may be used as a predictive biomarker in NSCLC patients.

Approximately 3–7% of lung tumours, mostly adenocarcinomas, harbour *ALK* fusions [42]. The various N-terminal fusion partners, echinoderm microtubule-associated protein-like 4 (*EML4*), Kinesin Family member 5B (*KIF5B*) and TRK-fused gene (*TFG*) genes in particular, promote dimerisation and, therefore, constitutive activity of *ALK* receptor tyrosine kinase. Signalling downstream of *ALK* fusions results in activation of cellular pathways involved in cell growth and proliferation [44]. *ALK*-rearrangements do not usually overlap with other oncogenic mutations found in NSCLC patients, e.g., *EGFR* or *KRAS* mutations. Clinically, the presence of *EML4-ALK* fusion is associated with resistance to *EGFR*-TKIs [45]. Gasparini et al. found that miR-504 expression in 17 *ALK*-positive lung tumours (normalised expression level of 1.8) was significantly lower than in 11 *EGFR*-mutated tumours (normalised expression level of 4.1; $p < 0.05$) [31]. In our study, we identified only two NSCLC patients with *ALK*-rearrangements among those with *EGFR/KRAS* wild-type tumours, so we could not use these data in statistical analyses. However, the mean expression levels of circulating miR-504 in *ALK*-positive patients (11.41 for miR-191 normalised data; 15.38 for miR-16 normalised data) were clearly lower than in *EGFR*-mutated subjects (13.24 and 17.01, respectively), which is consistent with the trend observed in tumour tissue.

Because few relevant studies have been published, the biological role of miR-504 in NSCLC remains unclear. A recent report suggested that miR-504 may function as a tumour suppressor in NSCLC [46]. MiR-504 was notably downregulated in NSCLC tissues compared to adjacent normal tissues. Lower miR-504 expression positively correlated with lymph node metastasis and advanced TNM stage in patients. Furthermore, upregulation of miR-504 significantly inhibited cell proliferation, cell invasion and epithelial–mesenchymal transition in NSCLC. RT-qPCR, western blotting analysis and luciferase reporter assays confirmed that miR-504 could bind to lysyl oxidase homolog 2 (*LOXL2*) 3' untranslated regions (UTRs) and regulate its expression. Moreover, miR-504 can function as a negative regulator of human p53 by binding to two sites in the p53 3' UTR [47]. Overexpression of miR-504 decreased p53 protein levels and activity in lung cancer cells (including p53 transcriptional activity), p53-mediated apoptosis, and cell cycle arrest in response to stress; it also promoted the formation of tumours

in vivo. In NSCLC cells, loss of p53 is associated with drug resistance to *EGFR* inhibitors and radiation [48].

Interestingly, miR-504 occurs within an intron of the fibroblast growth factor 13 (*FGF13*) gene, a member of the FGF homologous factor family [49]. *FGF13* and miR-504 probably share a common primary transcript. Expression of the *FGF13* locus, including miR-504, was negatively regulated by p53 in NSCLC. The inhibition of miR-504 expression by p53 may define a p53-regulatory negative feedback loop, possibly boosting p53 protein levels further in response to p53-activation signals.

MiR-504 is misregulated in a variety of cancer types, functioning as an oncogenic miRNA or a tumour-suppressive miRNA. In human glioma, miR-504 suppressed cell proliferation and induced cell apoptosis by targeting Forkhead Box P1 (*FOXP1*) [50]. In osteosarcoma, miR-504 directly targeted tumour protein p53-inducible nuclear protein 1 (*TP53INP1*), which enhanced osteosarcoma growth and promoted distant metastases [51]. In addition, miR-504 expression levels are frequently downregulated in both hepatocellular carcinoma (HCC) tissues and cell lines [52]. MiR-504 also functions as a tumour-suppressive miRNA that inhibits the proliferation and invasion of HCC cells by targeting Frizzled-7 (*FZD7*) and inhibiting Wnt/ β -catenin signalling. Bioinformatic analyses and dual-luciferase reporter assays confirmed that miR-504 directly targets the 3' UTR of *FZD7* mRNA. Furthermore, increasing evidence suggests that the Wnt/ β -catenin pathway is abnormally activated in NSCLC and may be an important mediator of drug resistance to *EGFR* inhibition. β -catenin levels were increased in NSCLC cells with oncogenic *EGFR* mutations, as well as in gefitinib-resistant cells, and inhibition of the Wnt/ β -catenin pathway re-sensitised cells to *EGFR* inhibitors and increased their efficacy in these tumours [53–55]. Recently, Jiang et al. reported that *EGFR*-activating mutations in NSCLC co-occurred with mutations in genes participating in most key signalling pathways and biological processes, including receptors of different classes, key regulators involved in genome and epigenome stability, the PI3K–AKT–mTOR pathway, and the TP53/apoptosis pathway [56]. In the study by Jiang et al., many of the Wnt/ β -catenin pathway-related genes linked to oncogenic *EGFR*s were mutated, and there was more variation in the Wnt/ β -catenin pathway than in the TP53/apoptosis and PI3K–AKT–mTOR pathways. Mutations in *RB1*, *JAK2*, *APC*, *JAK3*, *NF1* and *SMAD4* were predominantly associated with activated mutant *EGFR*s.

