Doi: 10.31276/VJSTE.60(4).31-38

Establishing a molecular protocol for detection of *EGFR* mutations in patients with non-small cell lung cancer

Thi Minh Ly Nguyen[1,2*], Quoc Dang Quan[3], Thi Bach Hue Nguyen[2], Hoang Chuong Nguyen[2]

[1]*Gentis Joint Stock Company*
[2]*University of Science, Vietnam National University, Ho Chi Minh city*
[3]*Center of Science and Technology Development, Ho Chi Minh Communist Youth Union*

Received 23 April 2018; accepted 28 September 2018

Abstract:

Lung cancer is one of the quickest and most fatal diseases and is divided into two types: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Approximately 50% of NSCLC patients harbour mutations in the epidermal growth factor receptor (*EGFR*) gene, spanning from exon 18 to exon 21, which are responsible for the tyrosine kinase domain of the *EGFR* protein. Currently, gefitinib and erlotinib are two drugs inhibiting EGFR which are used in the treatment of lung cancer. The indication for use of these drugs depends on the mutation status of *EGFR*. Resistance to treatment with these two drugs has been found with mutations in exon 20; however, mutations in exon 18, 19, 21 benefited from the same treatment. Therefore, we built a molecular protocol for detecting mutations in exon 21 of *EGFR* in NSCLC patients in Vietnam which, in this study, is based on the real-time polymerase chain reaction high resolution melting (PCR HRM) technique. We successfully designed pairs of primers for the real-time PCR HRM technique to detect mutations in exon 21 and for the Sanger sequencing method to confirm the real-time PCR HRM results. We also investigated the optimal concentration of $MgCl_2$ for the real-time PCR HRM protocol. Performance characteristics of the real-time PCR HRM protocol were evaluated in terms of specificity, limit of detection, repeatability, and it showed good results. Finally, we applied the molecular protocol to 172 clinical samples and compared these with the Sanger nucleotide sequencing method and the peptide nucleic acid (PNA) clamping method. The results of the comparison demonstrate that the molecular protocol constructed for detection of exon 21 mutations can be used in the treatment of NSCLC patients.

Keywords: ***EGFR*, exon 21, mutations, PNA Clamp.**

Classification numbers: **3.2, 3.5**

Introduction

Lung cancer is the leading cause of cancer mortality worldwide. This is especially the case in Vietnam. Lung cancer comprises two types: NSCLC, which accounts for approximately 80-85% of lung cancer patients, and SCLC, which accounts for approximately 15-20% of lung cancer patients [1]. NSCLC has a poor prognosis if it is diagnosed at an advanced stage.

EGFR is a member of the ErbB receptor tyrosine kinase family; it has an extracellular ligand-binding domain, a transmembrane domain, and an intracellular tyrosine kinase. *EGFR* (also called ErbB1, Her1) is a transmembrane receptor tyrosine kinase that transduces signals that are critical for cell proliferation, differentiation, and apoptosis [2]. Overexpression of *EGFR* has been observed in tumours in more than 60% of metastatic NSCLC patients and is correlated with a poor prognosis [3]. Overexpression or mutational activation of *EGFR* is associated with the development and progression of numerous human malignancies. Drugs as tyrosine kinase inhibitors (TKIs) that target the ATP-binding cleft of *EGFR* have been developed for the treatment of NSCLC patients with *EGFR* mutations [4].

EGFR is located on chromosome 7 at the 7p12 locus and it is in the proto-oncogene group. *EGFR* is 110 kb long, comprising 28 exons [5]. *EGFR* mutations associated with

*Corresponding author: Email: nguyenminhly76@gmail.com

NSCLC belong to four exons, from 18 to 21. Mutations in *EGFR* cause the *EGFR* protein to exist in a non-ligand-activated state that increases the sensitivity of the tumour or resistance to *EGFR* TKIs.

In the last few years, *EGFR* receptor inhibitor drugs such as erlotinib (Tarceva) and gefitinib (Iressa) have been used for the treatment of NSCLC and have been shown to have encouraging effects [6]. These two drugs inhibit *EGFR* autophosphorylation by inhibiting ATP binding to *EGFR* and reducing the affinity between *EGFR* and its ligand. Patients treated with gefitinib and erlotinib show good progression, such as significant tumour size reduction, longer life duration, and fewer side effects than conventional chemotherapy or radiation. The level of treatment response to the drug depends on the presence or absence of mutations in *EGFR*; patients should be tested for *EGFR* mutations prior to the administration of targeted therapies.

