The relationship between SNP rs895819 (A>G) on miRNA-27a and the breast cancer in the Vietnamese population

- Nguyen Ba Hung Phong
- Tran Thi Hong Minh
- Nguyen Thi Ngoc Thanh
- Nguyen Thi Hue
  University of Science, VNU-HCMC

(Received on 17th November 2015, accepted on 2nd December 2016)

ABSTRACT

Breast cancer (BC), the most common type of cancer among women worldwide, is a polygenic disease which is caused by the interaction of several genes. Understanding the genetic factors for early diagnosis of BC is crucial to ensure the survival of BC patients. MicroRNA 27a (miR-27a), an oncogenic miRNA, has been predicted to target the tumor suppressor ZBTB10 that can regulate many processes of cell. Single nucleotide polymorphism (SNP) rs895819 (miR-27a) alters the structure and function of miR-27a, which has been anticipated to reduce the risk of BC in different populations such as German and Chinese. This study aims to investigate the relationship between the existence of SNP rs895819 (A>G) and the risk of BC using the optimized high resolution melting (HRM) method. 106 BC samples and 117 healthy samples were genotyped using an optimal HRM protocol then statistical analysis was applied to examine the relationship of the SNP. In the case group, the risk G allele accounted for 36 % while in the control group it took up 32 %. Statistic result revealed that rs895819 (A>G) had no significant relationship with the breast cancer (OR=1.119; P=0.46676) in given case-control samples. Although the SNP is significantly related with BC in German or Chinese populations, it is not a potential marker for diagnosis in the Vietnamese population. Further studies investigating relationship between rs895819 (A>G) and breast cancer in the Vietnamese population is not recommended. In future, other SNPs should be investigated with the aim of identification efficient biomarker for early diagnosis of BC in Vietnamese.

Key words: Breast cancer, rs895819, high resolution melting (HRM), miR-27a

INTRODUCTION

Breast cancer (BC) is the leading type of cancer in women worldwide. Excluding lung cancer, BC is the most common cause of cancer death in women [1]. In 2002, there were 1,383,500 BC incidences and these caused 458,400 deaths worldwide [2]. Until 2012, 1,671,000 new cases of BC in women were estimated, caused 522,000 deaths over the world [3]. In Vietnam, the breast cancer incidence has increased significantly during the last decade, from a rate of 13.8 per 100,000 women in 2000 to 28.1 per 100,000 women in 2010, with an estimated of 12,533 breast cancer cases in the country [4]

BC starts when there is a malignant tumor, a group of cancer cells, growing in the tissue of the
breast. There are numerous risk factors that have been known to be the cause of BC including intrinsic factors (host genetic) and extrinsic factors (environmental factors). Although familial inherited genetic factors only account for 5–10% of BC cases, it is an essential factor in the prevention and early detection of BC [1]. Among the genetic factor, microRNAs (miRNA) and their polymorphisms recently have been investigated in the cancer research. MiRNA, a tiny, approximately 18–25 bases in length, non-coding piece of RNA, plays a significant role in the regulation of the gene expression [5]. The regulatory activity of miRNAs is based on their ability to bind to complimentary regions of their target messenger RNAs (mRNAs) to induce the translational repression or mRNA degradation. MiRNAs participate in many biological processes such as cell proliferation, differentiation, apoptosis and development. Any defection in activities of miRNAs can lead to improper function of these processes. Therefore, SNP, minute mutations that has enormous effects in miRNA activities, is a highly potential target for studies of cancer, including BC.