A mature miRNA can guide the miRNA-induced silencing complex (miRISC) to targets using sequence complementarity between the miRNA and sequences in the 3' UTRs of cognate mRNAs. However, experiments have shown that targeting can also occur in the 5' UTRs and open reading frames (ORFs) [57]. Hafner et al. found that among the exonic target regions, 50% of the target sites

corresponded to coding sequences (CDSs), whereas only 46% corresponded to 3' UTRs [58]. Chi et al. applied a high-throughput approach to isolate Argonaute-bound target sites, which indicated that target sites in the CDSs were as numerous as those located in the 3' UTRs [59]. Canonical mRNA–miRNA interactions can occur at the 5' region of the miRNA, called the seed (nucleotides 2–7). However, biologically important sites that do not respect canonical seed base pairing have been discovered for a number of targets [60]. Recent publications have highlighted both the abundance and functional importance of non-canonical 'seedless' miRNA binding sites in transcriptomes. Because of the degenerate nature of miRNAs outside the seed, the critical determinants of non-canonical sites are probably specific to individual miRNAs/target sites.

Accordingly, we analysed the *EGFR* gene sequence for miR-504 binding sites using the STarMirDB database (<http://sfold.wadsworth.org/starmirDB.php>), which can identify seeded and seedless sites in the 3' UTRs, CDSs and 5' UTRs of a gene [61]. The software predicted multiple binding sites within the 3' UTR sequence (seven seedless sites), CDS (six seeded sites) and 5' UTR sequence (four seedless sites). Interestingly, many single nucleotide variants were found within the predicted binding sites in *EGFR* sequence using the Ensembl genome database (<http://www.ensembl.org>), which makes the regulation of *EGFR* by miR-504 even more complicated. No canonical binding site for the miR-504 seed was found within the *EGFR* 3' UTR, and no seedless binding within exons 18–21 of the *EGFR* gene was suggested by the bioinformatic analysis. There is strong experimental evidence that some genes are directly targeted by miR-504, including *TP53*, *TP53INP1*, *MDM2*, *VEGFA*, *BAX* and *CDKN1A*. The implications of this for the *EGFR*-dependent signalling pathways are well documented (miR-TarBase, Release 7.0: 15 September 2017).

Although correlations had been reported by previous investigators, we found no significant relationships between *EGFR* mutation status and the expression levels of the other circulating miRNAs tested: miR-195, miR-122, miR-10b and miR-21. Zhang et al. showed that miR-122 was expressed differently in wild-type ($n=48$) compared to mutant *EGFR* carriers ($n=57$) with NSCLC (most cases were at the resectable stage; $p=0.018$) [16]. Overall survival in patients, especially those with advanced cancer and *EGFR* mutations, was associated with plasma miR-195 levels and miR-122 expression. Zhao et al. showed that upregulation of miR-195, miR-122 and miR-21 was significantly associated with *EGFR* mutations in both tumour tissues and plasma from 149 radically resected NSCLC patients (all $p=0.000$) [18]. The ROC analysis showed that the AUC values for detecting the *EGFR* mutation were 0.7 ($p=0.000$; 95% CI 0.618–0.782) for miR-195, 0.733 ($p=0.000$; 95% CI 0.655–0.812) for miR-122, and 0.8 ($p=0.000$; 95% CI 0.730–0.870) for miR-21. The

AUC of an optimum combination of four plasma miRNAs (miR-195, miR-122, miR-125 and miR-21) was 0.869 ($p=0.000$; 95% CI 0.808–0.930). Shen et al. showed that the expression levels of circulating miR-21 and miR-10b were significantly higher in radically resected NSCLC patients with *EGFR* mutations ($n=60$) than in those without such mutations ($n=68$; $p<0.001$) [17].

