Open Acess
Jurnal Biota Vol. 10 No. 2 (2024) https://doi.org/10.19109/Biota.v10i2. 21166

Research Article

Comparison of Two PCR Primer Sets for In-House Validation of GHSR Gene Variation Detection Employing Artificial Recombinant Plasmid Approach

Ahsanal Kasasiah[1,3], Jekmal Malau[1,4], Sekar Andjung Tresnawati[2], Priscinya Christiana Debora[1], Nur Komala Fitri[1], Saarah Hamidah Asmara Indratno[1], Asman Hitopik[1], Eriyanti Astika[1], Anisa Aula Rahma[1], Al Mukhlas Fikri[5]*

[1] Department of Pharmacy, Faculty of Health Science, University of Singaperbangsa Karawang, West Java, Indonesia.
[2] The National Population and Family Planning Board for West Java, West Java, Indonesia.
[3] Microbial Biotechnology Laboratory, UPA Laboratorium Terpadu, University of Singaperbangsa Karawang, West Java, Indonesia.
[4] Biology Laboratory, UPA Laboratorium Terpadu, University of Singaperbangsa Karawang, West Java, Indonesia.
[5] Department of Nutrition, Faculty of Health Science, University of Singaperbangsa Karawang, West Java, Indonesia
*Email: mukhlas.fikri@fikes.unsika.ac.id

Received: 22/12/2023	Revised: 26/04/2024	Accepted: 23/05/2024

ABSTRACT

Stunting is a significant global public health problem caused by long-term dietary deficits that affect many children worldwide. Both environmental and genetic factors, including variants in the GHSR gene, play a crucial role in stunted growth. This study used an artificial recombinant plasmid DNA method to evaluate two primer set combinations for identifying DNA variants in the GHSR gene. Selecting suitable primer sets for identifying GHSR genetic variants linked to stunting is essential, as evidenced by PCR and sequencing techniques. The target gene, based on the GHSR reference sequence, consists of eight DNA variations (ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, and A277P). A recombinant plasmid was created by inserting a 1000 bp fragment of the GHSR gene into the pUC57 backbone. Primer sets were chosen based on their capacity to amplify these eight genetic variations and were optimized and validated using PCR methods. PCR and bi-directional sequencing verified the existence of surrounding DNA and specific single nucleotide variants (SNVs). In our study, we discovered four changes in the DNA sequence (R159R G>A, C173R T>C, D246A A>C, and A277P G>C) using the E1_F2/E1_R3 primer pair. Additionally, a new combination of primers (E1_F1/E1_R3) effectively detected seven DNA sequence mutations (ΔQ36 del CAG, G57G C>T, P108L C>T, L118L C>T, R159R G>A, C173R T>C, and D246A A>C). We have developed a new combination of forward and reverse primers to identify seven SNVs in the GHSR gene, which could serve as a diagnostic tool in clinical laboratory settings.

Keywords: *DNA variant; GHSR gene; Primer set; PCR-sequencing; Stunting.*

Introduction

Stunting continues to become a significant global health concern attributable to persistent malnourishment. In 2020, the World Health Organization

(WHO) reported that nearly 149 million of under-five children globally experienced stunting [1]. Moreover, the stunting prevalence rate in Indonesia in 2022 remains high at 21.6%, affecting an estimated 4.7 million children based on the latest national survey [2]. This rate surpasses the WHO threshold of 20%, highlighting the urgency to address this issue.

Stunting is caused by two significant factors, including genetic and environmental factors (socio-economic status, education, and nutrition) [3-7]. Genetic factors have been taken an attention as several previous studies reported the significant correlation with stunting. For instance, research in Brazil revealed that 76.9% of malnourished toddlers were born from short mothers (height <145 cm) [8]. Furthermore, there is a strong genetic correlation due to the incidence of stunting tied to the height of both parents [4]. GHSR, GH1, GHRHR, STAT5B, IGF1, and COMP gene are believed to be linked to stunting [6]. Moreover, various studies have revealed the association of genetic variation factors with stunting incidence in children [6], with particular attention paid to the presence of eight DNA variations in exon 1 of the GHSR gene. These variations comprise ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, and A277P [9].

Biomolecular-based methods are widely recognised as robust method of detecting genetic variation. PCR and direct sequencing techniques have undergone extensive testing for multiple purposes of genetic variation analysis, including the analysis of genetic variation linked to antibiotic resistance [10], the detection of cancer-associated genetic variation [11], and the identification of genetic variation in pathogenic organisms [12]. However, the advancement of PCR for identifying gene variations linked to stunting is currently in a preliminary phase of investigation. In addition, the exploration for primer set pairs capable of amplifying eight genetic variations in GHSR genes ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P is not widely documented while these variations come from Japanese population.

This study aims to identify a set of primer combinations that can be a validated protocol for amplifying GHSR gene variations, namely ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P using Indonesian samples. Prior to direct testing on samples of stunted children, an artificial recombinant pUC57 plasmid containing a synthetic GHSR partial gene was employed as an amplification target. PCR optimization using the engineered synthetic plasmid facilitated the enumeration of target DNA copy numbers, which is pivotal for validating protocols under regulated conditions. We inspected the identification limit of standard plasmid detection and evaluated its performance on several blood samples.

Nevertheless, to our best knowledge it's the very first report comparing several primer set combinations to identify variations in the GHSR gene by using artificially recombinant plasmid DNA. The results emphasized the significance of selecting the appropriate primer set combinations and bi-directional sequencing for the detection of eight GHSR genetic variations, which have been linked to stunting incidence. As a result, this research has the potential enable cost-effective testing of multiple genetic variations.

Materials and methods

Materials

This research utilized nine blood samples from stunted and non-stunted children aged 1-5 years, collected from the villages of Cipicung and Sukatani in the Purwakarta Regency. Ethical approval for this study was provided by the Health Research Ethics Committee (KEPK) of Dr. Cipto Mangunkusumo National Central General Hospital and the Faculty of Medicine, University of Indonesia with

number KET-43/UN2.F1/ETIK/PPM.00.02/2024. Analysis was conducted at the Laboratory of Microbial Biotechnology, UPA Laboratorium Dasar, University of Singaperbangsa Karawang, from September to November 2023. The materials utilized in this study include Genomic DNA Mini Kit (Blood/Cultured Cell) (Geneaid, Taiwan), PCR MyTaq™ master mix (Bioline, USA), recombinant plasmid pUC57 (Genescript Biotech, Nanjing China), TE Buffer (Himedia, India), Nuclease-Free Water (NWF) (Himedia, India), Primer Forward and Primer Reverse (Macrogen, South Korea), Agarose Gel (Vivantis, Malaysia), and GelRed (Biotium, California).