It has been reported that the frequency of *EGFR* mutations in NSCLC patients is dependent on ethnicity, with a rate of 54.1% found in Asian populations [7]. Vietnam is one of the countries with a high rate of *EGFR* mutations and therefore *EGFR* mutation assessment before treatment is essential for successful treatment. In Vietnam, *EGFR* mutations are detected using Sanger sequencing or commercialized molecular kits. The sequencing method is not appropriate to detect gene mutations with a low frequency. On the other hand, commercialized kits are expensive and involve a complicated process. For example, the PNA clamping kit costs 70,000,000 VND for 50 reactions (about 1,400,000 VND per reaction), while the chemical reagents for the real-time PCR HRM assay cost approximately 120,000 VND per reaction. According to the manual, 'PNAClamp™ *EGFR* Mutation Detection Kit Ver. 2', the PNA clamping kit involves a complicated process. Real-time PCR HRM is a powerful technique in molecular biology for the detection of mutations, polymorphisms, and epigenetic differences in double-stranded DNA samples [8]. It is a simple, quick, and accurate method for detecting genetic modifications. In addition, it is more cost effective than other methods such as nucleotide sequencing, real-time PCR with *Taqman* probes, and DNA hybridization techniques with specific probes. However, it requires that the primer pair be well-designed so that homologous DNAs that differ in terms of only one nucleotide can be differentiated by melting curves. For these reasons, we established a molecular protocol based on the real-time PCR HRM technique to detect mutations in *EGFR*'s exon 21. This protocol will be useful for the treatment with gefitinib and erlotinib of NSCLC patients bearing *EGFR* mutations.

Materials and methods

Samples

A total of 172 lung cancer biopsy specimens was collected from Hanoi Oncology Hospital and Cho Ray Hospital. These patients were recommended for *EGFR* mutation testing by physicians due to clinical features associated with *EGFR* mutations. Of the 172 samples, 140 were diagnosed as being adenocarcinoma, and 32 were of other or unknown histology. The tissues were stained using the hematoxylin-eosin (HE) method. The tumour-rich areas were identified by geneticists with the aid of a microscope to ensure that sufficient material was available for the real-time PCR HRM assay. The bacterial strains, including *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus* were provided by Gia Dinh People's Hospital.

Reagents

All the chemical reagents for DNA extraction, PCR, real-time PCR HRM, and agarose gel electrophoresis were purchased from Qiagen, Kapa Biosystems, Merck, and Sigma. The nucleotide sequencing kit was supplied by Applied Biosystems. The primers were synthesized and supplied by Phu Sa Biochem. The PNAClamp™ *EGFR* Mutation Detection kit Ver. 2 was purchased from Panagene, Korea.

DNA extraction from formalin-fixed paraffin-embedded (FFPE) tissue

The biopsy specimen was fixed with formalin and was embedded in paraffin. The specimens were then stained with HE dye and examined under a microscope. The location of the suspected cancer cells was zoned and compared to the cell region on the tissue. A sharp knife was used to cut the tumour-rich areas, including a portion of the surrounding cells. The tissue was washed several times in xylene to dissolve the paraffin wax and then the xylene was removed by washing the tissue several times with ethanol before DNA was extracted. Genomic DNA was extracted using the QIAamp DNA FFPE Tissue kit (Qiagen, Germany) according to the manufacturer's protocol. The extracted DNA was then checked by means of the spectrophotometry method and maintained at 4°C until use.

Design of the HRM primers

In order to ensure the best performance for mutation detection using HRM, primer pairs were designed to create PCR products with sizes ranging from 50 to 120 bp. The primers were designed with Primer3Plus software. Each amplicon was analyzed using Umelt software online to ensure that it contained only a single melting peak. The primers were checked with OligoAnalyzer software and the Blast tool for their *in-silico* performance characteristics.

HRM assays

The real-time PCR HRM reaction was performed in a 48-well Eco Plate on the Eco™ Real-time PCR system (Illumina, USA). The 20 µl final volume of the reaction mixture contained: 1X KAPA HRM Fast Master mix, 3 mM $MgCl_2$, 100 nM of 21F1, 100 nM of 21R1, 10 ng of genomic DNA, and PCR-grade water. The KAPA HRM Fast Master mix contained an optimized concentration of EvaGreen dye and a highly engineered version of *Taq* DNA polymerase. The cycling and melting conditions were as follows: one cycle at 95°C for 3 min; 45 cycles at 95°C for 30s, 66°C for 30s, one cycle of 95°C for 1 min, and a melt from 70°C to 95°C, rising 0.2°C per second. HRM analysis was performed using Eco Software (Ver. 4.1). The normalized graph and the difference graph were used to analyze the data.

DNA sequencing

The amplification reaction of the PCR product for nucleotide sequencing contained the following components: 1X KAPA 2G master mix, 3.5 mM $MgCl_2$, 500 nM of 21F2 and 21R2, 30-50 ng of genomic DNA (if possible), and PCR-grade water. The PCR reaction was performed using the following conditions: initial denaturation at 95°C for 5 min; 40 cycles at 95°C for 20s, 60°C for 30s, 72°C for 30s, and one cycle at 72°C for 7 min.

The PCR product was generated by means of agarose gel electrophoresis. The band of interest was then cut into the 1.5 ml eppendorf to prepare the DNA purification. The PCR products were purified using the Zymo Gel Recovery kit (Zymo Research, USA). Next, the purified product was sequenced using the Big Dye Terminator v3.1 kit (Applied Biosystems, USA), in accordance with the kit manual. Following the Big Dye reaction, the product was cleaned with a sequencing clean kit. The sequencing products were run on a 3100 Genetic Analyzer (Applied Biosystems). The sequencing data was visualized using sequencing analysis v5.3.

