Constructing a molecular genotyping assay for rs11077 based on real-time polymerase chain reaction high resolution melting (PCR HRM) technique for the prognosis of hepatocellular carcinoma (HCC) patients

Thi Hao Pham[1], Thanh Huy Nguyen[1], Minh Phung Truong[1], Thi Thuy Hang Le[2], Quoc Dang Quan[3], Quang Tri Le[4], Hoang Chuong Nguyen[5*]

[1]Center for Research and Application in Bioscience
[2]Ho Chi Minh city University of Food Industry
[3]Center of Science and Technology Development, Ho Chi Minh Communist Youth Union
[4]7A Military Hospital
[5]University of Science, Vietnam National University Ho Chi Minh city

Received 5 March 2018; accepted 19 May 2018

Abstract:

***XPO5* codes for the nuclear transport factor exportin-5, which is a membrane-bound protein. This gene is responsible for the transport of pre-miRNA from the nucleus to the cytoplasmic compartments, thereby adjusting the whole miRNA expression level. The reduction of the miRNA levels was recorded when *XPO5* was knocked down. rs11077 is found in the 3′UTR region of *XPO5*, and this SNP might affect mRNA stability and be associated with the altered expression of *XPO5*. This leads to the universal suppression of miRNA expression profiles, thereby mediating the HCC survival. HCC patients bearing C/C and A/C genotypes of rs11077 had a survival rate of 60% after 3 years; and this rate was reduced to 24.7% with HCC patients bearing the A/A genotype. In this study, we constructed a molecular assay based on a real-time PCR HRM technique for rs11077 genotyping. We successfully designed the primer pair for the real-time PCR HRM of rs11077. We also found the optimal concentration of $MgCl_2$ to arrive at a clear differentiation of the three genotypes of rs11077. Thereafter, we characterised the analytical specificity and the precisions of the molecular assay. The SNP genotyping results were compared between the real-time PCR HRM and nucleotide sequencing. Finally, we evaluated the molecular assay on 123 human blood samples. The rs11077 genotyping assay in this study could be used for the prognosis of HCC patients.**

Keywords: ***hepatocellular carcinoma, real-time PCR HRM, rs11077, SNP.***

Classification numbers: ***3.2, 3.5***

Introduction

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer, and is also the 5th most common type of cancer in the world. Worldwide, more than 700,000 cases are diagnosed each year; and the high mortality rates make HCC the third leading cause of cancer deaths in the world, following lung cancer and stomach cancer. The incidence of HCC varies across the geographical regions of the world, as it relates to the frequency difference of risk factors, the most pertinent ones being chronic hepatitis B and C infection [1, 2]. Vietnam is among the countries with the highest rates of HCC in the world [3].

rs11077 in the 3'UTR region of *XPO5* has been shown to be related to HCC. This SNP consists of two alleles - C and A - that produce three different genotypes, namely A/A, C/C, and A/C. It was found that the 3-year survival rate for HCC patients with genotype A/C and C/C was 60%, while in HCC patients with genotype A/A, it was 24.7%. Therefore, making the distinction between the three genotypes of rs11077 in the *XP05* gene is necessary to be predictive for the treatment of HCC [4, 5].

There are many molecular methods for SNP genotyping such as real-time PCR, PCR-RFLP, ARMS-PCR, PCR-sequencing, and real-time PCR HRM. Among them, real-time PCR HRM has many advantages over other methods such as ease of operation and shorter process duration (compared to PCR-RFLP, PCR-sequencing), which does not require a complicated primer design and PCR optimisation (compared to ARMS-PCR). It also does not require

*Corresponding author: Email: hoangchuongnguyen@gmail.com

complicated analytical devices (automatic nucleotide sequencing machine). Finally, it also provides accurate genotyping result without requiring expensive chemical reagents (Taqman probe). For these reasons, we performed this research aiming at developing a molecular assay for rs11077 genotyping that could be used to genotype rs11077 in clinical practice.

Materials and methods

Reagents

The human blood samples and bacterial strains were supplied by Center for Research and Application in Bioscience (Ho Chi Minh city, Vietnam). The blood samples were collected from healthy people. All the chemical reagents for the DNA extraction, real-time PCR HRM, and agarose gel electrophoresis were purchased from Bioline, Merck and Sigma. The nucleotide sequencing kit was supplied by Applied Biosystems. The primers were synthesised and supplied by Phu Sa Biochem.

