HNUE JOURNAL OF SCIENCE DOI: 10.18173/2354-1059.2022-0051
Natural Sciences, 2022, Volume 67, Issue 3, pp. 153-161
This paper is available online at http://stdb.hnue.edu.vn

OPTIMIZING GENOTYPING PROCEDURE AND DETERMINING FREQUENCY OF SINGLE NUCLEOTIDE POLYMORPHISM RS7799039 BELONGING TO *LEP* GENE IN FOUR-YEAR-OLD BOYS IN HANOI

Nguyen Thi Trung Thu, Pham Bui Quang Minh and Le Thi Tuyet*
Faculty of Biology, Hanoi National University of Education

Abstract. The *LEP* (leptin) gene plays an important role in the regulation of food intake and energy metabolism. The single nucleotide polymorphism (SNP) rs7799039 located in the promoter of the *LEP* gene may affect the risk of obesity and other diseases such as diabetes, cardiovascular disease, cancer, etc. due to altering the gene expression. The objective of this study is to find out the optimum procedure for genotyping SNP rs7799039 by using the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method and determine the allele and genotype frequencies of this SNP in 61 four-year-old boys with normal nutritional status in Hanoi. The sequence of primers has been designed as follows: forward primer 5'-tttcctgtaattttcccatgag-3' and reverse primer 5'-aaagcaaagacaggcataaaa-3'. The thermal cycle for the PCR reaction includes initiation at 94 °C (3 minutes), 34 cycles of denaturation at 94 °C (30 seconds), annealing at 56 °C (30 seconds), extension at 72 °C (30 seconds), final extension at 72 °C (8 minutes), and then keep the products chilling at 4 °C. *Hha*I restriction enzyme was used to distinguish A and G alleles of rs7799039. PCR products are incubated with 0.5 µl *Hha*I restriction enzyme for 15 minutes at 37°C and then the products of restriction enzyme digestion are separated by gel electrophoresis. Genotype frequencies of our studied population are 62.3% AA, 31.1% AG, and 6.6% GG (P_{HWE} = 0.452).

Keywords: rs7799039, *LEP* gene, PCR-RFLP, genotyping, 4-year-old boys.

1. Introduction

The leptin (*LEP*) gene, also known as the obese gene, is located on chromosome 7q32.1 in the human genome, and has crucial roles in energy homeostasis, regulation of appetite as well as other physiological processes [1]. The *LEP* gene, which is expressed mainly in white adipose tissue, encodes for the hormone leptin - a 167-amino-acid secretory protein that acts on the satiety center in the hypothalamus [1]. In the brain, the major targets of leptin are two types of neurons in the arcuate nucleus called POMC

Received October 10, 2022. Revised October 21, 2022. Accepted October 28, 2022.
Contact Le Thi Tuyet, e-mail address: tuyetlt@hnue.edu.vn

(pro-opiomelanocortin expressing neurons) and AGRP (agouti-related protein-expressing neurons). The binding of leptin to these neurons leads to changes in gene expression, which results in an inhibitory impact on food intake and promotes energy expenditure [2]. Many mutations in the *LEP* gene, which alter the normal function or the stability of leptin, are demonstrated that linked to both monogenic obesity and polygenic obesity [2-3].

Based on the advantages of genomic analysis, genome-wide association (GWA), and meta-analysis, researchers have identified many single nucleotide polymorphisms (SNPs) of the *LEP* gene that may have an association with obesity such as rs7799039, rs2167270, rs4731426 [4]. Among these variants, the SNP rs7799039 (A/G) is a well-studied polymorphism showing a strong relationship with obesity in many studies [5-8]. SNP rs7799039 is a single nucleotide substitution that occurs in the promoter of the *LEP* gene, which may affect the gene expression due to altering the binding of transcription factors [9]. Previous studies indicated that rs7799039 is not only associated with obesity but also obesity-related traits such as diabetes mellitus, cardiovascular diseases, and cancer [10-12]. In addition, the effect of rs7799039 on an individual's phenotype depends on ethnic population, gender, age, nutrition status, physical activity, and other factors [3]. Thus, identifying the genotype of this polymorphism in the Vietnamese population will provide a useful database for studying the association of rs7799039 with obesity, as well as other diseases, and it is also beneficial for predicting and preventing obesity in childhood.