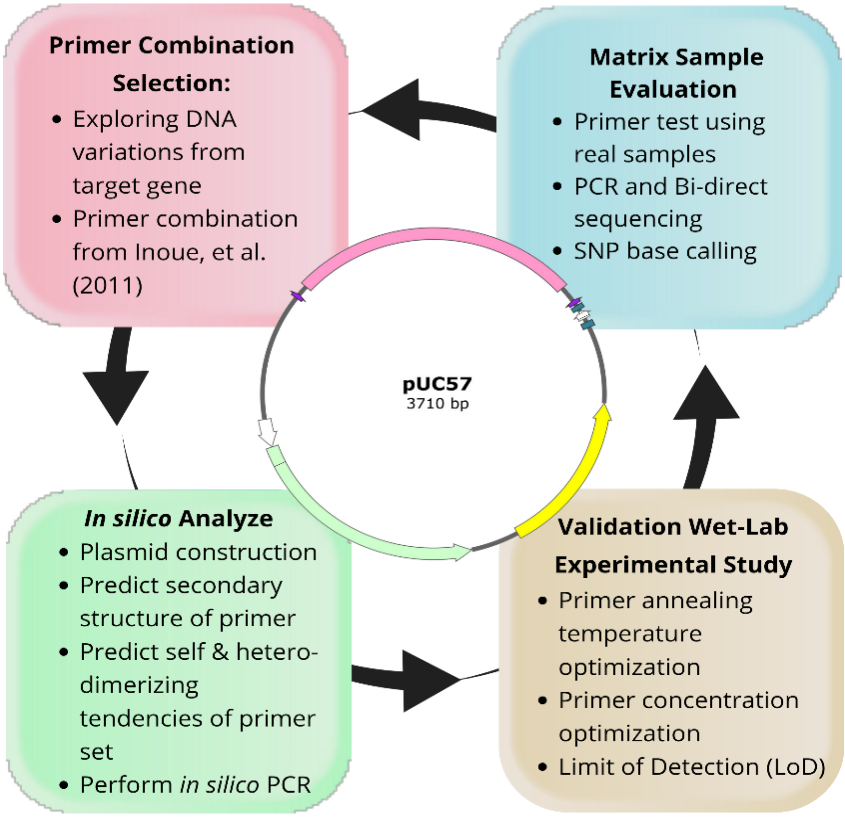

Figure 1. Schematic representation and experimental flowchart illustrating the comparison of dual primer set combinations to identify DNA variations within the GHSR gene.

Methods

1. *Schematic representation in this experimental*

In our research, we proceeded by selecting the reference sequence of the GHSR gene, which contains eight DNA variants (ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, and A277P, respectively). The steps in determining primer combinations to produce the best primer set are detailed in the schematic in Figure 2. This process involves in silico analysis, validation in laboratory experiments, evaluation of samples using diverse study matrices, and analysis of the results of the selected primer combinations. In this study, these steps were carried out comprehensively with the main goal of achieving optimal result quality. These steps were directed at improving the ability of the primer set to identify a large number of SNV variants in the GHSR gene. In silico analysis aimed to evaluate the match of primers with target sequences as well as specificity towards targeted DNA variants. Experimental validation in the laboratory was conducted to confirm the amplification efficiency and detection specificity of the

selected primers. Evaluation of clinical samples aims to assess the performance of primer performance in detecting GHSR gene variation sequences. Analysis of primer combination results was conducted to select the most effective primer combination in identifying SNVs in the gene. With a comprehensive and integrated approach, the steps in this study can make a significant contribution by improving the ability of the primer set to better identify GHSR gene variants [9].

2. *In-silico Construction of Recombinant Plasmids*

Recombinant plasmid construction is performed in silico approach. The GHSR gene sequence was obtained from the National Center for Biotechnology Database (NCBI) with accession number NG_021159 which was then used as a reference sequence. Backbone plasmid pUC57 with a base length of 2,710 bp (base pairs) was obtained from Addgene (www.addgene.org). Recombinant plasmid construction and gene insert annotation were performed using Snapgene Viewer software (www.snapgene.com). The Multiple Cloning Sites (MCS) pUC57 was selected as the partial insertion site for the GHSR gene. The restriction enzyme EcoRV was the cutting site of the pUC57 plasmid for the insertion site of the GHSR target gene (Figure 1). The recombinant plasmid as synthesized from Genescript Biotech (Nanjing, China). This plasmid has been designed as control to determine the optimization combination of primer set and the inserted gene is employed as DNA standard in this internal validation.

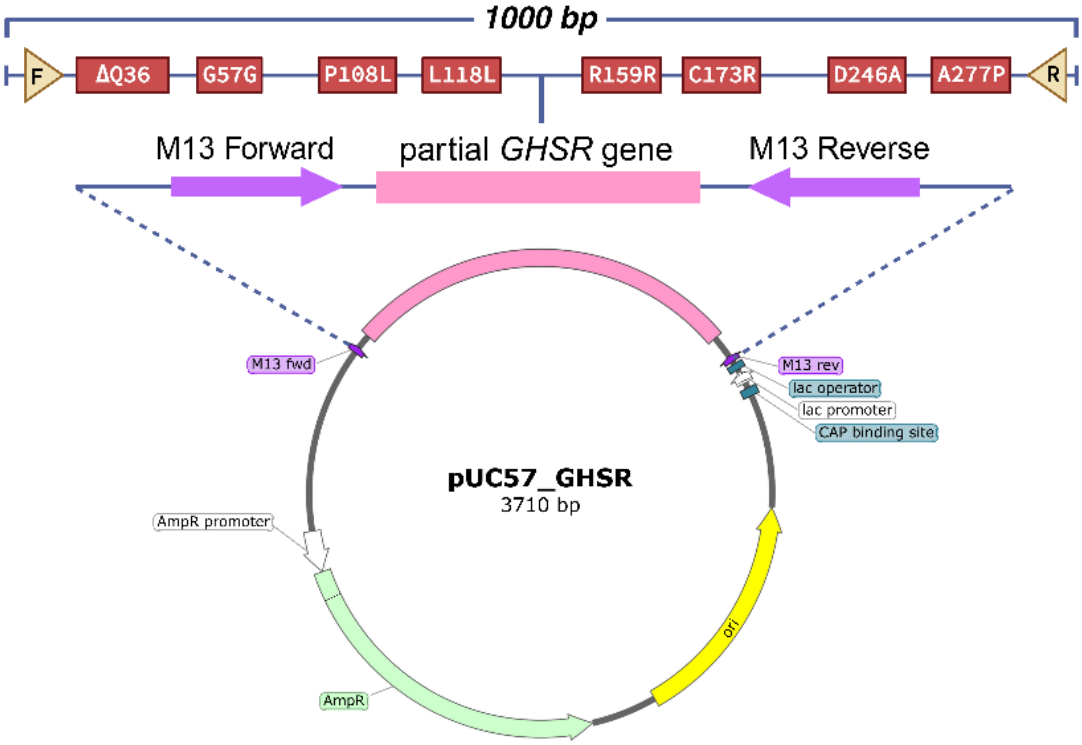

Figure 2. Plasmid mapping and DNA sequence annotation of pUC57 with partial insertion of GHSR gene, including the flanking DNA containing the ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P variations.