MiR-27a is an oncogenic miRNA located in the chromosome 19 (location 19q13.13, from nucleotide 13,836,440 to 13,836,517; 78 bp). Its important role in the breast cancer development has been demonstrated by many studies [6-8]. The oncogenicity of miR-27a is its expression in cancer cell down-regulates the expression of ZBTB10 in the mRNA/protein level. ZBTB10 is a suppressor of many specific protein (Sp) transcription factors such as Sp1, which induces the expression of the estrogen receptor (ER) and other Sp-dependent genes which are important for cell survival and angiogenesis such as vascular endothelial growth factor (VEGF), VEGF receptor 1 (VEGFR1), or VEGFR2. ER is an important protein in cells which has a role in the regulation of many processes in cells such as proliferation, apoptosis and cell cycle, supporting sustainability of cell. Dis-regulation of ER pathway can cause the dis-function of cells and may lead to cancer. Abnormally inhibited ZBTB10 causes overexpression of these Sp and Sp-dependent genes, and the overexpression of ER leads to the cancer development [6, 7, 8]. The SNP rs895819 is located in the terminal loop of pre-miR-27a. The alteration of the structure and function of miR-27a by this SNP can alter the ER pathway and finally affects the development of BC.

The effect of SNP rs895819 on the expression of BC has been investigated in many populations and these studies afforded contradictory results. In 2010 a study conducted on German women had confirmed that the rare G allele of the SNP had the protective effect against BC in the German population [9]. Later in 2013 a study carried out on Chinese population also yielded similar result [10]. However, a study performed on Italian population in 2012 was in contrast with those two studies by saying that the SNP rs895819 had no with BC [11]. The inconsistence in result of those studies has motivated us to carry out experiments to examine relationship between of the SNP to BC.

In this study, SNP rs895819 was screened on Vietnamese BC patients by HRM method. This method is based on PCR melting (dissociation) curve techniques and is supported by the innovative double-stranded DNA (dsDNA)–binding dyes along with next-generation real-time PCR instrumentation and analysis software. The DNA is first amplified by normal three-steps PCR, then undergone a short melting step where analysis software works, in the aid of signal from DNA-binding dye, to figure out the unique melting pattern of the DNA strand in the form of melting curve, representing the genotype of the SNP.
This study was conducted to examine the relationship between the SNP rs895819 which located on miR-27a in Vietnamese population, which had not been studied before. MiR-27a and its SNP were chosen due to their ability to indirectly target to ER, which play an important role in the BC development. The study was accomplished by using an optimized high resolution melting (HRM) method.

**MATERIALS AND METHODS**

**Collecting samples and DNA extraction**

The interested population in this study was Vietnamese one. Blood samples were collected from patients with positive or negative clinical diagnosis for breast cancer in Oncology Hospital from 2011 to 2014. Samples included 106 BC cases and 117 healthy controls. All of the patients were given the consent forms to sign on. The collected blood was stored in tubes containing EDTA in -20 °C for further use.

DNA from blood samples was extracted by salting-out method followed the protocol [12]. First white blood cells were isolated from the whole blood by centrifugation. The white blood cell was added with cell lysis buffer (TrisHCl 10 mM, sucrose 11 %, MgCl₂ 5 mM and Triton X100 1 %) to be lysed and released cellular components. Then the cell pellet was added with nuclei lysis buffer (TrisHCl 10 mM, SDS 1 %, EDTA 10 mM, sodium citrate 10 mM) to lyse the nuclei and release DNA. Then, the DNA was separated from other components and cell debris by adding NaCl (50 M) and absolute chloroform. The upper aqueous phase containing DNA was then transferred to a new eppendorf. The DNA was precipitated out of the solution by using absolute ethanol, followed by ethanol 70 %. The supernatant was discarded and the precipitated DNA was kept overnight for drying. Finally, the dried and clear DNA was dissolved in water and stored in -20 °C for further use. After extraction, DNA samples measured the absorbance by a NanoDrop 1000 Spectrophotometer (Thermo Scientific, USA). To be chosen for HRM analysis, the DNA sample must have the concentration of 10 ng/ul or higher and the purity (OD value A260/A280) is in the range of 1.6–1.9.

**Development of HRM protocol for genotyping**

This study implemented the HRM technique in which typical three-step thermal cycles are followed by a short heating of PCR product to reach the melting temperature. During the time of rising the temperature, the sensor inside the instrument captures the change of fluorescence signal emitted by dsDNA-binding dye. The signal is analyzed by software ad visualized in the form of the melting curve which represents for three genotypes of the SNP. The detail of HRM principle and result visualization is shown in Figure 1.