PNA Clamp

All samples were set up using the PNAClamp™ *EGFR* Mutation Detection kit (Ver. 2, Panagene, Korea). These samples were set up according to the manufacturer's instructions. The PNA Clamp data was visualized using 7500 software Ver. 2.3.

Results

Identification of the target DNA sequence of exon 21 of EGFR

Exon 21 of *EGFR* encodes the tyrosine kinase region of *EGFR*; its length is 156 bp. The sequence of exon 21 was loaded from GenBank and in Fig. 1, it is shown which mutations (if any) occur in codon L858R and codon L861Q.

ggcatgaactacttggaggaccgtcgcttggtgcaccgcgacctggcagccaggaacgtactggtgaaaacaccgcagcatgtc

aagatcacagattttgggc**t**ggccaaac**t**gctgggtgcggaagagaaagaatac catgcagaaggaggcaaa

Fig. 1. The exon 21 sequence loaded from GenBank. The two nucleotides t, in bold, occur from left to right at codon 858 and codon 861.

Oligonuclotide design

We designed the 21F1-21R1 primer pair for amplification of exon 21 by means of real-time PCR HRM and the 21F2-21R2 primer pair for the nucleotide sequencing of this exon. The nucleotide sequences of the primers were as follows:

21F1: 5' CGC AGC ATG TCA AGA TCA CAG 3' - 21R1: 5' GGT ATT CTT TCT CTT CCG CAC C 3'

21F2: 5' TGG CAT GAA CAT GAC CCT GAA 3' - 21R2: 5' CAG CCT GGT CCC TGG TGT C 3'

The 21F1-21R1 primer pair was designed to generate a PCR product of 66 bp in size. The small size of 66 bp helps to avoid other SNPs in the proximity of the L858R and L861Q mutations. In contrast, the 21F2-21R2 primer pair was designed to generate a PCR product of 295 bp which is suitable for sequencing by means of the Sanger technique.

In the next step, we checked the technical parameters of the primers such as Tm, % GC, and the free energy of the secondary structures using OligoAnalyzer software. The results are presented in Table 1.

Table 1. Technical parameters of the designed primers.

Parameters	Primer			
	21F1	21R1	21F2	21R2
Nucleotide	21	22	21	19
GC content (%)	52.4	50	47.6	68.4
Tm (°C)	56.5	55.6	56.9	60.6
Hairpin (kcal/mole)	0.04	1	-1.55	0.25
Self-dimer (kcal/mole)	-5.38	-4.41	-5.38	-3.55
Hetero-dimer (kcal/mole)	-5.12		-6.21	

The results in Table 1 show that the four primers met the specific requirements for working well in PCR. Finally, we tested the theoretical specificity of these primers using Blast software. The results show that the primers match the human DNA in the *EGFR* gene on chromosome 7 (data not shown).

In addition, using the Umelt software, we predicted the melting curve of 21F1-21R1 primers for wild and mutant forms (of both L858R and L861Q mutants). The melting curve peaks were clearly separated at Tm levels from 0.2°C to 0.7°C (data not shown). In conclusion, the designed primers were suitable for the subsequent experiments.

Building the real-time PCR HRM for exon 21 EGFR mutation detection

We set up a real-time PCR HRM reaction with the 21F1-21R1 primer pair on a few human DNA samples. All of them were positive in the real-time PCR. When analyzed by means of HRM, they showed different melting curves corresponding to the L858R and L861Q muations, as shown in Fig. 2.

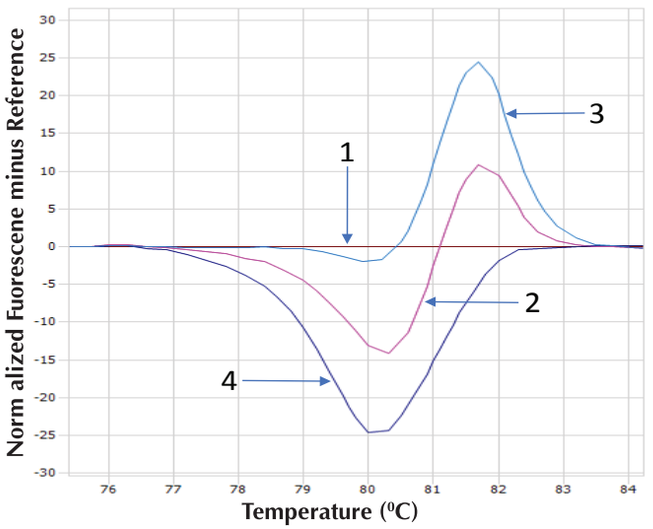


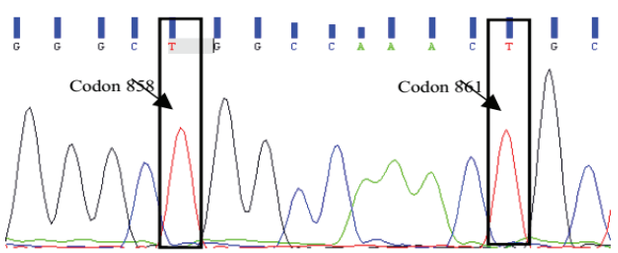
Fig. 2. The real-time PCR HRM results of exon 21 mutation detection. 1: the wild type, 2: the L858R heterozygous mutation, 3: the L858R homozygous mutation, 4: the L861Q heterozygous mutation.