Currently, there is no standard endogenous control for normalising circulating miRNA levels in the blood. In our study, the expression of circulating miRNAs was normalised to the levels of endogenous control miRNAs, miR-191 and miR-16. Using the comprehensive algorithm RefFinder, described above, we identified miR-191 as the most stable gene (comprehensive stability value = 1.41), followed by miR-16, which showed moderate stability (comprehensive stability value = 2.06) among all the miRNAs tested. Pelletier and Latham generated miRNA microarray data from dozens of normal and diseased human tissues and revealed ubiquitous and stably expressed normalisation candidates to evaluate using qRT-PCR [62]. In their study, miR-191 showed the highest consistency in its expression level across 13 normal tissues and five pairs of lung cancer/normal adjacent tissues. This miRNAs was statistically superior to the reference RNAs most commonly used in miRNA qRT-PCR experiments, such as 5S rRNA, U6 snRNA, or total RNA. Hu et al. selected miR-484 and miR-191-5p as the most stably expressed miRNAs and successfully used them as normalizers in serum from breast cancer patients [63].

MiR-16 was also used as an endogenous control to normalise expression data in many previous studies. In our study, the miR-16-normalised and miR-191-normalised expression data were similarly associated with *EGFR* mutation status. McDermott et al. demonstrated that combining miR-16 and miR-425 generated more reliable results than using either of these miRNAs alone, or using RNU6 from blood specimens collected from women with breast cancer and healthy volunteers [64]. Zhang et al. measured the levels of miR-16 and RNU6B using RT-qPCR in plasma samples from 20 NSCLC patients and 20 healthy controls and observed no significant differences between the two groups (miR-16, $p=0.158$; RNU6B, $p=0.557$) [65]. However, when they examined the stability of the two endogenous controls, there were no significant differences in miR-16 expression among the plasma samples kept at room temperature for 0, 2, 4 and 8 h ($p>0.05$), whereas the expression levels of RNU6B were significantly different after 4 and 8 h ($p<0.05$). The results from other studies that have analysed circulating miRNAs in patients with different tumour types also suggest that miR-16 is a good endogenous control [66–68]. However, when selecting a miRNA for normalisation, it is crucial to realise that there can be disease-specific differences in expression patterns because of biological variability among different tumour types [69].

Zhang et al. normalised the expression levels of circulating miR-122 and miR-195 using the exogenous spike-in control cel-miR-39 (synthetic *Caenorhabditis elegans* miRNA) [16]. This spike-in method can eliminate some deviations of the experimental process and make the results more reliable. However, this method does not correct for differences in sampling methods and the quality of samples [70]. A major drawback of using spike-in controls is that although the handling of experiments is taken into account, the quality of tissues, body fluids and extracellular vesicle samples are not. Sample age, body fluid collection procedures, experimental preparation, and storage procedures may produce changes in miRNA levels caused by cell lysis or miRNA degradation.

Zhao et al. and Shen et al. normalised the expression of target miRNAs relative to the expression of small noncoding RNA RNU6 genes, which are commonly used as internal controls for miRNA quantification assays [17, 18]. However, RNU6 is not an miRNA and does not reflect the biochemical character of miRNA molecules in terms of their transcription, processing, or tissue-specific expression patterns [70]. In addition, the efficiency of RNU6 extraction, reverse transcription, and PCR amplification may differ from that of miRNAs. Therefore, it may be best to normalise miRNAs using genes from the same RNA class, i.e., miRNAs [71].

The use of different miRNA normalisation strategies may partially explain discrepancies between our study and the three reports described above, concerning the expression of miR-195, miR-122, miR-10b and miR-21 in the plasma of *EGFR*-mutated versus wild-type NSCLC patients. Another important methodological consideration is haemolysis in the plasma samples. We ensured that none of our plasma samples showed signs of haemolysis using a two-step approach that included visual inspection and measuring the relative expression of the erythrocyte-specific miR-451 and the stable miR-23a.