3. *Primers analysis and evaluation*

The GHSR gene, containing eight points of genetic variation (ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P), was used as the target gene with gene bank accession number NG_021159

(NCBI). Primers sequences used in this study are shown in Table 1 [9]. Both primer sets were synthesised by Macrogen (Seoul, Korea). Analysis of primer quality was conducted using the NetPrimer online tool (https://www.premierbiosoft.com/netprimer/). Successfully evaluated primers were subsequently mapped using SnapGene viewer software to display primer attachment and polymorphism points/SNVs position.

4. DNA Plasmid Copy number Preparation

The recombinant plasmid pUC57 (Genescript Biotech, Nanjing, China), which inserted the GHSR target gene, was employed as standard DNA in the internal validation. Lyophilized pUC57 plasmid DNA was reconstituted prior to any use. To compute plasmid DNA copy number, a total of 4 µg lyophilized recombinant pUC57 was dissolved in 100 µl of non-flocculent water (NFW) to obtain 4 ng/µl. The tube was centrifuged and placed at room temperature for 5 minutes. A stock concentration of 40 ng/µl of pUC57 was then referred to as the stock plasmid. The tube was vortexed and swirled for 5 seconds each to ensure that the solution was homogenized and collected at the bottom of the tube. A 10x dilution was performed to obtain a working solution 4 ng/µl with total volume 100 µl. Estimated plasmid copy number was calculated based on Avogadro's number and the molecular weight of each nucleotide, (NA × C)/MW, where NA is the Avogadro constant expressed in mol^{-1}, C is the concentration expressed in g/µL, and MW is the molecular weight expressed in g/mol [13].

5. Primer Optimization

Two optimization tests were conducted, specifically for optimizing annealing temperature and primer concentration using pUC57 plasmid. During the optimization test for annealing temperature, eight different temperatures were applied, namely 67°C, 66.1°C, 64.6°C, 63.3°C, 59.6°C, 57.3°C, 55.9°C and 55°C. Conventional PCR was performed using a T100 thermocycler (Bio-Rad, USA). In PCR machine, the gradient program was entered to allow for variation in annealing temperature. PCR composition was conducted by preparing the reagent with a final volume of 20 µL, consisting of 10 µL of MyTaq™ Master mix (Bioline, USA), 1 µL of each of forward and reverse primers (final concentration of 500 nM), 1 µL of pUC57 plasmid (10^5 copy number/reaction) and 7 µL of nuclease-free water per PCR tube. The amplification method started with a predenaturation at 94°C for 5 min. This was followed by 35 cycles of denaturation at 94°C for 30 seconds and annealing for 1 minute using a previously determined gradient temperature. Extension was carried out at 72°C for 2 minutes and final extension was carried out at the same temperature for 5 minutes [14], [15].

While optimizing the primer concentration, four different concentrations were tested: 300 nM, 500 nM, 700 nM and 900 nM. The number of primers in the reagent was adjusted according to the primer concentration used. Each tube had a final volume of 20 µL, which was optimised for the annealing temperature. The PCR amplification protocol remained unchanged from the temperature optimization stage, and the annealing temperature was determined using the optimal temperature achieved during the annealing temperature optimization process [16].

6. PCR procedure

A total of 20 µl of optimized reaction mix containing 10 µl of MyTaq™ ace blend (Bioline, USA), 1 µl of DNA template (105 copy number/reaction), 1 µl of direct/reverse primer (final concentration, 500 nM) and nuclease-free water in the desired final volume. The amplification protocol began with 2 initial denaturation runs at 95°C, followed by 35 cycles of denaturation at 95°C for 15

seconds, annealing at 57°C for 30 seconds and extension at 72°C for 30 seconds and final extension at 72°C for 3 min. A control without template was included in each run and all runs performed were measured in three replicates. PCR amplification was performed using the T100 PCR system (Bio-Rad, USA)

7. *Evaluation Limit of Detection*

The recombinant pUC57 plasmid was used to assess the limit of detection (LOD) in the PCR system. Serial dilutions were performed on the target plasmid, consisting of 10^5, 10^4, 10^3, 10^2, 10^1 and 10^0 ccopies for LoD examination. Serial dilutions were made using TE as a buffer (Himedia, India). Target amplification was determined according to the PCR protocol outlined in the PCR procedure section.

8. *Primer Performance test with artificial recombinant plasmid*

In order to detect genetic variations in the GHSR gene, primer combinations 1 and 2 (as shown in Table 1) were analyzed using artificial recombinant plasmids carrying the target genetic variations. The PCR procedure section was followed for target amplification using a composition mix and PCR protocol. Negative controls were included in each run to ensure accuracy. Furthermore, the presence of PCR products of appropriate size were separated by electrophoresis at 70 volts for 45 minutes (Mupid-exu, Japan) using 1% agarose gel (Vivantis, Malaysia). Band visualization was performed using a UV transilluminator (Accuris-Benchmark Scientific, USA). Single band PCR products were then sequenced using Sanger sequencing capillary electrophoresis (Macrogen, Korea).

Table 1. DNA sequence of oligonucleotides used in this study

No	Primer Name	DNA Sequence Of Oligonucleotides	Amplicon Length	Primer Combination
1.	E1_F2 E1_R3	5'-CAGTGAGAGCTGCACCTACG-3' 5'-TAGCGACTCAGGGGGAAATA-3'	536 bp	F/R; Existing [9]
2.	E1_F1 E1_R3	5'-GTCCCAGAGCCTGTTCAGC-3' 5'-TAGCGACTCAGGGGGAAATA-3'	932 bp	F/R; New combination [9]

9. *Primer Performance test with multiple sample*

F/R primer combinations and predefined protocols were designed to detect eight genetic variations of the GHSR gene (ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P) using various sample types. The effectiveness of the primers was assessed in routine inspections by analyzing blood samples obtained from stunted and non-stunted children. Sampling procedures followed established ethical guidelines and used standardized protocols performed by medical professionals. Genomic DNA extraction was performed using the Genomic DNA Mini Kit (Blood/Cultured Cell) (Geneaid, Taiwan) following the manufacturer's procedures. DNA concentration was measured using a Nanodrop Microplate Reader Epoch 2 (BioTek, USA). Target amplification was performed following the mixture composition and PCR protocol specified in the PCR procedure section. A negative control was included in each run. PCR products of appropriate size were separated by electrophoresis at 70 volts 45 min (Mupid-exu, Japan) using 1% agarose gel (Vivantis, Malaysia). Band visualization was performed using a UV transilluminator (Accuris-Benchmark Scientific, USA). Single band PCR products were then proceeded for sequencing process using sanger sequencing capillary electrophoresis (Macrogen, Korea).

10. Data Analysis

Sequencing data was analyzed using specific software. Electropherograms were analyzed using Seq Scanner 2 software (Applied Biosystems, USA). Quality base calling analysis was performed using Snapgene viewer software. Trimming and consensus generation were performed using Bioedit software. Flanking DNA analysis and SNP calling obtained were evaluated using Snapgene viewer software.