There is a variety of methods to determine the genotype of SNP rs7799039 such as TaqMan allelic discrimination assays, genome sequencing [4, 13]. However, the application of these approaches to identify genotypes on a large scale, especially in Vietnam, is difficult to perform because of the limitation of equipment, high-cost chemicals, and biological materials. The polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) method is a more suitable method that can apply for genotyping in a large number of samples in Vietnam because this method is not required high-tech equipment, just basic equipment like thermal cyclers, gel electrophoresis machines. The principle of this method is based on the difference in length of restriction fragments created by a particular restriction enzyme. The substitution of nucleotide in the SNP leads to alteration of the restriction site of the restriction enzyme, thus resulting in different bands (corresponding with different alleles) that can be observed on the electrophoresis image. This cheaper method can be used for almost labs in our country and has been applied to determine the genotype of many SNPs in Vietnamese populations.

Up to now, there has not been any study that applies the PCR-RFLP method for genotyping SNP rs7799039 in the Vietnamese population. Thus, the aim of this study is to find out the optimum procedure for identifying the genotype of rs7799039 and to analyze the frequency of this SNP on four-year-old boys in Hanoi.

2. Content

2.1. Research methodology

2.1.1. Study objects

This research was carried out on 61 four-year-old boys in Hanoi that were selected randomly from the objects of the B2018-SPH50 project. The project was approved by the Medical Ethics Council of the Institute of Nutrition with Decision No. 343/VDD-QLKH on July 27, 2018. All children in this study had normal nutritional status according to the criteria of the World Health Organization (WHO, 2006). Children were classified with normal nutritional status as Z-score BMI by age and gender ranging from -2 to 2. Any children who had chronic or acute diseases such as tuberculosis, HIV/AIDS were excluded from this study. Genetic analysis procedures were performed in the central lab of Hanoi Medical University. The study was carried out from March 2022 to June 2022.

2.1.2. Research materials

DNA samples: DNA was extracted from the sample of the cheek mucosa cell using the GeneJET Genomic DNA Purification kit (Thermo, USA) according to the manufacturer's instructions. All the DNA samples have the A260/A280 ratio ranging from 1.8 to 2 to ensure the purification of the study.

Chemicals and biological materials: specific primers in 10 pmol/µl concentration (Invitrogen, USA), PCR master mix 2X (0.05 u/µl Taq DNA polymerase, 4 mM $MgCl_2$, dNTP 400 µM), New England Biolabs® (NEB) *Hha*I restriction enzyme (ER1851), Ultrapure™ 10x TBE buffer, RedSafe™ nucleic acid staining solution (Intron Biotechnology) and ΦX174 DNA HaeIII Digest markers.

Equipment: PCR machine (Eppendorf's Master cycler), Mupid-Exu gel electrophoresis system (Japan), Dry bath, fridge, centrifuge, Geldoc-It™ gel camera.

2.1.3. Methodology

**** Amplification of LEP rs7799039 sequences***

To amplify the DNA fragments containing the SNP rs7799039, suitable primers that are complementary to each end of the target sequence need to be designed. The primers used for PCR reaction were designed according to Amer et al [14]. The sequence of forward and reverse primers were 5'-tttcctgtaattttcccatgag-3' and 5'- aaagcaaagacaggcataaaa-3', respectively.

The melting temperature of the two primers used for the PCR reaction was about 52 - 60 °C [15]. To determine the best annealing temperature, two DNA samples were selected to carry out PCR reaction with the following testing temperature: 52 °C, 54 °C, 56 °C, 58 °C, and 60 °C. PCR amplification was carried out in a volume of 12 µL containing 2.4 µl of nuclease-free water, 0.8 µl *LEP* rs7799039 forward primer, 0.8 µl *LEP* rs7799039 reverse primer, 6 µL GoTaq® Green PCR master mix 2 X (Promega Corporation, USA), and 2 µL genomic DNA.