The effects of haemolysis on the expression of circulating miRNAs have been neglected in many studies. Recent studies have shown that haemolysis during blood collection can have a substantial impact on plasma/serum miRNA content. Kirschner et al. profiled the miRNA content of haemolysed and non-haemolysed plasma and red blood cells to generate a profile of circulatory miRNAs that were affected by haemolysis [72]. They found that miR-16, miR-195, miR-122 and miR-21 were likely to be affected by haemolysis, particularly in plasma from cancer patients. In contrast, miR-191 was unaffected by haemolysis in plasma from healthy individuals and cancer patients. Therefore, rigorous quality control of plasma/serum samples is a critical step before measuring circulating miRNA expression levels. Accordingly, various methods may be used to assess the level of haemolysis in plasma/serum, including simple visual inspection (i.e., identification of a visible pink colour), measuring haemoglobin using a biochemical analyser, measuring

haemoglobin by spectrophotometry at 414 nm, and comparing the ratio of red blood cell-enriched miR451a with the stable miR-23a-3p [73]. Measuring the relative expression of miR-451 and miR-23a proved to be the most sensitive method and could detect 0.001% haemolysis in serum.

Extensive variation among pre-analytical conditions and detection methods may account for conflicting results with specific miRNA biomarkers. Haemolysis and data normalisation are only two of the factors affecting the reliability and reproducibility of results pertaining to circulating miRNA expression. Other important factors include blood specimen collection procedures, the effects of fasting and the timing of blood sample collection, blood processing, plasma/serum storage times, miRNA extraction techniques, and miRNA detection methods [74–76].

Conclusions

MiR-504 expression is misregulated in human cancers, including NSCLC. In this study, for the first time, circulating miR-504 levels in plasma from NSCLC patients were measured and compared to previously published results on differential miR-504 expression in tumour tissue relative to the *EGFR* mutation status of NSCLC patients. Here, we demonstrated the feasibility and potential diagnostic value of circulating miR-504 expression in plasma to discriminate between *EGFR*-mutated and *EGFR* wild-type NSCLC, and also between *EGFR*-mutated and *KRAS*-mutated tumours. Significantly, the greatest discriminative power was achieved for NSCLC patients with *EGFR* exon 19 deletions, which are the strongest predictors of EGFR TKI efficacy. In our analysis, the sample quality control and normalisation strategies for the RT-qPCR data were the key to produce reliable results and evaluating the miRNAs. Consequently, we did not confirm the diagnostic capacity of circulating miR-195, miR-122, miR-10b and miR-21 for discriminating between *EGFR*-mutated and wild-type NSCLC, as reported by other groups.

Further research is needed to confirm the utility of miRNAs, either as predictive biomarkers or therapeutic targets for NSCLC. Currently, there is little or no consensus among published miRNA data due to a lack of standardised protocols for miRNA expression analyses and insufficient attention to pre-analytical variables and quality control. Therefore, we believe it is important to reproduce and revalidate the miRNA expression data published by other groups in independent studies.

Acknowledgements This study was supported by the Research Grant of Polish Ministry of Science and Higher Education for young investigators, and by the National Research Institute of Tuberculosis and Lung Diseases in Warsaw (statutory activity project no. 3/24), and by the research grant from The Count Jakub Potocki Foundation, Poland.

We thank to the Team of the Roche Diagnostics Poland who supported our research by reagents for *KRAS* mutation analysis. We thank the patients who participated in the study.