The sequence of SNP rs895819 region on the miRNA-27a was identified using Gene Bank database. The sequence and other informations of this SNP could be obtained from the web page http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=895819. As getting the SNP sequence, the primer design was carried out by the Primer3Plus online software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/). The parameters were set as: product size 80–150 bp, primer size 18–25 bp, primer Tm around 60–70 °C (65 °C is the optimum). Then the chosen pairs of primer were given to BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to check whether they were specific for the SNP sequence. Pairs of primer which had high specificity were then used to predict the HRM melting curve of their amplicons using the UmeltHet software (https://www.dna.utah.edu/hets/umh.php). The resolution was adjusted to “Very high – 0.1 °C” and other PCR components such as Mono⁺, Mg²⁺, DMSO concentration were also screened in order to create three distinct melting curves and peaks representing 3 genotypes of the SNP. Beside the...
primers for HRM analysis, an extra pair of primer for sequencing to confirm the genotypes of three positive control samples was also designed.

Primers and amplicons were described in Table 1.

Table 1. Primers for HRM analysis and sequencing

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Melting temperature</th>
<th>Amplicon length</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRM_Forward</td>
<td>5'-GGCAAGGCCAGAGGAGGTGA-3' (20 bp)</td>
<td>67.2 °C</td>
<td>105 bp</td>
</tr>
<tr>
<td>HRM_Reverse</td>
<td>5'-GGCCTGGAGGGCGAGGCTTA-3' (20 bp)</td>
<td>66.1 °C</td>
<td></td>
</tr>
<tr>
<td>Seq_Forward</td>
<td>5'-AGTACGCACGGGACAGGCGAG-3' (20 bp)</td>
<td>64.1 °C</td>
<td>303 bp</td>
</tr>
<tr>
<td>Seq_Reverse</td>
<td>5'-GGGATGGGATTTGGCTTCT-3' (20 bp)</td>
<td>64.5 °C</td>
<td></td>
</tr>
</tbody>
</table>

The HRM optimization procedure was composed of three steps: initial optimization, determination of three control genotypes and final optimization. In the initial optimization step, we aimed to find out the condition of three factors annealing temperature (Ta), MgCl$_2$ and DMSO concentration in which the HRM reaction yielded a good melting curve. To determine the optimal Ta, five PCR reactions with gradient annealing temperatures (60–68 °C) were run using Eppendorf Mastercycler. The reagents included Toptaq Mastermix 1X, 0.2 µM each primer and 50 ng DNA. For optimization of MgCl$_2$ and DMSO concentrations, different concentrations of MgCl$_2$ and DMSO were added to 10 µL reaction mixture and tubes were placed in 96-well plate of LightCycler 96 thermocycler (Roche Diagnostics, Germany). The concentrations of MgCl$_2$ and DMSO in this step were based on the prediction on U.meltHet. The reagents in the reactions included Light Cycler 480 Resolight Dye Mastermix 1X, 0.2 µM each primer HRM Forward/ HRM Reverse, 50 ng DNA and adjusting concentrations of DMSO and MgCl$_2$. Thermal cycles were set as the following: 5 minutes pre-incubation at 95 °C followed by 40 thermal cycles including 30 seconds denaturation at 95 °C, 30 seconds annealing at 66 °C and 30 seconds extension at 72 °C for each cycle; and continued by high resolution melting step including 90 seconds at 95 °C, then 60 seconds at 40 °C, 30 seconds at 65 °C and gradually increasing temperature from 65 to 95 °C. The process was ended by a hold cooling at 37 °C.

After having initial HRM protocol, few samples were applied in order to find out three control genotypes. Eight samples were run on HRM analysis and three distinct groups of melting curve, representing for three genotypes AA, AG and GG, were obtained. Random samples from each genotype were sequenced to confirm their exact genotype. The reagents for PCR reaction to prepare for sequencing included Toptaq Mastermix 1X, 0.2 µM each primer Seq_Forward/ Seq_Reverse and 50 ng DNA sample. Finally we had three positive controls which represented the three genotypes of the SNP. As three positive controls were determined, optimization had to be conducted again in order to obtain the clustered and distinct melting curves of three controls together. The adjustment of the MgCl$_2$ concentration was carried out one more time.
Figure 1. HRM principle and result visualization of three genotypes of a SNP. (A) Thermal setting-up of HRM analysis by LightCycler® 96 Real-time PCR system. (B1) Melting curves. (B2) Normalized melting curves. (B3) Melti
ng peaks. (B4) Difference plots