Primer design

A DNA fragment containing rs11077 was obtained from GenBank to be used as the template for the primer design. Two primer pairs were designed using the AnnHyb software, in which one pair was designated for genotyping rs1801133 using the real-time PCR HRM method and the other was designated for the nucleotide sequencing of this SNP. The oligo characteristics of these primers in terms of Tm, percentage of GC, free energy of the secondary structures (hairpin, homodimer, heterodimer) were examined using the OligoAnalyzer software to ensure good performance during PCR. Finally, selective binding to the target region containing rs11077 of these primers was checked using the Primer-Blast software.

DNA extraction

Whole blood was first treated with a red blood cell lysis buffer to remove the erythrocytes and obtain the lymphocytes. The lymphocytes were lysed with the guanidine thiocyanate-containing solution in the presence of silica particles. Proteins and other impurities were removed, and the genomic DNA was absorbed by the silica particles. The DNA-containing silica were washed with washing buffer and ethanol 70% to completely remove the remaining impurities and salts. Finally, the genomic DNA was eluted from the silica particles in nuclease-free water and kept at -20^{o}C until used.

Real-time PCR HRM

The 20 µl-volume real-time PCR reaction was set up in 0.1 ml tube with the following components: 10 µl of SensiFast^{TM} HRM master mix 2X (Bioline), 2 µl of the 10 µM-concentration CN5-CN6 primer pair, 5 µl of the genomic DNA template, and 3 µl of water. The reaction program was initiated at 95^{o}C for 120 seconds followed by 40 cycles at 95^{o}C for 10 seconds, 60^{o}C for 10 seconds, and 72^{o}C for 10 seconds. The HRM analysis on PCR product was started at 60^{o}C to 97^{o}C with 0.1^{o}C increment. The results were analysed using MyGo-Pro PCR software based on the melting curve shape on the normalised melting curves and melting point (Tm) of the amplified products.

Nucleotide sequencing

The PCR products containing rs11077 were obtained using PCR with the CN15-CN16 primer pair. These PCR products were purified before being labelled with the appropriate fluorescents. The fluorescent-labelled PCR products were analysed on the ABI 3500 genetic analyser. The nucleotide sequence of the target PCR product was analysed based on the fluorescence signals and was then compared to the original sequence containing rs11077 on GenBank nucleotide database.

Analytical specificity

The selective amplification of the CN13-CN14 primer pair was checked using real-time PCR with the genetic materials from bacteria such as *Escherichia coli, Staphylococcus aureeus, Pseudomonas aeruginosa, Shigella dysenteria, Vibrio cholera,* and *Klebsiella pneumoniae* that may co-exist in human body. In addition, the PCR with the 27F-1495R was performed on these genetic materials to prove that the negative results in the above real-time PCR were not a result of the PCR inhibition.

Precisions

The real-time PCR HRM for genotyping rs11077 was repeated five times in the same test conditions on the same day on the samples containing known C/C, C/A, and A/A genotypes to check the repeatability of the method. Similarly, the rs11077 genotyping protocol was repeated five times in various test conditions on the samples containing known C/C, C/A, and A/A genotypes to check reproducibility. The degree of deviation in the rs11077 genotyping result on the samples was assessed using the value of the coefficient of variation (CV) and the unit of calculation was expressed in

percentage.

Data analysis

All the data in this study was analysed with Excel (Microsoft Office 2013).

Results

Primer design

We designed the primer pair for genotyping rs11077 using a real-time HRM PCR assay and the primer pair for nucleotide sequencing of this SNP. The nucleotide sequences of the primers were shown as follows:

CN13: AGTACCTCCAAGGACCAGG

CN14: AAAGGGGATGTTAGCACTAAAGAC

CN15: CCTTTTGCTGCTGGGCTGG

CN16: TGAGTGGACCTTGAGGCTG

The CN13-CN14 primer pair was designed for genotyping rs11077, and the PCR product with this primer pair was 51 bp in size. In contrast, the CN15-CN16 primer pair was designed for sequencing rs11077 using the Sanger technique, and the PCR product with this primer pair was 300 bp in size.

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

Table 1. Technical parameters of the designed primers.