The thermal cycle for the PCR reaction was conducted as follows: the first denaturation for 3 minutes at 94 °C, 34 cycles of denaturation at 94 °C for 30 seconds, annealing at 52 - 60 °C for 30 seconds, elongation at 72 °C for 30 seconds, final extension at 72 °C for 8 minutes, and finally keeping the products chilling at 4 °C. Five µL of PCR

products were stained with RedSafe and electrophoresis on 3% agarose gel for 30 minutes at 100 V in 0.5 X TBE buffer, with ΦX174 DNA HaeIII Digest marker. The DNA bands were detected using a Geldoc-It^{TM} gel camera. The best annealing temperature will be chosen based on this testing result.

**** Determine optimum enzyme concentration for cutting reaction***

The *Hha*II restriction enzyme was selected (New England Bio-labs, Beverly, MA) to digest PCR product, using the restriction mapper database [14, 16-17]. The optimum enzyme concentration for DNA digestion was determined by testing the effectiveness of cutting reaction at different amounts of enzyme. Three DNA samples with different genotypes at the *LEP*-rs7799039 locus were selected to conduct a PCR reaction, and then the presence of PCR products was tested by gel electrophoresis. After that, PCR products were incubated with *Hha*I enzyme at three different enzyme concentrations: 1.0 µl, 0.5 µl, and 0.25 µl. The composition of the 10 µl enzymatic incubation reaction included: 1.0 µl of buffer (10x fast digest, Cutsmart), 1.0 µl/0.5 µl/0.25 µl of *Hha*I enzyme (NEB), and 5.0 µl of PCR product and the rest of water (nuclease-free). These mixtures then were incubated at 37 °C for 15 minutes. Finally, DNA fragments obtained from DNA digestion were separated by gel electrophoresis to evaluate the optimum amount of enzyme for the cutting reaction to occur.

**** Determination of LEP-rs7799039 polymorphism***

The representative 61 boys were genotyped to determine frequencies of genotypes and alleles of *LEP*-rs7799039 polymorphism.

**** Statistical analysis***: We used $\chi 2$ test and SPSS 16.0 software to compare genotype and allele frequencies (expressed in percentage) between groups. Genotype frequencies were compared and tested for Hardy Weinberg Equilibrium (HWE) by chi-square test. The p-value less than 0.05 represents a significant statistical difference.

2.2. Results and discussion

2.2.1. Optimizing the PCR-RFLP procedure for genotyping *LEP* rs7799039

**** Determine optimum annealing temperature for PCR reaction***

The final primers for the PCR reaction, which were designed based on Oligo 7 Primer Analysis software, had the following sequence: forward primer 5'-tttcctgtaattttcccatgag-3', reverse primer 5'-aaagcaaagacaggcataaaa-3'. According to recommendations, the melting temperature of the primer was approximately 56 to 58 °C [14]. To determine the best annealing temperature for the PCR reaction, the PCR gradient method was carried out. Therefore, to select the appropriate annealing temperature (Ta) for PCR, we conducted a test with 2 samples and negative control (H_2O) at 5 following annealing temperatures: 52 °C, 54 °C, 56 °C, 58 °C, and 60 °C. The result is shown in Figure 1.

The obtained PCR product had 242 bp. According to the electrophoresis image, the thickest bands were observed at 56 °C. Thus, this was the best annealing temperature. Based on the results, the optimum thermal cycle for PCR reaction was identified as follows: initiation at 94 °C in 3 minutes, 34 cycles of denaturation at 94 °C in 30 seconds, annealing at 56 °C in 30 seconds, extension at 72 °C in 30 seconds, final extension at 72 °C in 8 minutes, and finishing at 4 °C.

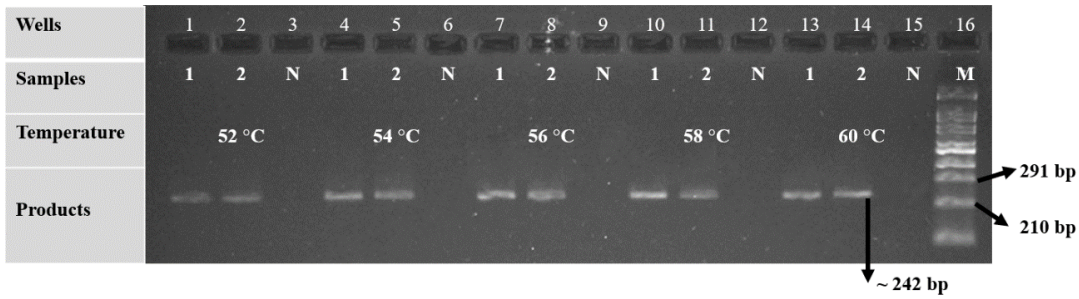

Figure 1. Electrophoresis image of PCR products at different annealing temperatures
Samples: 1 and 2, N: negative control (H_2O), M: marker ΦX174 DNA HaeIII Digest.