According to the prediction of the Umelt software, there were four different melting curves which corresponded to the expected curves of the wild type (curve 1), the L858R heterozygous mutation (curve 2), the L858R homozygous mutation (curve 3), and the L861Q heterozygous mutation (curve 4). The expected curve of the L861 homozygous mutation was not found in HRM results.

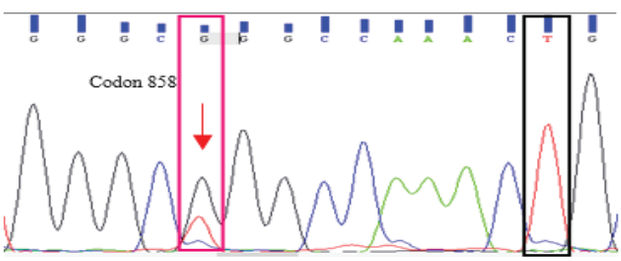
To verify the results of the real-time PCR HRM, we ran Sanger nucleotide sequencing on these four samples. The results of the nucleotide sequencing of the four samples are shown in Fig. 3.

The nucleotide sequencing results show that there were two peaks of thymine at codon 858 and codon 861 in the wild-type sample. There were two peaks of guanine and thymine at codon 858 in the L858R heterozygous mutation sample. The L858R homozygous mutation sample had one peak of guanine at codon 858. The L861Q heterozygous mutation sample had one peak of adenine and thymine at codon 861. Thus, the Sanger nucleotide sequencing confirmed that the mutation detection results using real-time PCR HRM were correct.

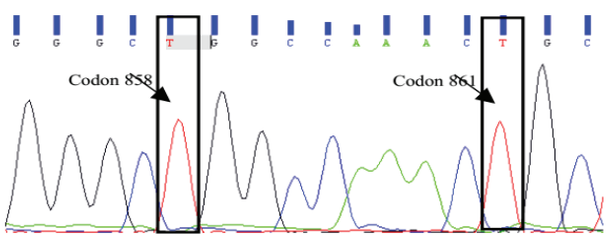
With these four samples, we performed the optimization of $MgCl_2$ concentration for the real-time PCR HRM.



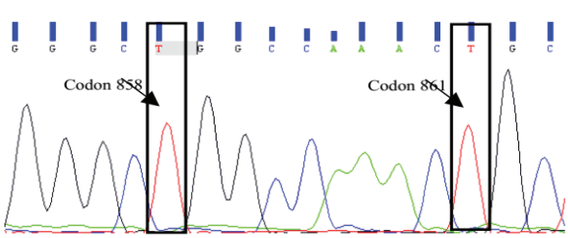
The wild-type sample



The L858R heterozygous mutation sample



The L858R homozygous mutation sample



The L861Q heterozygous mutation sample

Fig. 3. Results of the Sanger nucleotide sequencing of four samples of wild type, L858R heterozygous mutation, L858R homozygous mutation, and L861Q heterozygous mutation.