Results and Discussion

In-silico Construction of Recombinant Plasmid

Recombinant plasmid construction was performed using in-silico approach. The backbone plasmid pUC57 has a total length of 2,710 bp [17]. A 1000 bp long gene insert containing wild-type DNA sequences ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P was successfully produced and inserted through the EcoRV cutting site located between the M13 forward and reverse primers. As a result, a recombinant plasmid pUC57+insert MT gene was obtained with a length of 3,710 bp (Figure 3). The use of SnapGene Viewer software made it possible to simulate the gene insertion at the EcoRV site and confirm the compatibility of the target gene sequence ligated into the pUC57 plasmid [13], [18]. The results obtained were then used as a starting point for the primer optimization and validation process at a later stage.

Primers analysis and evaluation

The primer set combinations were selected based on the ability to amplify eight genetic variants revealed in the GHSR gene. The primer sets were selected based on the DNA sequences provided by Inoue et al. (2011). In this study, the comparison of the two primer sets tested consisted of existing primers and new primer combinations with the names E1_F2/E1_R3 and E1_F1/E1_R3, respectively. The E1_F2/E1_R3 primer targets four genetic variants in the GHSR gene namely R159R, A277P, D246A, and C173R (Figure 4A). These primers generated a 536-bp-long fragment located between positions 5409–5428 and 5925–5944. On the other hand, the E1_F1/E1_R3 primer set was designed to detect eight genetic variants, including ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, and A277P (Figure 4B). These primers amplified a 932-bp-long fragment of the GHSR gene, specifically within the binding region between positions 5013–5031 and 5925–5944. Results of mapping analysis of two primers are presented in Figure 4. The result of the evaluation of primer characteristics using Netprimer and Snapgene software on %GC, Tm value, amplicon length, and number of SNVs that can be detected are shown in Table 2.

Table 2. Evaluation of two primers set

Primer	Sequence (5'-3')	Sequence variant detected	Tm (°C)	GC (%)	Amplicon size (bp)
E1_F2	CAGTGAGAGCTGCACCTACG	R159R, A277P, D246A, C173R	56	60	536
E1_R3	TAGCGACTCAGGGGGAAATA		57	50	
E1_F1	GTCCCAGAGCCTGTTCAGC	ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P	57	63	932
E1_R3	TAGCGACTCAGGGGGAAATA		57	50	

Primer Optimization

The effectiveness of PCR is greatly dependent on the precise binding of a primer to its intended target sequence while avoiding any binding to non-target sequences. Therefore, it is crucial to improve the molecular interaction in order to reach this specificity [19]. The annealing

temperature (Ta) and concentration are key factors to optimize primer [20] [21]. Temperatures for primer annealing may vary between 55°C and 72°C and are dependent upon both the base composition and the sequence length. The guanine-cytosine (G-C) base pair is characterized by the existence of three hydrogen bonds, while the adenine-tin (A-T) base pair has only two hydrogen bonds. Therefore, DNA sequences with a larger proportion of G-C content would exhibit elevated melting temperatures and annealing temperatures. A temperature difference of 1-2 °C might lead to non-specific amplification [20] [22].

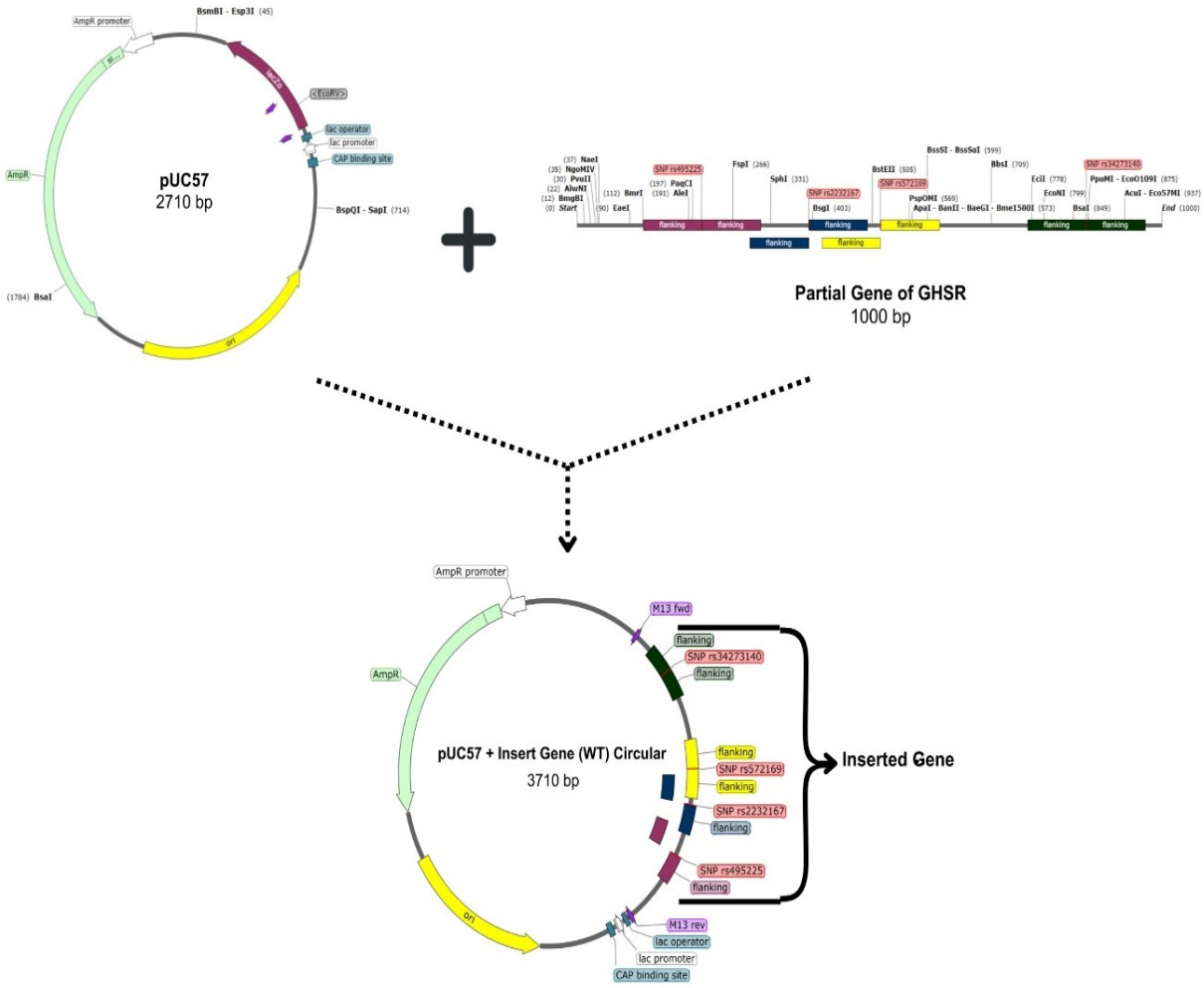

Figure 3. Mapping the construction of recombinant plasmid containing wild-type DNA sequences of ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, A277P using partial GHSR gene.