Genotyping and analysis

The optimal HRM conditions were applied on 106 cases and 117 controls for genotyping. The reagents in one reaction consisted of Lightcycler 480 Resolight dye Mastermix 1X, MgCl₂ 2.5 mM, HRM_Forward/Reverse primer 0.2 mM and 45 ng DNA sample. The setting-up of thermal cycles was the same as in the optimization step. In each running time, three positive controls and one negative control were included in the plate. The results were displayed using LightCycler® 96 SW 1.1 software. The abnormal-melting-curve samples and shifted-melting-curve samples were subjected to be run again. The repeatedly failed samples were eliminated from analysis. The samples that exhibited identified genotype were then applied to calculate genotypic frequency to prepare for the analysis. As in analysis, Chi-squared test was applied using STATA.

RESULT

Initial HRM conditions

For the Ta optimization, the PCR reaction with gradient temperature ranged from 60 to 68 °C was run and the result was analyzed using gel electrophoresis. The reactions exhibited good amplification and no extra bands in all Ta (data not shown). The reactions with Ta in the range of 60-66 °C, however, gave bolder and brighter band on the gel. As increasing Ta, the specificity of primers is increased, so 66 °C was chosen as the Ta for further HRM analysis.

For the MgCl₂ and DMSO concentration optimization, firstly the parameters that were predicted by using Umelt (2 mM Mg²⁺ and 10 % DMSO) were applied to the HRM reaction. The result, however, failed to identify genotypes. To check what were the optimal concentrations of MgCl₂ and DMSO, we had carried out several HRM reactions with different concentrations of DMSO and MgCl₂. Finally MgCl₂ 3 mM was
chosen as the initial condition for further screening and no DMSO was needed.

**Determination of three positive controls**

After obtaining HRM conditions for the first sample, it was proceeded to apply on 8 random samples in order to find out three distinct genotypes. Three groups of curves that could be assumed: 2 homozygote genotypes and a heterozygote genotype were obtained (Figure 2). The curves were well-separated into three groups of assumed genotype. Then selected samples, one from each group of curves was picked and sequenced to confirm the exact genotype of each sample.

Based on the sequencing result given in Figure 3, the homozygote AA was belong to the melting curve group which had a lower Tm of 88.4°C, while homozygote GG had Tm of 88.8°C. The sample melting at two different temperture showed two peaks A and G at the SNP position which represented the heterozygote genotype (Figure 3C). Those three genotype-determined samples were chosen to be three positive controls for later genotyping.

![Figure 2. HRM analysis for 8 random samples. (A) Melting peaks. (B) Different plots](image)

![Figure 3. Sequencing results of three control samples. (A) Homozygote of A allele. (B) Homozygote of G allele. (C) Heterozygote AG](image)
Optimal the HRM protocol for genotyping

As could be seen in Figure 2, the curves of three groups of genotype were not grouped very tightly together, especially heterozygote curves. On the other hand, 3 control samples for 3 genotypes have been determined, so it was decided to re-optimize the conditions with three controls to get a better condition for further analysis without confusing. Since MgCl$_2$ was the component that might influence the organization of the melting curve between three genotypes, MgCl$_2$ concentration was optimized again. HRM was run on three controls with different MgCl$_2$: 2.5 mM, 3 mM and 3.5 mM. The result revealed that at MgCl$_2$ 2.5 mM, three melting curves exhibited the best shape and discrimination.

Afterward, we got the final optimal protocol for the HRM analysis of the SNP rs895819 which was given in Table 2 and 3.

