bio-protocol

Open Access

Published: Feb 05, 2025

Protocol to Identify Unknown Flanking DNA Using Partially Overlapping Primer-based PCR for Genome Walking

Mengya Jia[1, 2], Dongqin Ding[1, 3, $], Xiaohua Liu[1, 3] and Haixing Li[1, 3, *]

[1]State Key Laboratory of Food Science and Resource, Nanchang University, Nanchang, China
[2]School of Life Sciences, Lanzhou University, Lanzhou, China
[3]Sino-German Joint Research Institute, Nanchang University, Nanchang, China
[$]Present address: Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China
*For correspondence: hxli@ncu.edu.cn

Abstract

Genome walking is a popular molecular technique for accessing unknown flanking DNAs, which has been widely used in biology-related fields. Herein, a simple but accurate genome-walking protocol named partially overlapping primer (POP)-based PCR (POP-PCR) is described. This protocol exploits a POP set of three POPs to mediate genome walking. The three POPs have a 10 nt 3′ overlap and 15 nt heterologous 5′ regions. Therefore, a POP can partially anneal to the previous POP site only at a relatively low temperature (approximately 50 °C). In primary POP-PCR, the low-temperature (25 °C) cycle allows the primary POP to partially anneal to site(s) of an unknown flank and many sites of the genome, synthesizing many single-stranded DNAs. In the subsequent high-temperature (65 °C) cycle, the target single-stranded DNA is converted into double-stranded DNA by the sequence-specific primer, attributed to the presence of this primer complement, while non-target single-stranded DNA cannot become double-stranded because it lacks a binding site for both primers. As a result, only the target DNA is amplified in the remaining 65 °C cycles. In secondary or tertiary POP-PCR, the 50 °C cycle directs the POP to the previous POP site and synthesizes many single-stranded DNAs. However, as in the primary PCR, only the target DNA can be amplified in the subsequent 65 °C cycles. This POP-PCR protocol has many potential applications, such as screening microbes, identifying transgenic sites, or mining new genetic resources.

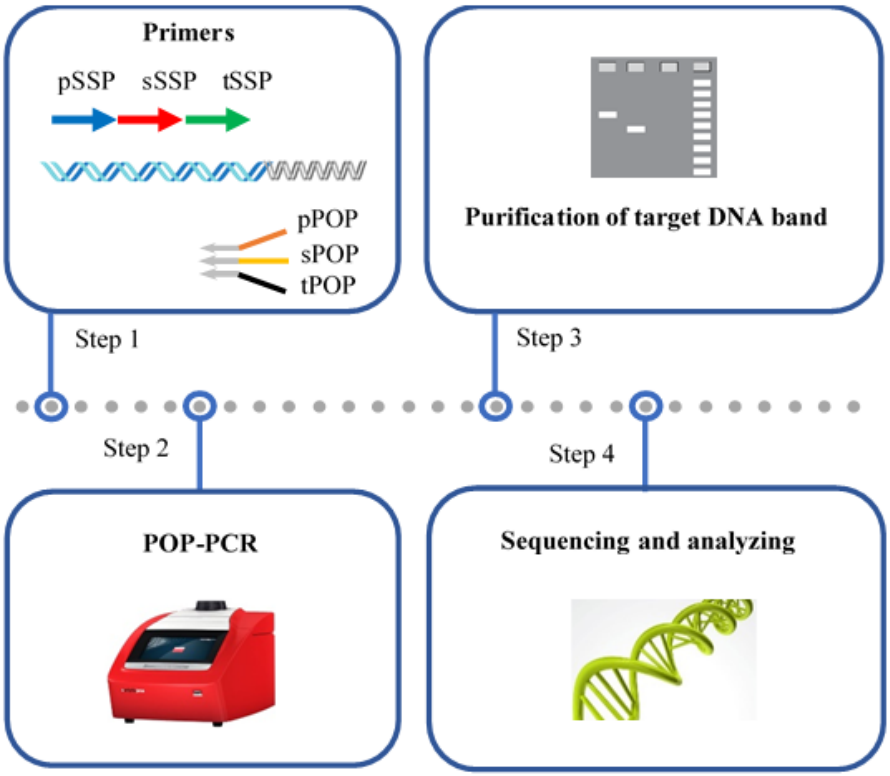
Key features

- This POP-PCR protocol, built upon the technique developed by Li et al. [1], is universal to genome walking of any species.
- The established protocol relies on the 10 nt 3′ overlap among a set of three POPs.
- The first two rounds of POP-PCRs can generally give a positive walking outcome.