For the optimization of Ta, control samples were prepared using recombinant pUC57 as standard DNA, which contained specific single nucleotide variants (SNVs) including ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, and A277P. The concentration of the standard DNA was 9.8 × 10^5 copies/µL. The annealing temperature (Ta) investigations were performed using a gradient PCR machine. Ta values ranging from 55 to 67°C were tested during the optimization process. The optimization results indicate that the optimal annealing temperature (Ta) for the existing primers and the new primer combinations (E1_F2/E1_R3 and E1_F1/E1_R3) was found to be 57°C despite the positive controls that were tested showed no

significant variation in the thickness of the bands at annealing temperatures of 59.6°C, 57.3°C, 55.9°C, and 55.0°C, indicating that both sets of primers were robust (Figure 5).

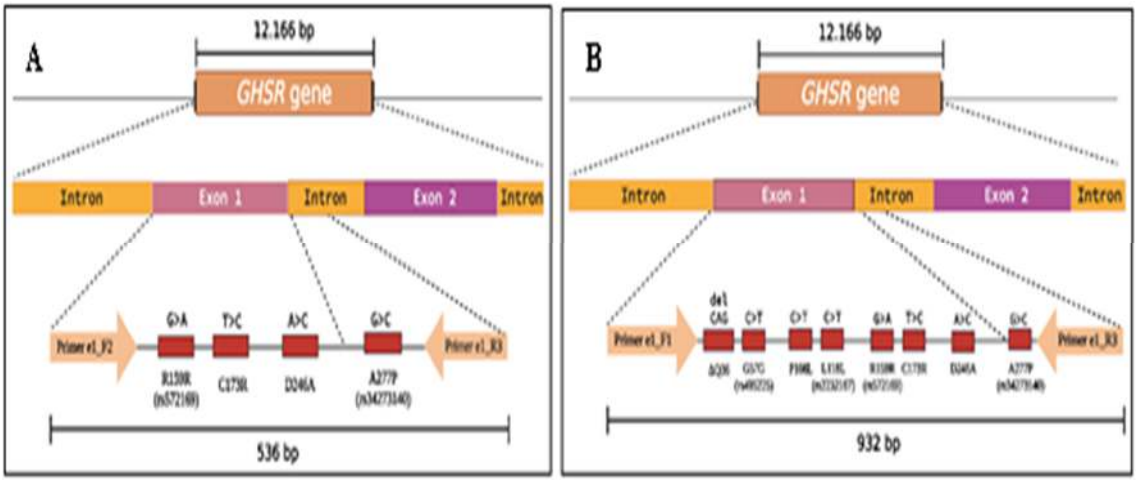

Figure 4. Position mapping of forward/backward primer attachment to partial GHSR gene. (A) Existing primer which produces a 536 bp amplification size; (B) New primer combination which produces a 932 bp amplification size

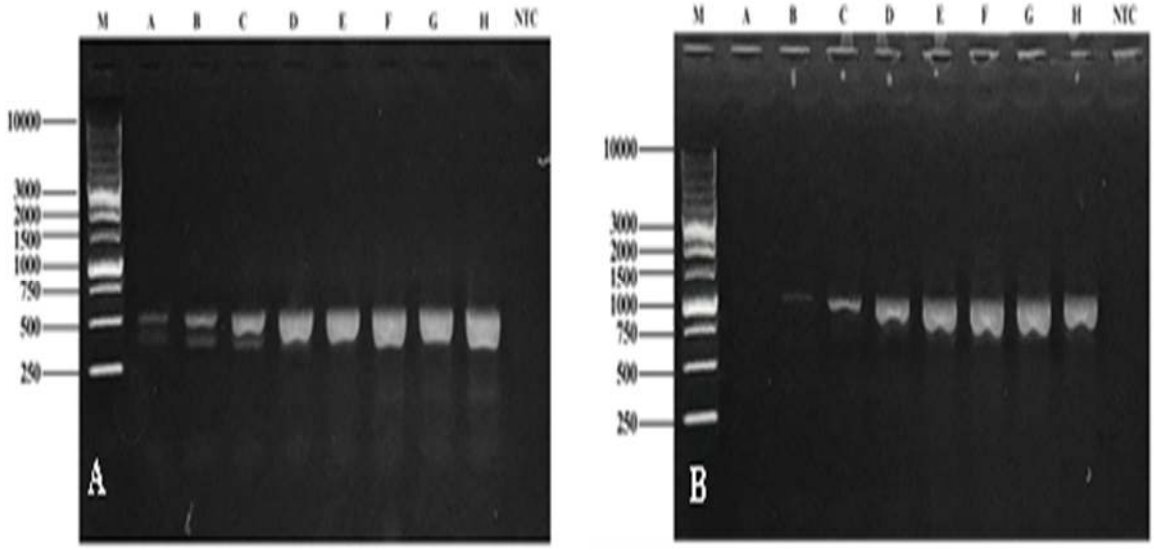

Figure 5. Optimization of two different primer sets in GHSR-PCR. M= DNA marker, A 67,0°C, B 66,1°C, C 64,6°C, D 62,3°C, E 59,6°C, F 57,3°C, G 55,9°C, H 55,0°C, NTC= negative control. (A) PCR amplification of E1_F2/E1_R3, amplicon length = ± 536 bp (B) PCR amplification of E1_F1/E1_R3 amplicon length = ± 932 bp.

Subsequently, further optimization stages were conducted to determine the optimal quantity of primer required to generate a sufficient amount of starter for DNA template amplification. The concentration range tested was 300 to 900 nM. The optimal primer concentration, which resulted in the highest quality DNA band with minimal or no primer dimers, was determined (Figure 6). Primer dimers occur due to self-annealing or annealing of primers to each other [23]. The result showed consistent band thickness through annealing temperatures of 500 nM, and 700

nM, suggesting that both primers were highly reliable [24]. For the E1_F2/E1_R3 primer pairs, a primer concentration of 500 nM demonstrated sufficient results in producing a solid band. On the other hand, for the E1_F1/E1_R3 primer pairs, the optimal concentration was found to be 700 nM, resulting in reduced formation of primer dimers.

Primer Validation

When considering the suitability for detecting GHSR genetic variation, determining should be based on PCR SNP genotyping. Thus, it is important to develop a set of primer combinations that can consistently detect the target DNA as low as possible. Plasmids were counted based on Avogadro's method and 10-fold serial dilutions were performed. Therefore, it is important to check the accuracy of the dilution, as the quality of the sample dilution is critical in determining the results. Based on the pattern of decreasing bandwidth of each lane. The detection limit for the specified conditions for the E1_F2/E1_R3 primer pair is 10 copies per reaction. In contrast, the E1_F1/E1_R3 primer pair showed a limit of detection of 100 copies per reaction. Data are presented in Table 3.

Table 3. The evaluation of limit of detection E1_F2/E1_R3 and E1_F1/E1_R3 primer sets using pUC57 recombinant plasmid.