Table 2. Thermal cycles of HRM analysis for genotyping SNP rs895819

<table>
<thead>
<tr>
<th>Thermal cycle steps</th>
<th>Temperature Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Preincubation</td>
<td>95 °C-300 s</td>
</tr>
<tr>
<td>2. Amplification</td>
<td>95 °C-30 s&lt;br&gt;66 °C-15 s&lt;br&gt;72 °C-30 s</td>
</tr>
<tr>
<td>3. High resolution melting</td>
<td>95 °C-60 s&lt;br&gt;40 °C-60 s&lt;br&gt;65 °C-30 s&lt;br&gt;95 °C-1 s</td>
</tr>
<tr>
<td>4. Cooling</td>
<td>37 °C-30 s</td>
</tr>
</tbody>
</table>

Table 3. Reagents of HRM analysis for genotyping SNP rs895819

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master mix</td>
<td>1X</td>
</tr>
<tr>
<td>MgCl$_2$</td>
<td>2.5 mM</td>
</tr>
<tr>
<td>DMSO</td>
<td>0 %</td>
</tr>
<tr>
<td>DNA</td>
<td>30 ng/ul</td>
</tr>
<tr>
<td>Forward and Reverse primer</td>
<td>0.2 µM</td>
</tr>
</tbody>
</table>

Genotyping and analysis

After having all the optimal conditions for HRM analysis, it was applied to 106 cases and 117 controls. HRM was run several times, each time we performed 16–24 reactions, including three positive controls and one negative control. The HRM successful rate was 92 %. Most of reactions yielded good result in which the curves in one group clustered tightly to each other and well-separated with other groups (Figure 4). Table 4 shows the detail genotyping as well as the analysis result of the SNP. Both case and control populations were in the Hardy – Weinberg equilibrium. The Chi-squared test showed the p-value for allele difference between two populations was P=0.45413, which was much higher than the threshold (P=0.05). Therefore, there was no significant association between this SNP and BC in Vietnamese population.
DISCUSSION

This study was carried out with the aim to examine the relationship of the SNP rs895819 with the aid of HRM analysis, a powerful molecular technique. In this study, a complete optimal HRM protocol was successfully developed to genotype the SNP in case-control samples. Compared to other techniques, HRM has some advantages such as the rapidity and high sensitivity. The sensitivity of HRM is the most important criterium. Even though the sequences are slightly different, the instrument can detect the difference and shows distinct curves. Furthermore, HRM is a single-step and closed-tube PCR, followed by a quick heating step and no requirement of gel electrophoresis for analysis as other technique such as Taqman genotyping or HPLC [13, 14]. In conclusion, the analysis result showed that the SNP rs895819 had no significant relationship with BC in Vietnamese population. Previous studies had investigated the relationship of rs895819 with BC in different populations. One of the initial studies was conducted by Yang et al. in 2010. The study revealed that SNP rs895819 had a protective effect on the BC development in German women. The rare G allele of rs895819 was less common in the cases than in the controls, indicating a reduced familial breast cancer risk ([G] vs. [A]: OR = 0.88, 95 % CI 0.78–0.99, P = 0.0287) [9]. Later in 2013, Ning Zhang et al. [10] showed a similar result in which people with the G allele exhibited a significantly decreased risk of BC.

Trang 46
relative to people carrying the A allele among the younger Chinese population ([G] vs. [A]: OR = 0.628, 95% CI 0.410–0.961; P = 0.032). On the other hand, a study conducted in 2012 by Irene Catucci et al. [11] performed on Italian population revealed a contradictory result. They investigated the effect of rs895819 on BC risk by genotyping the 1,025 Italian familial BC cases and 1,593 controls. The result failed to support the relationship between rs895819 and the breast cancer risk ([G] vs. [A]: OR = 0.91, 95% CI 0.80–1.02, P = 0.114). The contradictory between those studies implies that although the SNP was proved to have a functional effect on its miRNA and (due to the effect of miR-27a to ER) the cancer development, the relationship between this SNP and the risk of BC is varied among different populations. This SNP might be a risk factor of BC in German and Chinese population but could not be proved to associate with the BC in the Vietnamese population.