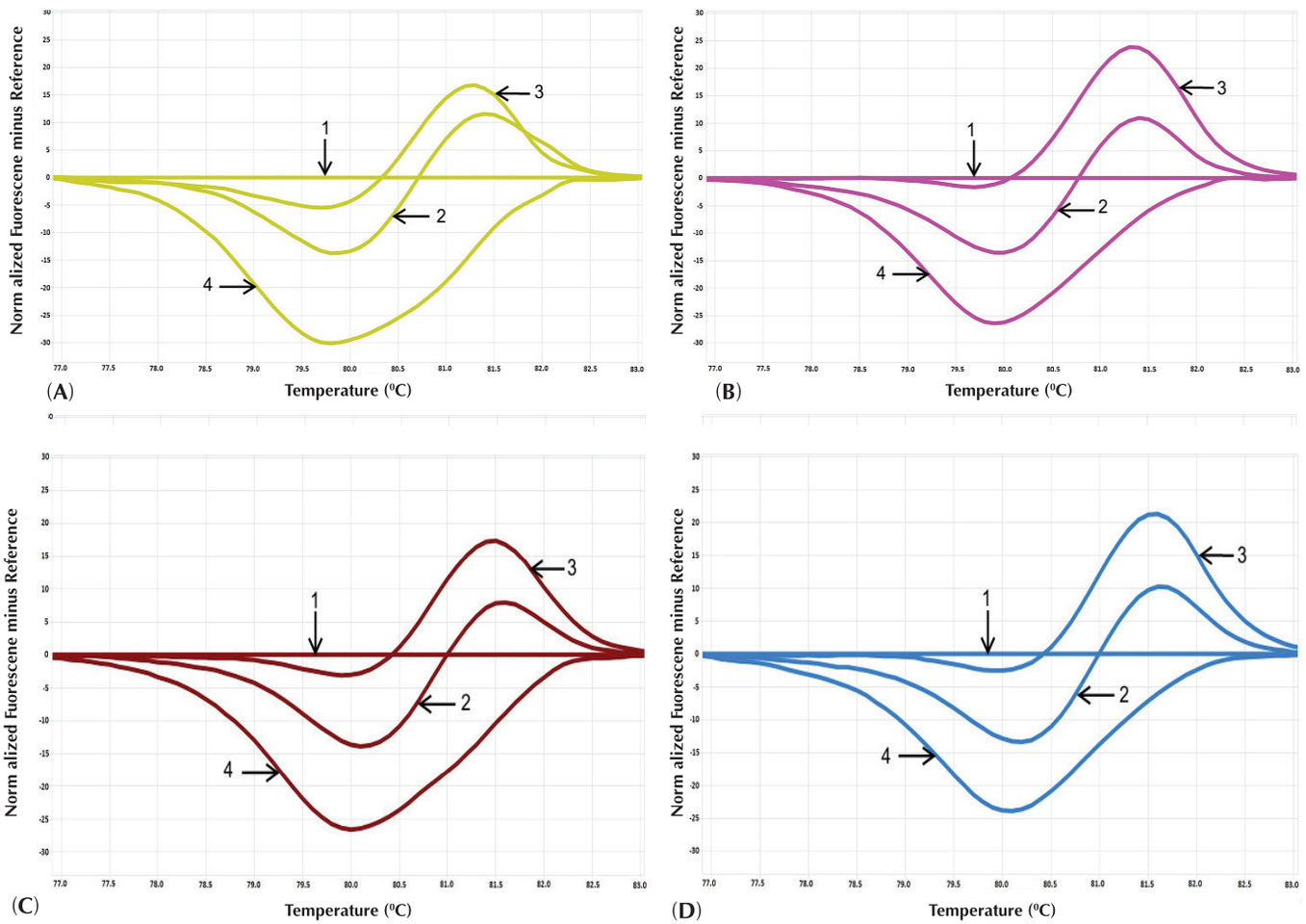


Fig. 4. Examination of the optimal $MgCl_2$ concentration for the real-time PCR HRM. (**A**) 1.5 mM, (**B**) 2 mM, (**C**) 2.5 mM, and (**D**) 3 mM.

Optimization of $MgCl_2$ concentration

$MgCl_2$ is the major component that influences the melting temperature of PCR products when analyzed using HRM. Therefore, we examined the optimum $MgCl_2$ concentration for distinguishing different melting curve types corresponding to the mutations at codons 858 and 861 of exon 21. We investigated the $MgCl_2$ concentrations of 1.5 mM, 2 mM, 2.5 mM, 3 mM, and 3.5 mM. The results of the optimization of $MgCl_2$ concentration are shown in Fig. 4.

The results in Fig. 4 show that the four melting curve genes corresponding to the wild type, the L858R heterozygous mutation, the L858R homozygous mutation, and the L861Q heterozygous mutation were more clearly distinguished when the $MgCl_2$ concentrations were increased. At the $MgCl_2$ concentration of 3.5 mM, a negative control has an amplified signal (data not shown). Thus, we selected the $MgCl_2$ concentration of 3 mM for the real-time PCR HRM.

Analytical sepcificty of the real-time PCR HRM protocol

Analytical specificity of the real-time PCR HRM protocol was demonstrated by means of the selective amplification of the human DNA region containing exon 21 of *EGFR* by means of the 21F1-21R1 primer pair. To effect this, we first performed restriction enzyme analysis of the PCR product of the 21F1-21R1 primer pair because this PCR product is too small to be analyzed using Sanger nucleotide sequencing. The restriction enzymes used in this experiment were the *Mbo*I enzyme, which cuts the target product into two fragments of 14 bp and 52 bp, and the *Dpn*I, which cuts the target product into two fragments of 16 bp and 50 bp. The results of the restriction anlysis are shown in Fig. 5.

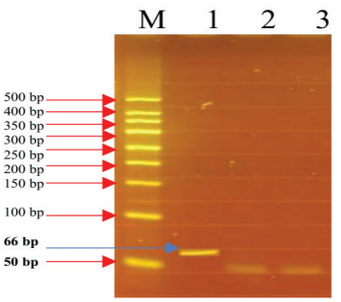


Fig. 5. Restriction analysis of the PCR product of 21F1-21R1. M: DNA ladder, 1: the untreated PCR product, 2: the PCR product treated with *Mbo*I, 3: the PCR product treated with *Dpn*I.

The results in Fig. 5 show that the restriction products from the treatment of the 21F1-21R1 PCR product with *Mbo*I and *Dpn*I had the expected sizes when compared to the DNA ladder. DNA bands of approximate 50 bp appeared in lane 2 and lane 3, while the bands of 12 bp and 14 bp were too small and they were run out of the gel. This result confirmed that the 21F1-21R1 amplified the target region in exon 21 of *EGFR* containing the L858R and L861Q mutations.

We also investigated the selective amplification of the 21F1-21R1 primer pair on the genetic material of various agents, including human and human pathogenic bacteria (*Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus*) in the real-time PCR. We selected these bacteria for the selective amplification experiment because they can coexist in humans as a result of infection and the total DNA extracted from human samples can contain bacterial DNAs. The results of the selective amplification of the 21F1-21R1 primer pair are presented in Fig. 6.