Assigned copy numbers per reaction	Existing Primer E1_F2/E1_R3	New Primer Combination E1_F1/E1_R3
9.8×10^5	+	+
9.8×10^4	+	+
9.8×10^3	+	+
9.8×10^2	+	+
9.8×10^1	+	-
9.8×10^0	-	-
H_2O	-	-

Note : (+) Detected, (-) Undetected

SNP Genotyping Verification using DNA standard (recombinant pUC57)

In this study, the two primer pairs showed different performances. Although there is an alteration in the forward primer, all other components remain constant. Several factors associated with primer usage in PCR can affect the limit of detection. Primers that exhibit high specificity can bind to target DNA even at low concentrations by preventing cross-reactivity with non-specific sequences. PCR conditions may also influence binding efficiency to target sequences, particularly the optimal annealing temperature for primers [25], [26].

Prior to evaluating clinical samples, control samples were performed using pUC57 106 copy/µL as the standard DNA, which contained a reference DNA (wild-type) of single nucleotide variants (SNVs), including ΔQ36, G57G, P108L, L118L, R159R, C173R, D246A, and A277P. Analysis of base calling was conducted using conventional Polymerase Chain Reaction (PCR) and bi-directional sequencing. The result showed that existing primers E1_F2/E1_R3 produce an amplicon fragment of approximately 536 bp and successfully identify four DNA wild-type sequence variants, specifically R159R, C173R, D246A, and A277P. In contrast, the new combination E1_F1/E1_R3 generated an amplicon fragment of approximately 932 bp that was able to identify seven DNA wild-type of sequence variants, namely ΔQ36, G57G, P108L, L118L, R159R, C173R, and

D246A (Table 4). The results of PCR and electropherogram of existing E1_F2/E1_R3 and new combination E1_F1/E1_R3 primers are shown in Supplementary 1 and 2, respectively. Based on the sequence data, the new combination primers as a new set of alternative primers showed strong performance with perfect base calling matches.

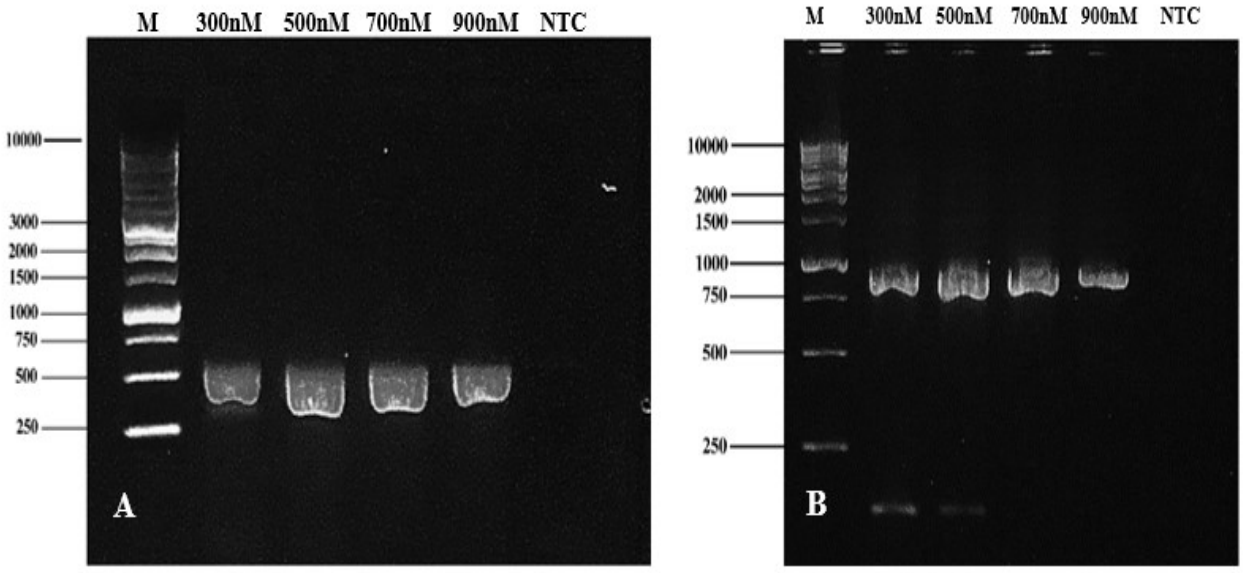

Figure 6. Primer concentration optimization for GHSR genes. Each lane on the agarose gel represents different primer concentrations. M = DNA marker 1 kb; NTC = negative control. (A) PCR amplification of E1_F2/E1_R3, amplicon length = ± 536 bp (B) PCR amplification of E1_F1/E1_R3 amplicon length = ± 932 bp

Assessment of Primer PCR using Clinical Samples

An evaluation of the suitability of primer sets was conducted by employing PCR and bi-directional sequencing on nine samples with an average DNA concentration of 35-45 ng/ul (Supp 3). All extracted genomic DNA samples showed high quality, with no smear banding pattern seen (Supp 4). Subsequent to applying of both existing primers and novel primer combinations, successful amplification was obtained for all samples, which produced amplicon lengths 536 bp and 932 bp for E1_F2/E1_R3 and E1_F1/E1_R3 primer sets (Figure 7).

Further performance testing was conducted for this primer using SNP genotyping by Sanger sequencing. All tested samples were confirmed positive and contained a flanking DNA of sequence variant (wild-type or mutant). Existing primers E1_F2/E1_R3 successfully identified four sequence variants, specifically R159R, C173R, D246A, and A277P. In contrast, the new combination E1_F1/E1_R3 was able to identify seven DNA wild-type sequence variants, namely ΔQ36, G57G, P108L, L118L, R159R, C173R, and D246A (Table 5). Repeated testing of individual samples indicated 100% identity of variant calls, as evidenced by electroferograms in the same region of the DNA flanking region. In addition, the genotyping results of the standard recombinant DNA plasmid pUC57 and human clinical samples were in excellent agreement for the wild-type and mutant genotypes of the tested sample pairs (Table 5). The data demonstrate the successful combination of the newly developing primers to detect seven sequence variants in the GHSR gene.