**** Determine optimum enzyme concentration for cutting reaction***

PCR products of three selected DNA samples were detected by gel electrophoresis. DNA segments containing the rs7799039 sequence (242 bp) were successfully amplified in all DNA samples. This result was represented in Figure 2. To determine the genotypes of samples, we used *Hha*I restriction enzyme to digest the PCR products because this enzyme has the restriction site $\begin{matrix}5'\ldots GCG^{\blacktriangledown}C\ldots 3'\\ 3'\ldots C_{\blacktriangle}GCG\ldots 5'\end{matrix}$ that matches with the sequence containing rs7799039. After restriction digestion, different alleles of rs7799039 can be distinguished by gel electrophoresis. Figure 3 showed the electrophoresis photo of DNA fragments derived from cutting reactions at different enzyme concentrations.

The replacement of nucleotide G with nucleotide A at the site of rs7799039 leads to the loss of the restriction site of the *Hha*I enzyme, thus this enzyme only cuts DNA fragments of allele G, but not allele A. The products of DNA digestion by restriction enzyme include two types of DNA fragments: 182 bp fragments and 60 bp fragments. These DNA fragments with a length of 60 bp can not be observed on the electrophoresis photo due to their small size. Therefore, the two homozygous AA and GG genotypes would exhibit only a single band of 242 bp and 182 bp, respectively. The heterozygous AG genotype would exhibit two bands 242 bp and 182 bp.

Regarding evaluating the effectiveness of DNA digestion at different amounts of restriction enzyme, the electrophoresis image at the enzyme concentration of 1.0 µl (following the instruction of the producer) was used as the control. Three DNA samples 1, 2, and 3 had different genotypes: AA, AG, GG, respectively due to the difference in the number and the location of bands that appeared on the electrophoresis image. At the concentration of 0.25 µl, the result can not be used to distinguish between AG and GG genotypes because both samples exhibited two bands (242 bp and 182 bp), which meant the cutting reaction incompletely occurred. In contrast, at the enzyme concentration of 0.5 µl, the GG genotype exhibited only a single band, which was different from the band patterns of the AG genotype, and the bands were easy to observe as those at 1.0 µl enzyme concentration. This indicated that at the enzyme concentration of 0.5 µl, the cutting reaction also occurred with the same effectiveness as at 1.0 µl enzyme concentration. Thus, the optimum enzyme concentration for restriction enzyme digestion is 0.5 µl. By using the *Hha*I enzyme at the concentration of 0.5 µl, we can save up to 50% amount of the enzyme needed for genotyping protocol.

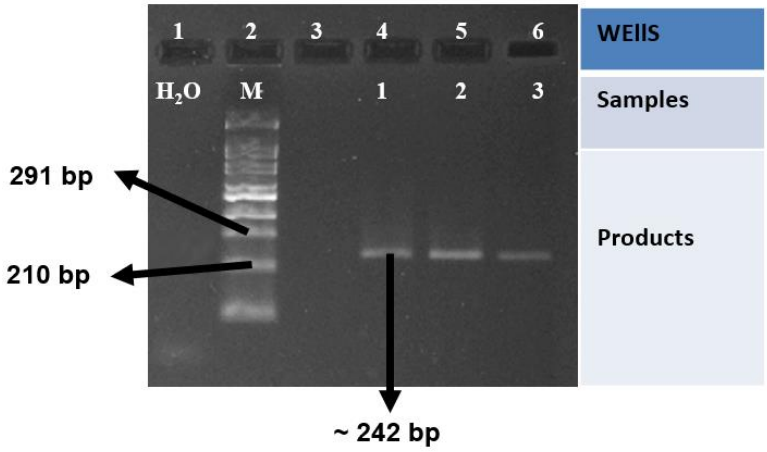

Figure 2. Electrophoresis image of PCR products derived from three DNA samples (1, 2, and 3) that were selected for testing the effectiveness of cutting reaction at different enzyme concentrations