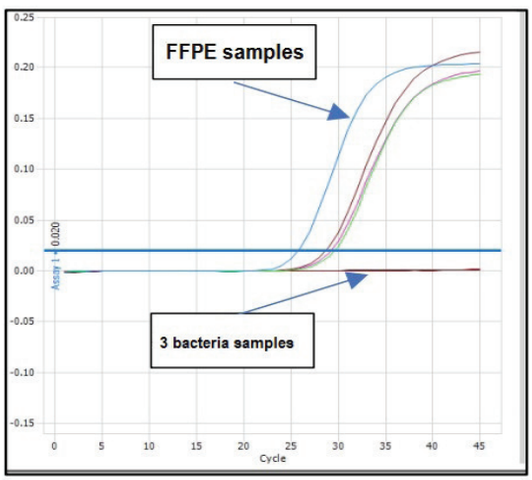


Fig. 6. Selective amplification of the 21F1-21R1 primer pair on various genetic materials from humans and bacteria.

The results in Fig. 6 show that only the human DNA samples generated a positive result in the real-time PCR reaction with the 21F1-21R1 primer pair; DNA samples from the bacteria produced negative results in the reaction with the same primer pair. To confirm that the negative results in the real-time PCR with the 21F1-21R1 primer pair on the bacterial DNA samples were truly negative, we performed a PCR with the 8F-534R primer pair on these DNA samples. This is a primer pair specific to the 16S rRNA gene of all eubacteria. The results showed that the DNA samples from *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus* were positive for the PCR with the 8F-534R primer pair (data not shown). These results confirmed that the negative results of the real-time PCR with the 21F1-21R1 primer pair on the bacterial DNA samples were truly negative. Thus, the real-time PCR HRM protocol with the 21F1-21R1 primer pair was specifically designed for the detection of the L858R and L861Q mutations in exon 21 of *EGFR*.

Limit of detection of the HRM method

Most of the FFPE tissue samples of NSCLC cases obtained for this study were small in size. Therefore, the amount of DNA extracted from these samples were too small for analysis. To check the limit of detection of the real-time PCR HRM protocol to detect the L858R and L861Q mutations, we performed the real-time PCR HRM on a range of amounts of DNA, from 1 to 15 ng, with four samples of the wild-type sample, the L858R homozygous mutation sample, the L858R heterozygous mutation sample, and the L861Q heterozygous mutation sample. The results of the limit of detection of the HRM method are shown in Table 2.

Table 2. Results of the limit of detection of the HRM method (in four replicates).

Sample	1 ng	5 ng	7 ng	10 ng	15 ng
Detections (normal + mutation)/total run					
The wild-type sample	4/4	4/4	4/4	4/4	4/4
The L858R homozygous mutation sample	4/4	4/4	4/4	4/4	4/4
The L858R heterozygous mutation sample	4/4	4/4	4/4	4/4	4/4
The L861Q heterozygous mutation sample	4/4	4/4	4/4	4/4	4/4

The results in Table 2 show that the molecular protocol could detect DNA concentrations at 1 ng. At 1 ng DNA levels, all the samples were amplified and the melting curve analysis showed the distinction between the wild type and mutant.

Repeatability

We detected the mutations in exon 21 of *EGFR* using the real-time PCR HRM protocol on four samples of the wild-type sample, the L858R homozygous mutation sample, the L858R heterozygous mutation sample, and the L861Q heterozygous mutation sample five times in the same experiment batch to measure the repeatability of the protocol. The repeatability results were expressed by the coefficient of variation (CV) and the CV values were calculated as a percentage. The results are shown in Table 3.

Table 3. Repeatability calculation of the real-time PCR HRM for exon 21 mutation detection.

Repeat	Wild-type sample	L858R homozygous mutation sample		L858R heterozygous mutation sample		L861Q heterozygous mutation sample
1	81	79.8	81.6	79.5	81.2	81.6
2	81.1	79.7	81.6	79.4	81.2	81.7
3	81.2	79.9	81.7	79.5	81.2	81.6
4	81.1	79.8	81.7	79.4	81	81.7
5	81.2	79.8	81.7	79.4	81.2	81.8
Average	81.13	79.80	81.66	79.44	81.16	81.68
Standard deviation	0.08367	0.07071	0.05477	0.05477	0.0894	0.08367
% CV	0.00103	0.00089	0.00067	0.00069	0.0011	0.00102
Total % CV	0.00090					

According to Table 3, the CV value of the repeatability test was 0.00090%. This CV proves the high precision of the real-time PCR HRM protocol for the detection of mutations in exon 21 of *EGFR*.

Evaluating the real-time PCR HRM protocol on 172 human DNA samples

We evaluated the performance of the real-time PCR HRM protocol for detection of the mutations in exon 21 of *EGFR* on 172 human DNA samples which were provided by Hanoi Oncology Hospital and Cho Ray Hospital. The detection of these mutations was also performed using the Sanger nucleotide sequencing and the PNAClamp™ EGFR Mutation Detection kit. The results of mutation detection in exon 21 of *EGFR* are shown in Table 4.

Table 4. Results of *EGFR* mutation detection by real-time PCR HRM, nucleotide sequencing, and the PNA clamping method.