Author contributions ASz designed the study. MB, GCz, PR, WK, EW, ABW, TO, JCW recruited NSCLC patients and evaluated medical data. RL and DG provided pathological diagnosis. ASz, MF, KD, AZ, SR carried out the experiments. ASz, MF and JCW discussed the results. ASz and MF prepared the manuscript, and they revised it with JCW. All authors have given approval to the final version of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Castellanos E, Feld E, Horn L (2017) Driven by mutations: the predictive value of mutation subtype in EGFR-mutated non-small cell lung cancer. *J Thorac Oncol* 12:612–623
- Singh M, Jadhav HR (2018) Targeting non-small cell lung cancer with small-molecule EGFR tyrosine kinase inhibitors. *Drug Discov Today* 23:745–753
- Lee DH (2017) Treatments for EGFR-mutant non-small cell lung cancer (NSCLC): the road to a success, paved with failures. *Pharmacol Ther* 174:1–21
- Chorostowska-Wynimko J, Szpechcinski A (2007) The impact of genetic markers on the diagnosis of lung cancer: a current perspective. *J Thorac Oncol* 2:1044–1051
- Han X et al (2014) Suitability of surgical tumor tissues, biopsy, or cytology samples for epidermal growth factor receptor mutation testing in non-small cell lung carcinoma based on chinese population. *Transl Oncol* 7:795–799
- Mok TS et al (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361:947–957
- Costa DB, Kobayashi S, Tenen DG, Huberman MS (2007) Pooled analysis of the prospective trials of gefitinib monotherapy for EGFR-mutant non-small cell lung cancers. *Lung Cancer* 58:95–103
- Kim ES et al (2008) Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (INTEREST): a randomised phase III trial. *Lancet* 372:1809–1818
- Rolfo C et al (2018) Liquid biopsy for advanced non-small cell lung cancer (NSCLC): a statement paper from the IASLC. *J Thorac Oncol* 13:1248–1268
- Diaz LA Jr, Bardelli A (2014) Liquid biopsies: genotyping circulating tumor DNA. *J Clin Oncol* 32:579–586
- Karachaliou N, Sosa AE, Molina MA, Centelles Ruiz M, Rosell R (2017) Possible application of circulating free tumor DNA in non-small cell lung cancer patients. *J Thorac Dis* 9:S1364–S1372
- Li Z, Zhang Y, Bao W, Jiang C (2014) Insufficiency of peripheral blood as a substitute tissue for detecting EGFR mutations in lung cancer: a meta-analysis. *Target Oncol* 9:381–388
- Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP (2008) The impact of microRNAs on protein output. *Nature* 455:64–71
- Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RH, Cuppen E (2005) Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* 120:21–24
- Florczyk M, Szpechcinski A, Chorostowska-Wynimko J (2017) miRNAs as biomarkers and therapeutic targets in non-small cell lung cancer: current perspectives. *Target Oncol* 12:179–200
- Zhang H, Su Y, Xu F, Kong J, Yu H, Qian B (2013) Circulating microRNAs in relation to EGFR status and survival of lung adenocarcinoma in female non-smokers. *PLoS One* 8:e81408
- Shen Y, Tang D, Yao R, Wang M, Wang Y, Yao Y, Li X, Zhang H (2013) microRNA expression profiles associated with survival, disease progression, and response to gefitinib in completely resected non-small-cell lung cancer with EGFR mutation. *Med Oncol* 30:750
- Zhao Q et al (2015) Circulating miRNAs is a potential marker for gefitinib sensitivity and correlation with EGFR mutational status in human lung cancers. *Am J Cancer Res* 5:1692–1705
- Khan J, Lieberman JA, Lockwood CM (2017) Variability in, variability out: best practice recommendations to standardize pre-analytical variables in the detection of circulating and tissue microRNAs. *Clin Chem Lab Med* 55:608–621
- Dugan L, Leech L, Speroni KG, Corriher J (2005) Factors affecting hemolysis rates in blood samples drawn from newly placed IV sites in the emergency department. *J Emerg Nurs* 31:338–345
- Blondal T, Jensby Nielsen S, Baker A, Andreassen D, Mouritzen P, Wrang Teilm M, Dahlsveen IK (2013) Assessing sample and miRNA profile quality in serum and plasma or other biofluids. *Methods* 59:S1–S6
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034
- Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64:5245–5250
- Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnol Lett* 26:509–515
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ Method. *Methods* 25:402–408
- Xie F, Xiao P, Chen D, Xu L, Zhang B (2012) miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol Biol* 80:75–84
- Metz CE (1978) Basic principles of ROC analysis. *Semin Nucl Med* 8:283–298
- Zweig MH, Campbell G (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 39:561–577
- Travis WD, Brambilla E, Burke AP, Marx A, Nicholson AG (2015) WHO classification of tumours of the lung, pleura, thymus and heart. International Agency for Research on Cancer, Lyon
- Sobin LH, Gospodarowicz MK, Wittekind C (2009) UICC TNM classification of malignant tumours, 7th edn. Wiley-Liss, New York
- Gasparini P et al (2015) microRNA classifiers are powerful diagnostic/prognostic tools in ALK-, EGFR-, and KRAS-driven lung cancers. *Proc Natl Acad Sci USA* 112:14924–14929
- Choi YW, Choi JH (2017) Does the efficacy of epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor differ according to the type of EGFR mutation in non-small cell lung cancer? *Korean J Intern Med* 32:422–428
- Sun JM et al (2011) The different efficacy of gefitinib or erlotinib according to epidermal growth factor receptor exon 19 and exon 21 mutations in Korean non-small cell lung cancer patients. *J Cancer Res Clin Oncol* 137:687–694