N: negative control (H_2O), M: marker ΦX174 DNA HaeIII Digest

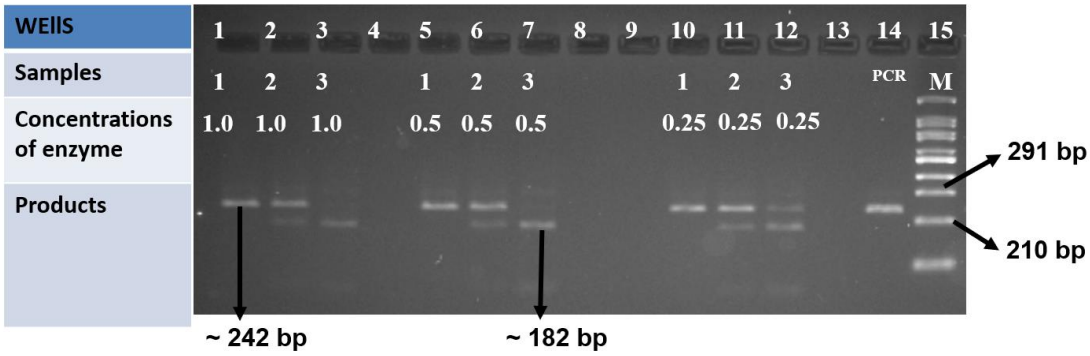

Figure 3. Electrophoresis image of incubated products at different enzyme concentrations

The HhaI enzymes were incubated with PCR products at three concentrations: 1.0 µl, 0.5, and 0.25 µl to test the effectiveness of cutting reaction. Three DNA samples 1, 2 and 3 had different genotypes: AA, AG, and GG, respectively. Restriction enzyme digestion was completely done at 0.5 µl but not at 0.25 µl. M: marker ΦX174 DNA HaeIII Digest

Following this optimum procedure, a large number of samples can be genotyped without using too many expensive chemicals or high-tech equipment. We have successfully applied this procedure for genotyping rs7799039 on a large scale with about 18 samples per agarose gel (Figure 4). The genotype of each sample can be quickly identified based on the forming of different restriction fragments, which correspond to different types of bands appearing on the electrophoresis photo.