Table 4. The evaluation variant calls E1_F2/E1_R3 and E1_F1/E1_R3 primer sets using pUC57 recombinant plasmid

	8 GHSR Variant Sequence							
	ΔQ36	**G57G**	**P108L**	**L118L**	**R159R**	**C173R**	**D246A**	**A277P**
New Combination Primers E1_F1/E1_R3	+ CAG/CAG	+ C/C	+ C/C	+ C/C	+ G/G	+ T/T	+ A/A	-
Existing Primers E1_F2/E1_R3	-	-	-	-	+ G/G	+ T/T	+ A/A	+ G/G

Note : (+) Detected (Wildtype), (-) Undetected

Table 5. Representative existing and new combination primers evaluation for SNVs identification using 4 clinical samples

Sample	**Sequence Variant**											
	E1_F2/E1_R3				**E1_F1/E1_R3**							
	R159R	**C173R**	**D246A**	**A277P**	**ΔQ36**	**G57G**	**L118L**	**P108L**	**R159R**	**C173R**	**D246A**	**A277P**
A1	+ G/G	+ T/T	+ A/A	+	+ CAG/CAG	+ C/C	+ C/C	+ C/C	+ G/G	+ T/T	+ A/A	-
A2	+ G/G	+ T/T	+ A/A	+	+ CAG/CAG	+ C/C	+ C/C	+ C/C	+ G/G	+ T/T	+ A/A	-
A3	+ G/G	+ T/T	+ A/A	+	+ CAG/CAG	+ C/C	+ C/C	+ C/C	+ G/G	+ T/T	+ A/A	-
A4	+ G/G	+ T/T	+ A/A	+	+ CAG/CAG	+ C/C	+ C/C	+ C/C	+ G/G	+ T/T	+ A/A	-

Note : (+) Detected (Wildtype), (-) Undetected

SNPs, also known as single nucleotide polymorphisms, are genetic variations that occur in more than 1% of the population [27]. These variants are applied in several disciplines such as pharmacogenomics and as biological markers for genetic diseases. Recent research has highlighted the significance and efficiency of identifying SNPs. Among the various methods, PCR-based detection stands out as widely used and primer design is a critical step in PCR studies. This study aims to compare the primer set combinations used in the previous study conducted by Inoue et al. (2011) with our newly developed combinations to optimize the method for identifying GHSR gene mutations associated with stunting in children. In the previous study, the first set of primers, E1_F2/E1_R3, was used to target four genetic variants in the GHSR gene. However, the E1_F1/E1_R3 primer set, which was a novel combination in this study, was designed to find eight genetic variants. Nevertheless, the new primer sets combination were only capable of amplifying seven SNPs, indicating the necessity for further enhancements. In order to amplify eight SNPs, it is necessary to identify the right combination of primers.

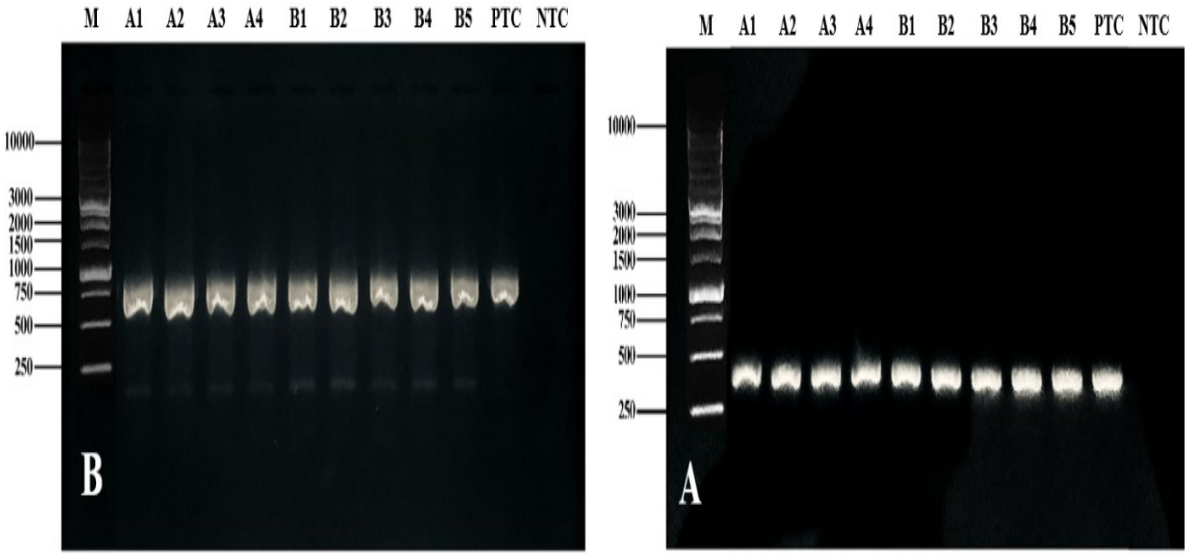

Figure 7. (A) PCR amplification of GHSR partial gene using E1_F2 forward and E1_R3 reverse on human genomic DNA. Amplicon length = ± 536 bp. M = DNA marker 1 kb; A1-B5 = genomic DNA carrying sequence variant; PTC = positive control NTC = negative control. (B) PCR amplification of GHSR partial gene using E1_F1 forward and E1_R3 reverse on human genomic DNA. Amplicon length = ± 932 bp. M = DNA marker 1 kb; A1-B5 = genomic DNA carrying sequence variant; PTC= positive control NTC = negative control.

Conclusions

In this research, we conducted a comparative analysis of established and novel forward/reverse primer combinations for detecting and identifying sequence variants within the GHSR gene. Our study successfully identified four wild-type DNA sequence variants—R159R, C173R, D246A, and A277P—utilizing E1_F2/E1_R3 as the established primer set and seven wild-type DNA sequence variants—Q36, G57G, P108L, L118L, R159R, C173R, and D246A—using E1_F1/E1_R3 as a novel primer combination. Both primer sets were successfully optimized and validated employing PCR technique. In addition, the existing and new primers were evaluated for SNVs genotyping using artificial recombinant plasmids and nine clinical

samples, which showed a perfect SNVs calling concordance rate of 100%. This investigation introduces an innovative forward and reverse primer combination for the detection and identification of seven SNVs in the GHSR gene, positioning it as a potential tool for clinical laboratory diagnostics.

Acknowledgments

The funding for this research was provided by the National Competitive Research Program under the Indonesian Ministry of Education, Culture, Research, and Technology. The award number for this funding is 254/SP2H/UN64.10.PP/2023. We would like to extend our appreciation to Universitas Singaperbangsa Karawang for their accessibility to research facilities and infrastructure.

References

[1] K. F. Michaelsen, L. M. Neufeld, and A. M. Prentice, *Global landscape of nutrition challenges in infants and children*. Karger Medical and Scientific Publishers, 2020, doi: 10.1159/000503315.

[2] Kementerian Kesehatan RI, "BUKU SAKU hasil survei status gizi indonesia (SSGI)," 2022. [Online]. Available: https://kesmas.kemkes.go.id/assets/uploads/contents/attachments/09fb5b8ccfdf088080f2521ff0b4374f.pdf

[3] P. Vonaesch *et al.*, "Factors associated with stunting in healthy children aged 5 years and less living in Bangui (RCA)," *PLoS One*, vol. 12, no. 8, p. e0182363, 2017, doi: 10.1371/journal.pone.0182363.

[4] R. R. El Kishawi, K. L. Soo, Y. A. Abed, and W. A. M. W. Muda, "Prevalence and associated factors influencing stunting in children aged 2-5 years in the Gaza Strip-Palestine: A cross-sectional study," *BMC Pediatr*, vol. 17, no. 1, p. 210, 2017, doi: 10.1186/s12887-017-0957-y.

[5] S. Fatima, I. Manzoor, A. M. Joya, S. Arif, and S. Qayyum, "Stunting and associated factors in children of less than five years: A hospital-based study," *Pak J Med Sci*, vol. 36, no. 3, 2020, doi: 10.12669/pjms.36.3.1370.