34. Choi YW et al (2018) EGFR exon 19 deletion is associated with favorable overall survival after first-line gefitinib therapy in advanced non-small cell lung cancer patients. *Am J Clin Oncol* 41:385–390
35. Krawczyk P et al (2017) Comparison of the effectiveness of erlotinib, gefitinib, and afatinib for treatment of non-small cell lung cancer in patients with common and rare EGFR gene mutations. *Oncol Lett* 13:4433–4444
36. Szepechinski A, Duk K, Wojcik P, Kupis W, Rudzinski P, Bryl M, Czyzewicz G, Orlowski T, Chorostowska-Wynimko J (2017) P3.01-041 The robustness of allele-specific qPCR assays for detection of EGFR mutations in plasma cell-free DNA from NSCLC patients. *J Thorac Oncol* 12:S2216
37. He WJ, Li WH, Jiang B, Wang YF, Xia YX, Wang L (2015) MicroRNAs level as an initial screening method for early-stage lung cancer: a bivariate diagnostic random-effects meta-analysis. *Int J Clin Exp Med* 8:12317–12326
38. Wang H, Wu S, Zhao L, Zhao J, Liu J, Wang Z (2015) Clinical use of microRNAs as potential non-invasive biomarkers for detecting non-small cell lung cancer: a meta-analysis. *Respirology* 20:56–65
39. Witwer KW (2015) Circulating microRNA biomarker studies: pitfalls and potential solutions. *Clin Chem* 61:56–63
40. Roman M, Baraibar I, Lopez I, Nadal E, Rolfo C, Vicent S, Gil-Bazo I (2018) KRAS oncogene in non-small cell lung cancer: clinical perspectives on the treatment of an old target. *Mol Cancer* 17:33
41. Ferrer I, Zugazagoitia J, Hertzberg S, John W, Paz-Ares L, Schmid-Bindert G (2018) KRAS-Mutant non-small cell lung cancer: from biology to therapy. *Lung Cancer* 124:53–64
42. Mayekar MK, Bivona TG (2017) Current landscape of targeted therapy in lung cancer. *Clin Pharmacol Ther* 102:757–764
43. Lindeman NI et al (2018) Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. *Arch Pathol Lab Med* 142:321–346
44. Golding B, Luu A, Jones R, Vilorio-Petit AM (2018) The function and therapeutic targeting of anaplastic lymphoma kinase (ALK) in non-small cell lung cancer (NSCLC). *Mol Cancer* 17:52
45. Lo Russo G et al (2017) Concomitant EML4-ALK rearrangement and EGFR mutation in non-small cell lung cancer patients: a literature review of 100 cases. *Oncotarget* 8:59889–59900
46. Ye MF, Zhang JG, Guo TX, Pan XJ (2018) MiR-504 inhibits cell proliferation and invasion by targeting LOXL2 in non small cell lung cancer. *Biomed Pharmacother* 97:1289–1295
47. Hu W et al (2010) Negative regulation of tumor suppressor p53 by microRNA miR-504. *Mol Cell* 38:689–699
48. Huang S, Benavente S, Armstrong EA, Li C, Wheeler DL, Harari PM (2011) p53 modulates acquired resistance to EGFR inhibitors and radiation. *Cancer Res* 71:7071–7079
49. Bublik DR et al (2017) Regulatory module involving FGF13, miR-504, and p53 regulates ribosomal biogenesis and supports cancer cell survival. *Proc Natl Acad Sci USA* 114:E496–E505
50. Cui R et al (2016) A tumor-suppressive microRNA, miR-504, inhibits cell proliferation and promotes apoptosis by targeting FOXP1 in human glioma. *Cancer Lett* 374:1–11
51. Cai Q, Zeng S, Dai X, Wu J, Ma W (2017) miR-504 promotes tumour growth and metastasis in human osteosarcoma by targeting TP53INP1. *Oncol Rep* 38:2993–3000
52. Quan H, Li B, Yang J (2018) MicroRNA-504 functions as a tumor suppressor in hepatocellular carcinoma through inhibiting Frizzled-7-mediated-Wnt/beta-catenin signaling. *Biomed Pharmacother* 107:754–762
53. Togashi Y et al (2015) Inhibition of beta-Catenin enhances the anticancer effect of irreversible EGFR-TKI in EGFR-mutated non-small-cell lung cancer with a T790M mutation. *J Thorac Oncol* 10:93–101