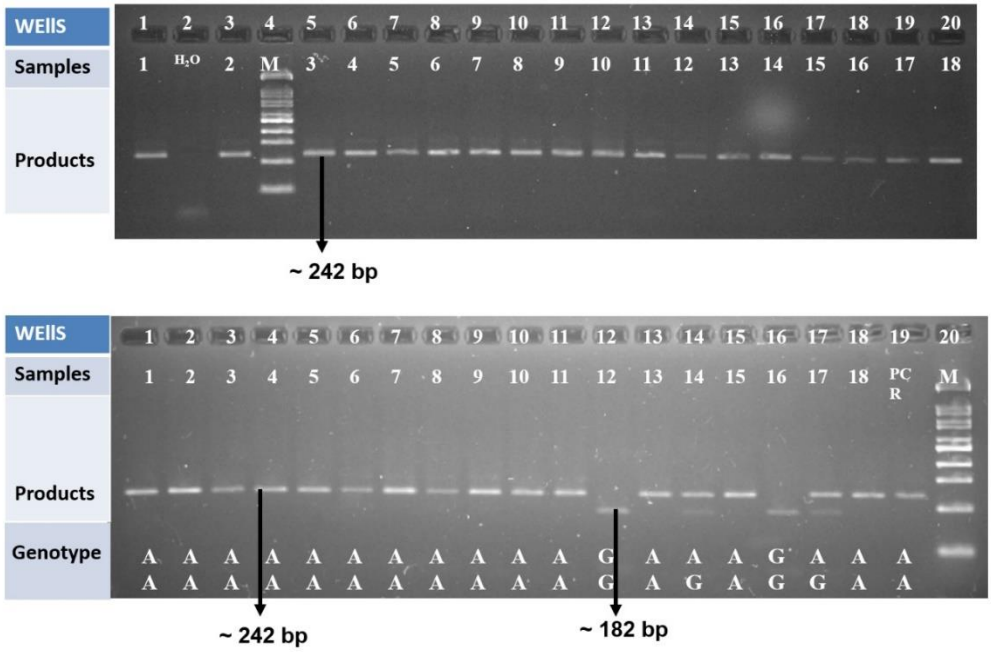

Figure 4. Applying optimum procedure for genotyping on a large scale

PCR: PCR product, M: marker ΦX174 DNA HaeIII

2.2.2. Frequencies of *LEP* rs7799039 polymorphism in 4-year-old boys in Hanoi

DNA samples of 61 four-year-old boys with normal nutrition status were genotyped at the SNP rs77993039 locus by PCR-RFLP method. Following the mentioned procedure, genotypes of all DNA samples were successfully detected. The distribution of alleles and genotypes at the *LEP* rs7799039 polymorphism is represented in Table 1.

Table 1. Genotype and allele frequencies of LEP rs7799039 polymorphism in Hanoi four-year-old boys

Frequencies		Samples (n = 61)
Genotype	AA	38 (62.3%)
	AG	19 (31.1%)
	GG	4 (6.6%)
Allele	A	95 (77.9%)
	G	27 (22.1%)
P_{HWE}		0.452

HWE: Hardy-Weinberg equilibrium. P value was from the χ2 test

The data showed that among four-year-old boys in Hanoi, the AA genotype had the highest frequency (62.3%), and the lowest frequency belonged to the GG genotype (6.6%). The distribution of genotypes at *LEP* rs7799039 polymorphism was in Hardy-Weinberg equilibrium (P = 0.452). In our studied population, the minor allele was the G allele with the frequency of 22.1%, and the major allele was the A allele (77.9%). Frequencies of these alleles were nearly the same as the allele frequencies of the Japanese population in Tokyo, Japan (frequency of the G allele was 23.6%) [18], and the dominance of the A

allele was also a similar characteristic compared with other East Asian populations such as the Han Chinese population in Beji, China (frequency of A allele = 75.7%), the Kinh population in Ho Chi Minh City, Vietnam (frequency of A allele = 67.7%) [18]. The difference in allele frequencies between populations around the world may be explained by ethnic features, nutrition factors, physical activities, economic and social factors... Due to the limitation of the sample size, further studies on large scale need to be conducted to investigate the allele and genotype frequencies of the entire Vietnamese population at the *LEP* rs7799039 locus.

3. Conclusions

Our study found the optimum procedure for genotyping of *LEP* rs7799039 polymorphism by the PCR-RFLP method. The sequences of primers are forward primer 5'-tttcctgtaattttcccatgag-3' and reverse primer 5'-aaagcaaagacaggcataaaa-3'. The composition of the PCR reaction includes 2.4 µl of nuclease-free water, 6.0 µl PCR Master Mix, 0.8 µl *LEP* rs7799039 forward primer, 0.8 µl *LEP* rs7799039 reverse primer, and 2.0 µl DNA sample. The thermal cycle for the PCR reaction was carried out as follows: initiation at 94 °C in 3 minutes, 34 cycles of denaturation at 94 °C in 30 seconds, annealing at 56 °C in 30 seconds, extension at 72 °C in 30 seconds, final extension at 72 °C in 8 minutes, and finishing by keeping the products chilling at 4 °C. PCR products were incubated with *Hha*I restriction enzyme, and the composition of incubation reaction includes 3.0 µl of water (nuclease-free), 1.0 µl of buffer (10x fast digest, Cutsmart), 0.5 µl of *Hha*I enzyme (NEB), and 5.0 µl of PCR product. This mixture was incubated at 37 °C for 15 minutes. The enzyme digestion products were stained with RedSafe and then separated by electrophoresis on 3% agarose agar at 100 V for 30 minutes, 0.5 X TBE buffer.

Genotype frequencies of *LEP* rs7799039 polymorphism in normal Hanoi boys aged 4 years old are AA - 62.3%, AG - 31.1%, GG - 6.6%. This population is in Hardy-Weinberg equilibrium due to its genotype distribution ($P = 0.452$).

REFERENCES

[1] Münzberg H, Morrison CD, 2015. Structure, production and signaling of leptin. *Metabolism: Clinical and Experimental*, Vol. 64, No. 1, pp. 13-23.