[6] W. R. W. Taib and I. Ismail, "Evidence of stunting genes in Asian countries: A review," *Meta Gene*, vol. 30, p. 100970, 2021, doi: 10.1016/j.mgene.2021.100970.

[7] A. D. Laksono, N. E. W. Sukoco, T. Rachmawati, and R. D. Wulandari, "Factors related to stunting incidence in toddlers with working mothers in Indonesia," *Int J Environ Res Public Health*, vol. 19, no. 17, p. 10654, 2022, doi: 10.3390/ijerph191710654.

[8] J. D. Y. Orellana *et al.*, "Intergenerational association of short maternal stature with stunting in Yanomami indigenous children from the Brazilian Amazon," *Int J Environ Res Public Health*, vol. 18, no. 17, p. 9130, 2021, doi: 10.3390/ijerph18179130.

[9] H. Inoue *et al.*, "Identification and functional analysis of novel human growth hormone secretagogue receptor (GHSR) gene mutations in Japanese subjects with short stature," *J Clin Endocrinol Metab*, vol. 96, no. 2, pp. E373–E378, 2011, doi: 10.1210/jc.2010-1570.

[10] C. Katsukawa, T. Kenri, K. Shibayama, and K. Takahashi, "Genetic characterization of Mycoplasma pneumoniae isolated in Osaka between 2011 and 2017: Decreased detection rate of macrolide-resistance and increase of p1 gene type 2 lineage strains," *PLoS One*, vol. 14, no. 1, p. e0209938, 2019, doi: 10.1371/journal.pone.0209938.

[11] P. Bohanes *et al.*, "Pharmacogenetic analysis of INT 0144 trial: Association of polymorphisms with survival and toxicity in rectal cancer patients treated with 5-FU and radiation," *Clinical Cancer Research*,

vol. 21, no. 7, pp. 1583–1590, 2015, doi: 10.1158/1078-0432.ccr-14-0857.

[12] K. S. Ku, R. K. Chodavarapu, R. Martin, M. D. Miller, H. Mo, and E. S. Svarovskaia, "Sequencing analysis of NS3/4A, NS5A, and NS5B genes from patients infected with hepatitis c virus genotypes 5 and 6," *J Clin Microbiol*, vol. 54, no. 7, pp. 1835–1841, 2016, doi: 10.1128/jcm.00238-16.

[13] U. Nuraeni *et al.*, "Droplet digital PCR versus real-time PCR for in-house validation of porcine detection and quantification protocol: An artificial recombinant plasmid approach," *PLoS One*, vol. 18, no. 7, pp. e0287712–e0287712, 2023, doi: 10.1371/journal.pone.0287712.

[14] P. E. P. Ariati, I. G. P. Wirawan, and M. M. V Sasadara, "Optimization of primer and polymerase chain reaction conditions to amplify COI locus for identification of purnajiwa (euchresta horsfieldii (lesch.) benn.) collected from bedugul, bali," *IOP Conf Ser Earth Environ Sci*, vol. 913, no. 1, p. 12069, 2021, doi: 10.1088/1755-1315/913/1/012069.

[15] D. Aisya, M. E. Poerwanto, D. Wicaksono, and J. D. Ortuoste, "Optimization of the detection method for xanthomonas axonopodis on chili pepper seeds using polymerase chain reaction with two difference primers," *BIO Web Conf*, vol. 69, p. 1033, 2023, doi: 10.1051/bioconf/20236901033.

[16] M. Park, J. Won, B. Y. Choi, and C. J. Lee, "Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR," *Exp Mol Med*, vol. 52, no. 6, pp. 963–977, 2020, doi: 10.1038/s12276-020-0452-7.

[17] Y. Goulev, A. Matifas, V. Heyer, B. Reina-San-Martin, and G. Charvin, "COSPLAY: An expandable toolbox for combinatorial and swift generation of expression plasmids in yeast," *PLoS One*, vol. 14, no. 8, p. e0220 10.1371/journal.pone.0220694.

[18] M. I. Fahri, R. M. Alatiffa, S. I. Yanti, I. Prakoso, and A. N. Mashitah, "In silico recombinant plasmid design of pHA171 with phdABCD insertion for ethidium bromide degradation," *Acta Biochimica Indonesiana*, vol. 4, no. 1, p. 7, 2021, doi: 10.32889/actabioina.7.

[19] M. McPherson and S. Møller, *PCR*, 2nd ed. Taylor & Francis, 2006.

[20] J. Obradovic *et al.*, "Optimization of PCR conditions for amplification of GC-RichEGFRPromoter sequence," *J Clin Lab Anal*, vol. 27, no. 6, pp. 487–493, 2013, doi: 10.1002/jcla.21632.

[21] R. Setyawati and S. Zubaidah, "Optimasi konsentrasi primer dan suhu annealing dalam mendeteksi gen leptin pada sapi peranakan ongole (PO) menggunakan polymerase chain reaction (PCR)," *Indonesian Journal of Laboratory*, vol. 4, no. 1, p. 36, 2021, doi: 10.22146/ijl.v4i1.65550.

[22] Y. M. D. Lo and K. C. A. Chan, "Introduction to the polymerase chain reaction," *Methods Mol Biol*, vol. 336, pp. 1–10, 2006, doi: 10.1385/1-59745-074-X:1.

[23] T. C. Lorenz, "Polymerase chain reaction: Basic protocol plus troubleshooting and optimization strategies," *Journal of Visualized Experiments*, vol. 63, no. 63, 2012, doi: 10.3791/3998.

[24] M. Masnaini, A. Achyar, M. Chatri, D. H. Putri, Y. Ahda, and Irdawati, "Primer design and optimization of pcr methods for detecting mixed rat meat in food samples," *Adv Biol Sci Res*, pp. 282–289, 2023, doi: 10.2991/978-94-6463-166-1_37.

[25] U. K. Hanapi, M. N. M. Desa, A. Ismail, and S. Mustafa, "A higher sensitivity and efficiency of common primer multiplex PCR assay in identification of meat origin using

NADH dehydrogenase subunit 4 gene," *J Food Sci Technol*, vol. 52, no. 7, pp. 4166–4175, 2014, doi: 10.1007/s13197-014-1459-7.

[26] D. Silalahi, I. G. P. Wirawan, and M. M. V Sasadara, "Optimization of annealing temperature for amplification of ehoscnola locus in pranajiwa (euchresta horsfieldii) plant collected from mountains, urban and coastal areas in Bali," *IOP Conf Ser Earth Environ Sci*, vol. 913, no. 1, p. 012059, 2021, doi: 10.1088/1755-1315/913/1/012059.

[27] K. S. Allemailem *et al.*, "Single nucleotide polymorphisms (SNPs) in prostate cancer: Its implications in diagnostics and therapeutics," *Am J Transl Res*, vol. 13, no. 4, pp. 3868–3889, 2021, [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8129253/.