54. Nakata A et al (2015) Elevated beta-catenin pathway as a novel target for patients with resistance to EGF receptor targeting drugs. *Sci Rep* 5:13076
55. Casas-Selves M et al (2012) Tankyrase and the canonical Wnt pathway protect lung cancer cells from EGFR inhibition. *Cancer Res* 72:4154–4164
56. Jiang J, Protopopov A, Sun R, Lyle S, Russell M (2018) Genomic profiling on an unselected solid tumor population reveals a highly mutated Wnt/beta-catenin pathway associated with oncogenic EGFR mutations. *J Pers Med* 8:E13
57. Xu W, San Lucas A, Wang Z, Liu Y (2014) Identifying microRNA targets in different gene regions. *BMC Bioinform* 15(Suppl 7):S4
58. Hafner M et al (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141:129–141
59. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. *Nature* 460:479–486
60. Flamand MN, Gan HH, Mayya VK, Gunsalus KC, Duchaine TF (2017) A non-canonical site reveals the cooperative mechanisms of microRNA-mediated silencing. *Nucleic Acids Res* 45:7212–7225
61. Rennie W et al (2016) STarMirDB: a database of microRNA binding sites. *RNA Biol* 13:554–560
62. Peltier HJ, Latham GJ (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* 14:844–852
63. Hu Z et al (2012) Serum microRNA profiling and breast cancer risk: the use of miR-484/191 as endogenous controls. *Carcinogenesis* 33:828–834
64. McDermott AM, Kerin MJ, Miller N (2013) Identification and validation of miRNAs as endogenous controls for RQ-PCR in blood specimens for breast cancer studies. *PLoS One* 8:e83718
65. Zhang H, Mao F, Shen T, Luo Q, Ding Z, Qian L, Huang J (2017) Plasma miR-145, miR-20a, miR-21 and miR-223 as novel biomarkers for screening early-stage non-small cell lung cancer. *Oncol Lett* 13:669–676
66. Wang B, Zhang Q (2012) The expression and clinical significance of circulating microRNA-21 in serum of five solid tumors. *J Cancer Res Clin Oncol* 138:1659–1666
67. Liu GH, Zhou ZG, Chen R, Wang MJ, Zhou B, Li Y, Sun XF (2013) Serum miR-21 and miR-92a as biomarkers in the diagnosis and prognosis of colorectal cancer. *Tumour Biol* 34:2175–2181
68. Summerer I et al (2015) Circulating microRNAs as prognostic therapy biomarkers in head and neck cancer patients. *Br J Cancer* 113:76–82
69. Poel D, Buffart TE, Oosterling-Jansen J, Verheul HM, Voortman J (2018) Evaluation of several methodological challenges in circulating miRNA qPCR studies in patients with head and neck cancer. *Exp Mol Med* 50:e454
70. Schwarzenbach H, da Silva AM, Calin G, Pantel K (2015) Data normalization strategies for microRNA quantification. *Clin Chem* 61:1333–1342
71. Chugh P, Dittmer DP (2012) Potential pitfalls in microRNA profiling. *Wiley Interdiscip Rev RNA* 3:601–616
72. Kirschner MB, Edelman JJ, Kao SC, Valley MP, van Zandwijk N, Reid G (2013) The impact of hemolysis on cell-free microRNA biomarkers. *Front Genet* 4:94
73. Shah JS, Soon PS, Marsh DJ (2016) Comparison of methodologies to detect low levels of hemolysis in serum for accurate assessment of serum microRNAs. *PLoS One* 11:e0153200
74. Zampetaki A, Mayr M (2012) Analytical challenges and technical limitations in assessing circulating miRNAs. *Thromb Haemost* 108:592–598

75. Moldovan L, Batte KE, Trgovcich J, Wisler J, Marsh CB, Piper M (2014) Methodological challenges in utilizing miRNAs as circulating biomarkers. *J Cell Mol Med* 18:371–390
76. Farina NH, Wood ME, Perrapato SD, Francklyn CS, Stein GS, Stein JL, Lian JB (2014) Standardizing analysis of circulating microRNA: clinical and biological relevance. *J Cell Biochem* 115:805–811

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.