Parameters	Primer			
	CN13	*CN14*	*CN15*	*CN16*
Nucleotide	19	24	19	19
GC content (%)	57.9	41.7	63.2	57.9
Tm (°C)	55.8	54.9	59.8	56.9
Hairpin (kcal/mole)	-1.58	0.49	-0.6	-0.65
Self-dimer (kcal/mole)	-4.67	-4.9	-3.14	-4.67
Hetero-dimer (kcal/mole)	-4.5		-4.67	

The results in Table 1 showed that the four primers met the specific requirements that allowed them to work well in the PCR. Finally, we tested the theoretical specificity of these primers using the Blast software. The results showed that the primers matched only the human DNA in the target gene *XPO5* (data not shown). In conclusion, the primers that we designed were suitable for the subsequent experiments.

Building the real-time PCR HRM for rs11077 genotyping

With the CN13-CN14 primer pair, we set up a real-time PCR HRM reaction on five human DNA samples. The results are illustrated in Fig. 1.

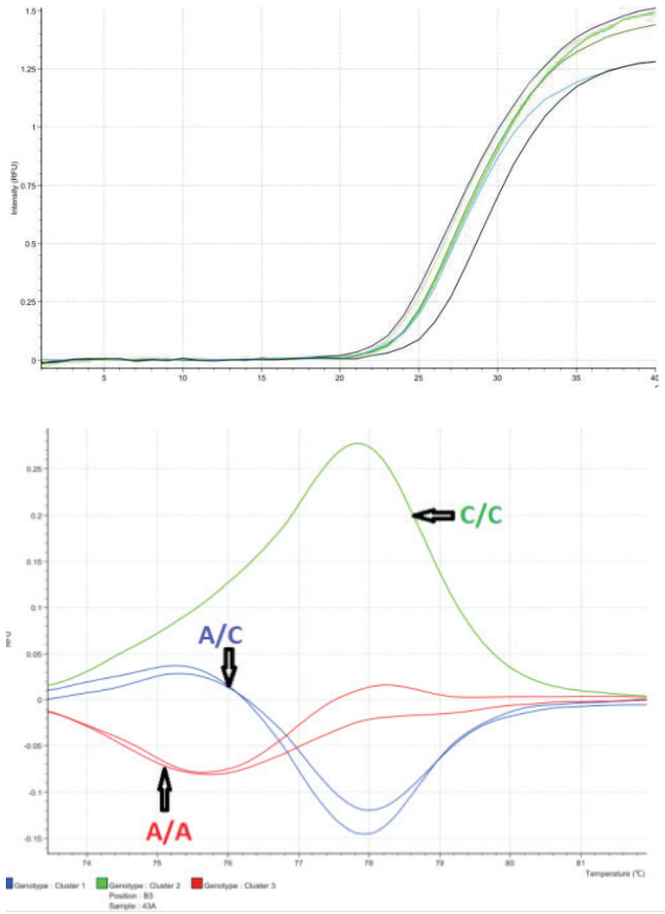

Fig. 1. The real-time PCR HRM results on five human DNA samples.

The results in Fig. 1 show that the real-time PCR results were positive for five DNA samples. While analysing the melting curve using HRM software, three different melting curve patterns that corresponded to the three genotypes A/A, A/C, and C/C of rs11077 were observed. The predicted A/A, A/C, and C/C genotypes based on the 3 melting curve models were confirmed by Sanger's nucleotide sequencing.

For rs11077 nucleotide sequencing, we used the CN15-CN16 primer pair, as the PCR product from the CN13-

CN14 primer pair was not suitable for Sanger nucleotide sequencing because of its small size (51 bp). The results of the nucleotide sequences are shown in Fig. 2.