Results/methods	HRM	SEQ	PNA Clamp
L858R mutation	45	40	44
L861Q mutation	2	2	2
No mutation	125	126	124
Sample invalid/suspected	0	4	2
Total sample with results/total sample	172/172	168/172	170/172

HRM: high resolution melting; SEQ: sequencing; PNA Clamp: peptide nucleic acid clamping.

The number of samples successfully analyzed by means of the three methods were 172, 168, and 170 for the real-time PCR HRM protocol, the Sanger nucleotide sequencing, and the PNAClamp™ *EGFR* Mutation Detection kit, respectively. In these 172 samples, 166 (96.5%) showed identical results for all three methods. The real-time PCR HRM protocol detected 45 samples with the L858R mutation while only 40 and 44 samples with this mutation were detected by the Sanger nucleotide sequencing method and by the PNAClamp™ *EGFR* Mutation Detection kit, respectively. Six cases that did not match among three methods were samples 19, 72, 24, 34, 171, and 172. The results of these samples using the three methods are shown in Table 5.

Table 5. The results of the samples using the three methods.

Sample	HRM	GTT	PNA Clamp
19	Mutation L858R heterozygous	Wild type	Mutation L858R
72	Mutation L858R heterozygous	Wild type	Mutation L858R
24	Mutation L858R heterozygous	Suspected wild type	Mutation L858R
34	Mutation L858R heterozygous	Suspected wild type	Mutation L858R
171	Wild type	Invalid	Invalid
172	Mutation L858R heterozygous	Invalid	Invalid

As can be seen in Table 5, the real-time PCR HRM protocol and the PNAClamp™ *EGFR* Mutation Detection kit produced identical results for four samples (19, 72, 24, and 34) while the Sanger nucleotide sequencing produced different results for the same samples. Moreover, the real-time PCR HRM protocol was successful for the last two samples (171 and 172) while the PNAClamp™ *EGFR* Mutation Detection kit and the Sanger nucleotide sequencing produced invalid results. Further experiments need to be conducted to confirm the results of these six samples.

Discussion

The appearance of the targeted therapy made the cancer treatment procedure more efficient and less toxic than traditional radiotherapy/chemotherapy. Tyrosine kinase inhibitors of *EGFR* are good examples of targeted therapy for NSCLC patients with *EGFR* mutations. Among *EGFR* mutations, L858R and L861Q in exon 21 were proven to confer selective sensitivity to treatment with some TKIs such as erlotinib and gefitinib [6]. Thus, a molecular protocol to detect these beneficial mutations would be useful for the treatment of NSCLC patients with *EGFR* mutations.

With the L858R mutation, the leucine amino acid (L) is replaced by the amino acid arginine (R), meaning that the nucleotide T is replaced by the nucleotide G. Therefore, with L858R there will theoretically be three forms: T/T (wild type), T/G (heterozygous), and G/G (homozygous). With the L861Q mutation, the leucine amino acid (L) is replaced by glutamine amino acid (Q), meaning that the nucleotide T is replaced by the nucleotide A. Hence, there are three forms: T/T (wild type), T/A (heterozygous), and A/A (homozygous). It is interesting to note that we did not find any references which mentioned that the two mutations (L858R and L861Q) are concomitant, which means that a molecular protocol based on the real-time PCR HRM technique can be applied to detect the L858R and L861Q mutations. In fact, we did not find any sample out of the 172 from NSCLC patients which had the concurrent L858R and L861Q mutations in this study. As mentioned above, as real-time PCR HRM is a powerful technique for the detection of mutations with simple manual, we selected this technique to establish a molecular protocol for detection of the L858R and L861Q mutations in *EGFR*.

The PCR products for HRM analysis often have a size of 100-300 bp; however, in this study, we designed the primer pair to amplify the DNA region of 66 bp. The small size of the PCR product enhances the differentiation of homologous DNA sequences that differ only at the L858R and L861Q positions. Moreover, the size of 66 bp helps to avoid SNPs in the proximity of these two mutations. In fact, there are many SNPs in the region surrounding L858R and L861Q, some of which can be listed here: rs1433831615, rs104886012, rs148934350, rs397517134, rs764700695, and rs1471553524. These SNPs may interfere with the melting curves generated by the primer pair in this study. When theoretically analyzed by means of HRM, the 66 bp PCR product generates five melting curve patterns corresponding to the wild type, the L858R heterozygous mutation, the L858R homozygous mutation, the L861Q heterozygous mutation, and the L861Q homozygous mutation. In fact, we obtained four melting curves corresponding to the wild type, the L858R heterozygous mutation, the L858R homozygous mutation, and the L861Q heterozygous mutation using the real-time PCR HRM protocol on

20 human DNA samples. These mutations detected by the real-time PCR HRM protocol were confirmed by the Sanger nucleotide sequencing method.