[2] Loos R, Yeo G, 2022. The genetics of obesity: from discovery to biology. *Nature Reviews Genetics*, Vol. 23, No. 2, pp. 120-133.

[3] Fairbrother U, Kidd E, Malagamuwa T, Walley A, 2018. Genetics of Severe Obesity. *Current diabetes reports*, Vol. 18, No. 10, pp. 85.

[4] Dasgupta S, Salman M, Siddalingaiah LB, Lakshmi GL, Xaviour D, Sreenath J, 2014. Genetic variants in leptin: Determinants of obesity and leptin levels in South Indian population. *Adipocyte*, Vol. 4, No. 2, pp. 135-140.

[5] Nesrine Z, Haithem H, Imen B, Fadoua N, Asma O, Fadhel NM, Ali B, 2018. Leptin and Leptin receptor polymorphisms, plasma Leptin levels and obesity in Tunisian volunteers. *International Journal of Experimental Pathology*, Vol. 99, No. 3, pp. 121-130.

[6] Boumaiza I, Omezzine A, Rejeb J, Rebhi L, Ouedrani A, Ben Rejeb N, Nabli N, Ben Abdelaziz A, Bouslama A, 2012. Relationship between leptin G2548A and leptin receptor Q223R gene polymorphisms and obesity and metabolic syndrome risk in Tunisian volunteers. *Genetic testing and molecular biomarkers*, Vol. 16, No. 7, pp. 726-733.

[7] Shahid A, Rana S, Mahmood S, Saeed S, 2015. Role of leptin G-2548A polymorphism in age- and gender-specific development of obesity. *Journal of Biosciences*, Vol. 40, No. 3, pp. 521-530.

[8] Sahin DS, Tumer C, Demir C, Celik MM, Celik M, Ucar E, Gunesacar R, 2013. Association with leptin gene C.-2548 G>A polymorphism, serum leptin levels, and body mass index in Turkish obese patients. *Cell Biochem Biophys*, Vol. 65, No. 2, pp. 243-247.

[9] Hoffstedt J, Eriksson P, Mottagui-Tabar S, Arner P, 2002. A polymorphism in the leptin promoter region (-2548 G/A) influences gene expression and adipose tissue secretion of leptin. *Horm Metab Res*, Vol. 34, No. 7, pp. 355-359.

[10] Bains V, Kaur H, Badaruddoza B, 2020. Association analysis of polymorphisms in LEP (rs7799039 and rs2167270) and LEPR (rs1137101) gene towards the development of type 2 diabetes in North Indian Punjabi population. *Gene,* Vol. 754, 144846.

[11] Tang W, Kang M, Liu C, Qiu H, 2019. Leptin rs7799039 (G2548A) polymorphism is associated with cancer risk: a meta-analysis involving 25,799 subjects. *OncoTargets and therapy*, Vol. 12, pp. 2879-2890.

[12] Xiao P, Shi J, Liu X, 2019. Associations of leptin and leptin receptor genetic variants with coronary artery disease: a meta-analysis. *Bioscience reports*, Vol. 39, No. 6, BSR20190466.

[13] He J, Fang Y, Lin X, Zhou H, Zhu S, Zhang Y, Yang H, Ye X, 2016. The Relationship Between Gene Polymorphism of Leptin and Leptin Receptor and Growth Hormone Deficiency. *Med Sci Monit*, Vol. 22, pp. 642-646.

[14] Amer T, El-Baz R, Mokhtar A-R, El-Shaer S, Elshazli R, Settin A, 2017. Genetic polymorphisms of IL-23R (rs7517847) and LEP (rs7799039) among Egyptian patients with hepatocellular carcinoma. *Archives of Physiology and Biochemistry*, Vol. 123, No. 5, pp. 279-85.

[15] https://genome.ucsc.edu/cgi-bin/hgPcr. Date of access February 20, 2022.

[16] http://www.restrictionmapper.org/. Date of access February 22, 2022.

[17] https://international.neb.com/products/r0139-hhai#Product%20Information. Date of access February 22, 2019.

[18] https://asia.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=7:128238230-128239230;v=rs7799039;vdb=variation;vf=731260722. Date of access February 22, 2022.