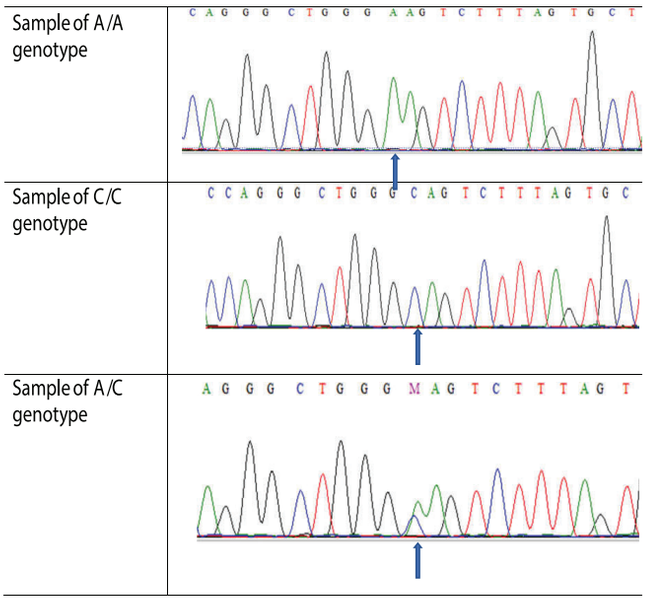

Fig. 2. Results of Sanger's nucleotide sequencing of the A/A, A/C, and C/C genotypes.

The results in Fig. 2 show the occurrence of the peaks corresponding to the nucleotides of rs11077. Specifically, the sample containing the genotype A/A contained a peak of Adenine, and the sample containing the genotype C/C contained a peak representing Cytosine. Samples containing genotype A/C contain the presence of a double peak, representing Adenine and Cytosine. Thus, the result of Sanger's nucleotide sequencing confirmed that the rs11077 genotyping results obtained by using real-time PCR HRM were completely accurate.

Optimisation of $MgCl_2$ concentration

$MgCl_2$ is the major component that influences the melting temperature of PCR products when analysed by HRM. Therefore, we investigated the optimal $MgCl_2$ concentration for distinguishing between the three melting curve patterns corresponding to three genotypes A/A, A/C, and C/C of rs11077. In the PCR master mix for real-time PCR HRM with unknown concentration of $MgCl_2$, we investigated added $MgCl_2$ concentrations in the real-time PCR HRM as follows: 0 mM, 0.5 mM, 1 mM, 1.5 mM, 2 mM, and 2.5 mM. The results of the optimisation of $MgCl_2$ concentration are shown in Fig. 3.