The non-association of rs895819 in Vietnamese population can be explained by some reasons. The SNP may have a down regulation effect on miR-27a but this effect may be interrupted by activities of other SNPs located in this miRNA or protein ZBTB10 was down-regulated by other SNPs or transcription factors, making the initiation of BC unchanged. In another scenario, this SNP may have significant effect on ER and through that inhibits the initiation of BC but it does not affect the risk of developing the disease because other SNPs or genetic factors may contribute to development of the disease.

This is the first time the SNP rs895819 was examined in the Vietnamese population using the novel HRM method. The method exhibited effectiveness in terms of sensitivity, specificity, task and time saving for the detection of genetic polymorphism. In the future, the protocol to genotype this SNP could be applied to any population is not only Vietnamese.

CONCLUSION

In conclusions, a HRM analysis was successfully developed for genotyping SNP rs895819. This method exhibited many advantages that make it become a potential choice for detecting genetic variants for diagnosis of diseases. This study also showed that in 97 breast cancer cases and 100 healthy samples the frequencies of A and G allele of rs895819 are 64.44% and 35.56% in the cases group; whereas in the control group, accounting for 68% and 32%, respectively. The SNP rs895819 in miR-27a was not associated with the breast cancer in the Vietnamese population (OR = 1.173, 95% CI 0.772–1.782, P = 0.45413). In the future more SNPs need to be investigated in order to find out effective genetic markers for diagnosis of the breast cancer in Vietnamese.

Acknowledgement: This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 106-YS.01-2013.09. The authors thank to the Oncology Hospital – HCMC who supported for the sample collection.
Liên quan giữa SNP rs895819 (A>G) trên miR-27a và bệnh ung thư vú ở người Việt Nam

- Nguyễn Bá Hùng Phong
- Trần Thị Hồng Minh
- Nguyễn Thị Ngọc Thanh
- Nguyễn Thị Huệ

Trưởng Đại học Khoa học Tự nhiên, ĐHQG-HCM

TÓM TẮT

Ung thư vú (UTV), loại ung thư phổ biến nhất ở phụ nữ toàn thế giới, là bệnh di truyền gây ra bởi sự tương tác của nhiều gen khác nhau. Việc hiểu được những yếu tố di truyền liên quan đến UTV nhằm chẩn đoán sớm bệnh góp phần làm tăng khả năng sống sót của bệnh nhân. MicroRNA-27a (miR-27a), một miRNA gây ung thư, được dự đoán là có tác dụng trên ZBTB10, một phần tử ức chế khối u, nhằm điều hòa nhiều hoạt động chuyển hóa của tế bào. Điểm đa hình (SNP) rs895819 thay đổi cấu trúc và chức năng của miR-27a, điều này được chứng minh là làm giảm nguy cơ UTV ở một số quốc gia, như Đức và Trung Quốc. Tóm tắt nghiên cứu này nhằm khảo sát mối liên quan giữa SNP rs895819 (A>G) và UTV bằng cách sử dụng phương pháp phân tích nhiệt độ nóng chảy (High Resolution Melting). 106 mẫu DNA UTV và 117 mẫu DNA khỏe mạnh được xác định kiểu gen bằng phương pháp HRM đã tốt ưu hóa và phân tích thống kê để áp dụng để kiểm tra mối liên quan giữa SNP và UTV. Allel G chiếm 36% trong nhóm bệnh UTV và 32% trong nhóm khỏe mạnh. Kết quả thống kê cho thấy rằng rs895819 (A> G) không có liên quan đáng kể đến ung thư vú (OR = 1,119; P = 0,46676). Mặc dù SNP này có liên quan đáng kể đến UTV của người Đức và Trung Quốc, nhưng không phải là một dấu ấn sinh học tiềm năng để chẩn đoán UTV ở người Việt Nam. Trong tương lai, các SNP khác cần được nghiên cứu để xác định các dấu ấn sinh học tiềm năng trong chẩn đoán sớm UTV ở người Việt Nam.

Từ khóa: UTV, rs895819, phân tích độ nóng chảy, miR-27a

REFERENCES

[7]. S.U. Mertens-Talcott, et al., The oncogenic microRNA-27a targets genes that regulate specificity protein transcription factors and the G2-M checkpoint in MDA-MB-231