The 21F1-21R1 primer pair used for the real-time PCR HRM protocol had characteristics suitable to working properly in the real-time PCR HRM reaction [9]. The size of the primers was 21-22 bp, with % GC in the range from 40-60%. There are no more than three guanine or cytosine at the 3' end. The Tm difference between the two primers was not more than 5°C. The free energy of the secondary structures (hairpin, homodimer, and heterodimer) of the primers did not exceed the experimental parameter of -9 kcal/mol which may interfere with the binding to the target DNA strands during PCR. The Blast results showed that the two primers annealed to the human DNA sequences. In the experiment, the 21F1-21R1 primer pair amplified the target region of 66 bp, which was then confirmed by restriction analysis with *Mbo*I and *Dpn*I. These two REs were capable of cutting the target 66 bp sequence at their recognition sites to produce specific DNA fragments. *Mbo*I cut the target product into two fragments of 14 bp and 52 bp at 5'-^ GATC-3'. *Dpn*I cut the target product into two fragments of 16 bp and 50 bp at 5'-GA^ TC-3'. In addition, the 21F1-21R1 primer pair demonstrated its ability to selectively amplify human DNA among several other genetic materials. These results demonstrate the high analytical specificity of the real-time PCR HRM protocol. Other performance characteristics of the real-time PCR HRM protocol concerning limit of detection and repeatability were good, with a limit of detection of 1 ng of total DNA, and CV of 0.0009%.

Finally, we evaluated the molecular protocol for the detection of L858R and L861Q in 172 samples of human DNA collected from hospitals where NSCLC patients were being treated. We also detected the L858R and L861Q mutations in these 172 samples using two other methods: Sanger nucleotide sequencing and the method using the PNAClamp™ *EGFR* Mutation Detection kit. All three methods showed identical results for 166 of the 172 (96.5%) samples. Discrepancies occurred in the results of the three methods with 6 of the 172 (3.5%) samples. The real-time PCR HRM protocol successfully detected mutations in all of the 172 samples, while the number of samples that were invalid or were suspected to be invalid was four for Sanger nucleotide sequencing and two for the PNAClamp™ kit. We could not reperform the detection of the mutations using Sanger nucleotide sequencing and the PNAClamp™ kit on the discrepant samples due to the limits of the study; however, further experiments could be undertaken to show the utility of the real-time PCR HRM protocol in comparison with other methods of detecting L858R and L861Q mutations.

Conclusions

In this study, we established the molecular protocol based on the real-time PCR HRM technique for the detection of the L858R and L861Q mutations in exon 21 of *EGFR*. The performance characteristics of the genotyping protocol, in terms of analytical specificity, limit of detection, and repeatability, were good. Application of this mutation detection protocol to 172 human DNA samples showed good results in comparison with Sanger nucleotide sequencing and the PNA clamping method. These results demonstrate the utility of the real-time PCR HRM protocol for the detection of exon 21 mutations in clinical samples, which is essential for the treatment of NSCLC patients with *EGFR* mutations.

ACKNOWLEDGEMENTS

This study was conducted by Gentis Joint Stock Company. We thank Hanoi Oncology Hospital and Cho Ray Hospital for providing the FFPE samples and Gia Dinh People's Hospital for providing the bacterial strains.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

[1] D.T. William, B. Elisabeth, G.N. Andrew, Y. Yasushi, H.M.A. John, B. Mary, R.C. Lucian, D. Sanja, D. Edwina, B.F. Douglas, G. Kim, R.H. Fred, YuichiIshikawa, M.K. Keith, N. Masayuki, P. Giuseppe, A.P. Charles, T. Ming Sound, W. Ignacio (2015), "The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification", *Journal of Thoracic Oncology*, **10(9)**, pp.1243-1260.

[2] S.H. Roy (2004), "Review of epidermal growth factor receptor biology", *Int. J. Radiat. Oncol. Biol. Phys.*, **59(2)**, pp.21-26.

[3] S.H. Yeh, C.F. Lin, F.L. Kong, H.E. Wang, Y.J. Hsieh, J.G. Gelovani, R.S. Liu (2013), "Molecular imaging of nonsmall cell lung carcinomas expressing active mutant *EGFR* kinase using PET with [(124)i]-morpholino-IPQA", *BioMed Research International*, doi: 10.1155/2013/549359.

[4] S. Ivana, P. David (2017), "Next-generation EGFR tyrosine kinase inhibitors for treating EGFR-mutant lung cancer beyond first line", *Frontiers Medicine*, **3**, p.76, doi.org/10.3389/fmed.2016.00076.

[5] EGFR gene, https://ghr.nlm.nih.gov/gene/EGFR#location.

[6] J. Köhler, M. Schuler (2013), "Afatinib, erlotinib and gefitinib in the first-line therapy of EGFR mutation-positive lung adenocarcinoma: a review", *Onkologie.*, **36(9)**, pp.510-518, doi. org/10.1159/000354627.

[7] S. Yuankai, S.K.A. Joseph, T. Sumitra, S. Sankar, T. Chun-Ming, T.K. Mai, H. Karin, Y.I. Med, C. Gerardo, Y. Pan-Chyr (2014), "A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER)", *Journal of Thoracic Oncology*, **9(2)**, pp.154-162.

[8] G.H. Reed, J.O. Kent, C.T. Witlwer (2007), "High-resolution DNA melting analysis for simple and efficient molecular diagnostics", *Pharmacogenomics*, **8(6)**, pp.597-608.

[9] E. Prediger, *Designing primers and probes*, Integrated DNA Technologies Guideline, https://eu.idtdna.com/pages/education/decoded/article/designing-pcr-primers-and-probes.