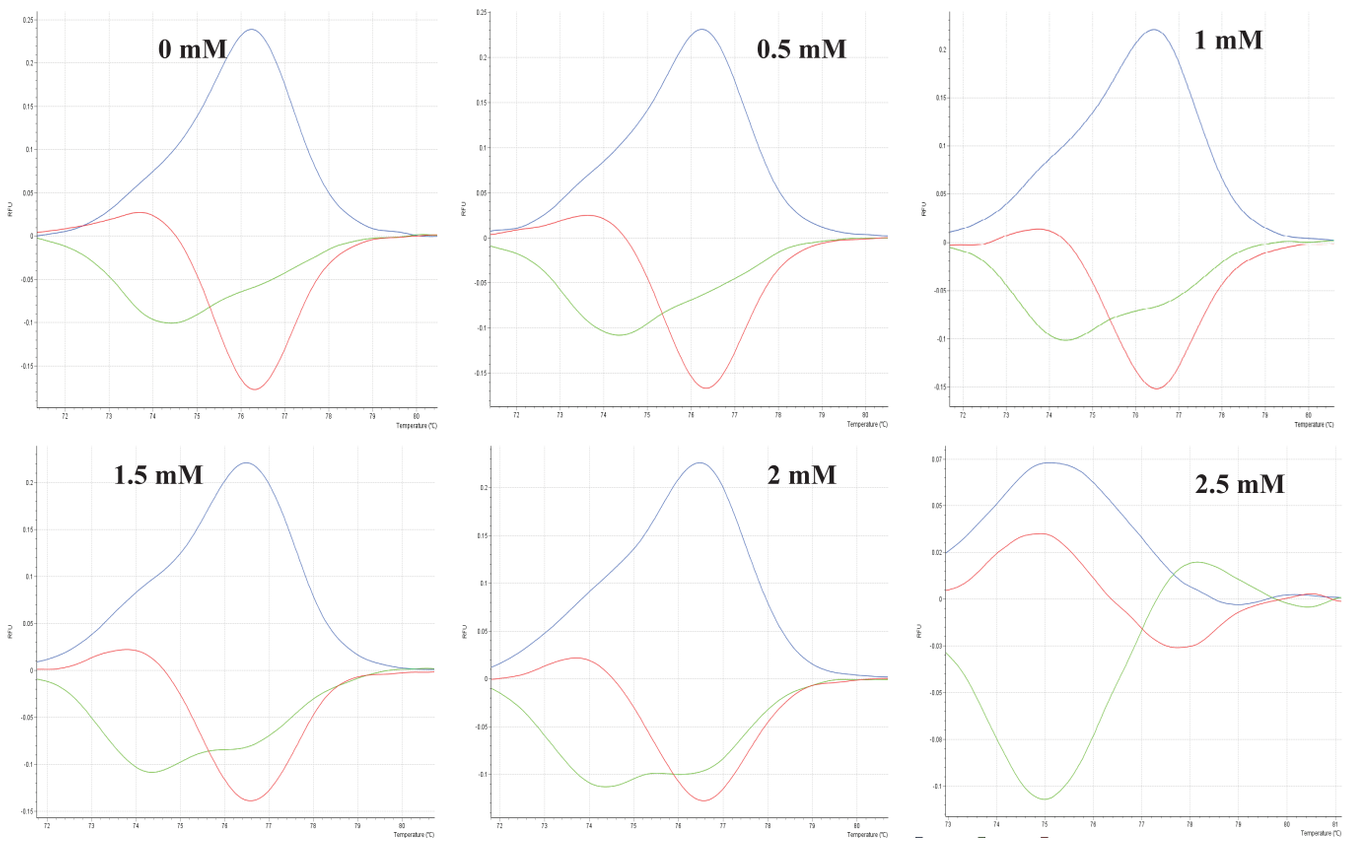

Fig. 3. Optimisation of $MgCl_2$ concentrations for the real-time PCR HRM for genotyping rs11077.

The results in Fig. 3 show that the three melting curves that correspond to the three genotypes A/A, A/C, and C/C of rs11077 were most clearly distinguished at the added $MgCl_2$ concentrations of 0 mM, 0.5 mM, and 1 mM. The other $MgCl_2$ concentrations presented unspecific curves. With this result, we chose 0 mM as the added concentration of $MgCl_2$ for subsequent studies.

Analytical specificity of the real-time PCR HRM protocol

Analytical specificity of the real-time PCR HRM assay was demonstrated by the selective amplification of the human DNA region containing rs11077 by the target primer pair. In this experiment, we investigated the selective amplification of the CN13-CN14 primer pair on the genetic material of various agents, including human and bacteria (*Escherichia coli, Staphylococcus aureus, Pseudomonas aeruginosa, Shigella dysenteriae, Vibrio cholerae,* and *Klebsiella pneumoniae*) in the real-time PCR. The results of the selective amplification of the CN13-CN14 primer pair are presented in Fig. 4.

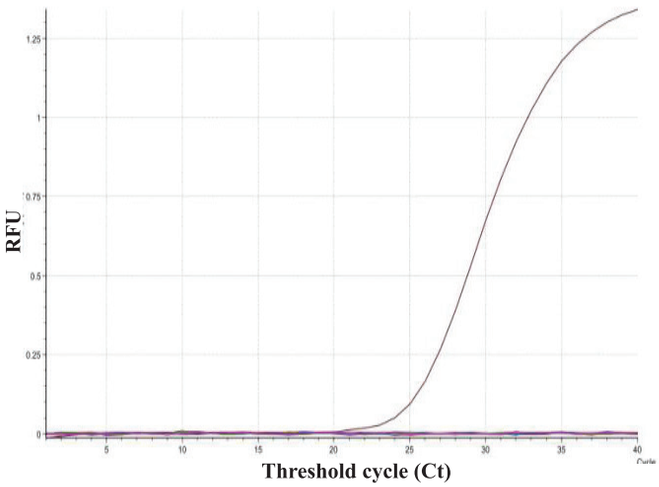

Fig. 4. Selective amplification of the CN13-CN14 primer pair on various genetic materials from human and bacteria.

The results in Fig. 4 show that only the human DNA sample gave a positive result in the real-time PCR reaction with the CN13-CN14 primer pair, whereas DNA samples from the bacteria produced negative results when reacting with the same primer pair. Next, to confirm that the negative results in the real-time PCR reaction with the CN13-CN14 primer pair on the bacterial DNA samples were truly negative, we performed the PCR reaction with the 27F-1495R primer pair on these DNA samples. 27F-1495R is the primer pair specific for the 16S rRNA gene of all eubacteria with the sequences as follows: 27F (GAGAGTTTGATCCTGGCTCAG) and 1495R (CTACGGCTACCTTGTTACGA). The 27F-1495R primer pair gives the PCR product of ~ 1.4 kb. The PCR results are shown in Fig. 5.

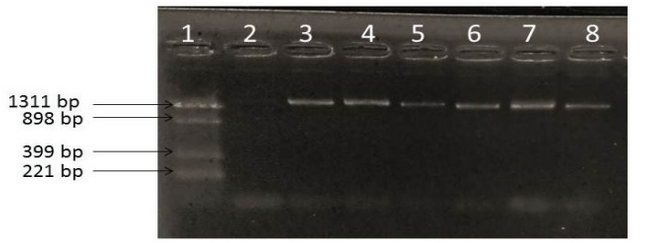

Fig. 5. The PCR results of the bacterial DNA with the 27F-1495R primer pair. Lane 1: DNA ladder, lane 2: negative control; lane 3-8: DNA from *Escherichia coli, Staphylococcus aureus, Pseudomonas aeruginosa, Shigella dysenteriae, Vibrio cholerae*, and *Klebsiella pneumoniae* respectively.

The results in Fig. 5 show that the DNA samples from *Escherichia coli, Staphylococcus aureus, Pseudomonas aeruginosa, Shigella dysenteriae, Vibrio cholerae,* and *Klebsiella pneumoniae* were positive for PCR with the 27F-1495R primer pair. This result confirmed that the negative results in the real-time PCR reaction with the CN13-CN14 primer pair on the bacterial DNA samples were truly negative. Thus, the real-time PCR HRM assay was specifically designed to genotype rs11077 in human.

Precisions

We genotyped rs11077 using the real-time PCR HRM assay on three samples of A/A, A/C, and C/C genotypes five times in the same experiment batch to measure the repeatability of the assay. Additionally, we genotyped rs11077 using the real-time PCR HRM assay on three samples of A/A, A/C, and C/C genotypes five times in the five experiment batches to measure the reproducibility of the assay. The repeatability and reproducibility were expressed by the coefficient of variation (CV), and the CV values were calculated as percentages. The CV value of the repeatability test of the real-time PCR HRM assay was 0.083%, and the reproducibility test results produced a CV value of 0.353%. These CV values proved the high precision of the real-time PCR HRM protocol for genotyping rs11077.

Evaluating the real-time PCR HRM protocol on 123 human blood samples

We evaluated the performance of the real-time PCR

HRM protocol for genotyping rs11077 on the 123 human blood samples that were supplied by Center for Research and Application in Bioscience. The real-time PCR HRM results are shown in Fig. 6.

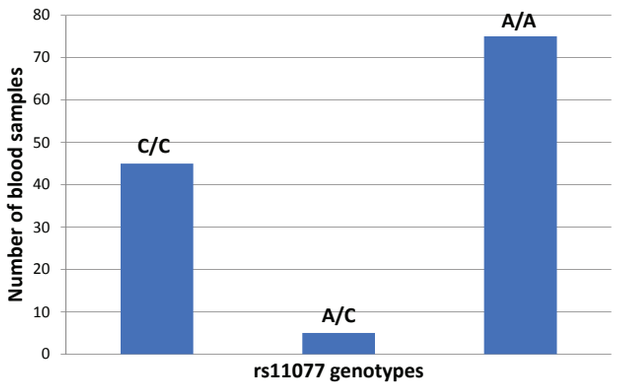

Fig. 6. rs11077 genotyping results on 123 human blood samples using the real-time PCR HRM assay.

Based on the real-time PCR HRM results on 123 human blood samples, 75 samples were found to contain genotype A/A, 5 samples contained genotype A/C, and 43 sample contained genotype C/C. We also performed nucleotide sequencing on 10 random samples out of the 123 samples to confirm the genotyping results by the real-time PCR HRM assay (supplementary data). The comparison showed that the genotyping results on rs11077 between the real-time PCR HRM assay and the Sanger's nucleotide sequencing were identical.

Discussion

There have been many studies on the association of SNPs with human diseases. In HCC, it has been found that a number of SNPs have been associated with this disease, particularly during the progression of the disease. The different genotypes of an SNP can affect the disease differently, which means a certain genotype of an SNP that can make a person more susceptible to develop a disease than another. Among the SNPs related to HCC, such as rs3783553, rs16405, rs99985, rs2910164, and rs11614913, we chose rs11077 to construct a molecular assay for genotyping using real-time PCR HRM. This is because this SNP is heavily involved in the prognostic of HCC patients.

Sanger's nucleotide sequencing is the gold standard for genotyping SNPs, but it is seemed impractical to employ the same in this method in the clinical practice for genotyping rs11077 owing to its complicated process and the required expensive apparatus. In research, other molecular methods such as PCR-RFLP, real-time PCR, molecular hybridisation were used for SNP genotyping. In this study, we chose the real-time PCR HRM method to genotype rs11077, because it has a simple operation procedure but still gives accurate results. In order to genotype rs11077 by distinguishing the melting curve based on HRM analysis, the size of the target PCR product must be so small that a single nucleotide change in the homologous nucleotide sequences can change the melting curve of these nucleotide sequences. Typically, the target PCR size for HRM analysis is from 100-300 bp; however, in this study, we designed the primer pair amplifying the target PCR of only 51 nucleotides. The smaller the target PCR, the more distinct will be the genotype of a SNP. In addition, the primer pair designed for the small target PCR product will avoid other SNPs that are adjacent to the target SNP, as these SNPs will interfere with the melting curves of the target SNP. Through an *in silico* analysis, the size of PCR from CN13-CN14 primer pair was found to be 51 bp, excluding neighbouring SNPs of rs11077. Moreover, the CN15-CN16 primer pair occupies most of the nucleotide sequence of the PCR product except for the position of rs11077, indicating that different human DNAs containing rs11077 will be distinguished merely by rs111077. However, the size of the target PCR product was 51 bp with the pair of CN13-CN14 primers, which was not long enough to be analysed by Sanger's nucleotide sequencing. Thus, we designed the CN15-CN16 primer pair to target the Sanger nucleotide sequence to confirm that the rs11077 subtype is real-time PCR HRM. The PCR product size by the CN15-CN16 is 300 bp, which is sufficient for sequencing by the Sanger technique. The results of sequencing by the Sanger technique confirmed that the genotyping results by real-time PCR HRM on rs11077 were accurate.

Finally, we evaluated the performance of the real-time PCR HRM assay for genotyping rs11077 on the 123 clinical blood samples. Results showed that there were 75 samples bearing genotype A/A, 5 samples carrying genotype A/C, and 43 samples carrying genotype C/C. According to the work of Liu, et al. (2014), people carrying genotype A/A will have a poor treatment prognosis with a 3-year survival rate of 24.7% [4]. In this study, the proportion of people

carrying the genotype A/A accounted for 60.09%, which might be one of the important factors contributing to the high mortality of HCC patients in the Vietnamese population. Knowledge of the genetic information that influences the cancer phenotype is essential for health authorities to control primary liver cancer in the community.

Conclusions

We successfully constructed a molecular assay based on the real-time PCR HRM technique for genotyping rs11077, an SNP involved in the prognosis of HCC patients. The real-time PCR HRM for genotyping rs11077 exhibited good performance in terms of analytical specificity, repeatability, and reproducibility. When evaluated on 123 clinical human blood samples, the real-time PCR HRM assay delivered accurate genotyping results. The assay can be developed to be a prognosis tool for the treatment of HCC patients.

ACKNOWLEDGEMENTS

We thank the Center for Research and Application in Bioscience for the human blood samples and bacterial strains used in this study.

REFERENCES

[1] Y.A. Ghouri, I. Mian, J.H. Rowe (2017), "Review of HCC: Epidemiology, etiology, and carcinogenesis", *Journal of Carcinogenesis*, **16(1)**, doi: 10.4103/jcar.JCar_9_16.

[2] L.R. Roberts, G.J. Gores (2005), "HCC: molecular pathways and new therapeutic targets", *Seminars in Liver Disease,* **25(2)**, pp.212-225.

[3] R.X. Zhu, W.K. Seto, C.L. Lai, M.F. Yuen (2016), "Epidemiology of HCC in the Asia-Pacific region", *Gut and Liver*, **10(3)**, pp.332-339.

[4] S. Liu, J. An, J. Lin, Y. Liu, L. Bao, W. Zhang, J.J. Zhao (2014), "Single nucleotide polymorphisms of microRNA processing machinery genes and outcome of HCC", *PLoS One*, **9(3)**, p.e92791.

[5] R. Yi, Y. Qin, I.G. Macara, B.R. Cullen (2003), "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs", *Genes & Development*, **17(24)**, pp.3011-3016.