Keywords: Genome walking, Partially overlapping primer, Sequence-specific primer, Partial annealing, Primary POP-PCR, Secondary POP-PCR, Tertiary POP-PCR, DNA sequencing

This protocol is used in: PLoS One (2015), DOI: 10.1371/journal.pone.0120139

Graphical overview



Background

In life science–related areas, genome walking is often employed to obtain unknown genomic DNA flanking known DNA [2–5]. Genome walking has significantly advanced areas such as genetics, biotechnology, and microbiology [6–9]. Many genome-walking methods have been proposed during the past four decades. Although the rationales or experimental steps involved are diverse, these methods can generally be classified into two groups: (i) genome treatment-dependent PCR and (ii) random PCR. The former requires digesting genomic DNA and then ligating the digested DNA with a cassette/adaptor/linker prior to PCR amplification. A random PCR, however, requires just a partial annealing of the walking primer to the unknown flank in the low-temperature (around 30 °C) cycle. A random PCR-based walking protocol has garnered much more attention than other protocols because it omits the time-consuming genome treatment process [10–14]. To date, approximately a dozen techniques of random PCR-based genome walking have been established, such as panhandle (or racket) PCR, thermal asymmetric interlaced PCR, and primer extension refractory PCR [6,12,15–18]. These techniques rely on the random annealing of the arbitrary walking primer to the unknown flanking genomic region, and the subsequent differential amplification between target DNA and non-target DNA. The reason for this differential amplification is that the annealing efficiency of the walking primer is lower than that of the paired sequence-specific primer (SSP). However, these techniques are still comprised of complex PCR steps or unsatisfactory amplification specificity [19–21]. Therefore, researchers have been actively attempting to devise a practical random PCR–based genome-walking technique.

Here, a novel genome-walking protocol of partially overlapping primer-based PCR (POP-PCR) is detailed for efficient genome walking. A POP-PCR set, comprising three rounds of nested amplifications, relies on the 10 nt 3′ overlap among the three POPs to overcome non-target amplification. In each round of amplification, the POP has only one chance of partial annealing, as only one 25 °C or 50 °C cycle is performed in this amplification. Thus, only the target product can be amplified, while the non-target product is removed. In addition, unlike other walking methods, POP-PCR uses a unique POP for each round of amplification, which helps reduce non-target background arising from the other POPs. The proposed protocol can be used in many aspects of life science–related areas, like isolating promoters, unveiling new genetic resources, or identifying T-DNA [1,22–24].

Materials and reagents

Biological materials

1. Genomic DNA of rice, given by the Lab of Dr. Xiaojue Peng at Nanchang University
2. Genomic DNA of *Pichia pastoris* GS115, extracted using Dr. GenTLE (from Yeast) High Recovery kit (Takara, Dalian, China) according to the supplier's instruction
3. Genomic DNA of *Levilactobacillus brevis* NCL912 [25–27], extracted using the method described by Li et al. [26]
4. Genomic DNA of human blood, extracted using the method described by Gustafson et al. [28]

Reagents

1. 10× LA PCR buffer (Mg^{2+} plus) (Takara, catalog number: RR042A)
2. 6× Loading buffer (Takara, catalog number: 9156)
3. LA Taq polymerase (hot-start version) (Takara, catalog number: RR042A)
4. dNTP mixture (Takara, catalog number: RR042A)
5. DL 5,000 DNA marker (Takara, catalog number: 3428Q)
6. 1× TE buffer (Sangon, catalog number: B548106)
7. Agarose (Sangon, catalog number: A620014)
8. 1 M NaOH (Yuanye, catalog number: B28412)
9. 0.5 M EDTA (Solarbio, catalog number: B540625)
10. Green fluorescent nucleic acid dye (10,000×) (Solarbio, catalog number: G8140)
11. Tris (Solarbio, catalog number: T8060)
12. Boric acid (Solarbio, catalog number: B8110)
13. Agarose Gel DNA Purification kit V2.0 (Takara, catalog number: DV805A)
14. Primers (Sangon)

pPOP1: AGTCAGCGTCCAGGT*AGTCAGTCTC*
sPOP1: TCAGGTCCAAGGTCA*AGTCAGTCTC*
tPOP1: CTCAGCGTGTTCGTC*AGTCAGTCTC*
pPOP2: CAGTCAGTCTCAGGT*CGTCTCCAGT*
sPOP2: AGCAGGTCAGTTACA*CGTCTCCAGT*
tPOP2: TCAGTCAGTCAGTTG*CGTCTCCAGT*
pPOP3: CGCTTCAGATGGTAC*AGTGCAGTCA*
sPOP3: ACACGATCCCAAGGT*AGTGCAGTCA*
tPOP3: GTTACTCAGGTCCCA*AGTGCAGTCA*
pPOP4: GCCTTGAACTGGACC*TGATCGACTG*
sPOP4: CATGACCGTGCTGAG*TGATCGACTG*
tPOP4: TGGACTGTGCTACCT*TGATCGACTG*
gadA-SSP1(5′): CATTTCCATAGGTTGCTCCAAGGTC
gadA-SSP2(5′): ACGTCATCTCAGTTGTTAGCCAACC
gadA-SSP3(5′): AGCCGGTTTGCTTTCAAATGATTCT
gadA-SSP1(3′): TGCGGATACTGATAACAAGACGACA
gadA-SSP2(3′): GGATTGAGAAAGAACGTACGGGTGA
gadA-SSP3(3′): TCCTGCATATCGGTAACGCCCAATC
ALDOA-SSP1(5′): AAATGCTGCAGCCTCCCTCTCACCC
ALDOA-SSP2(5′): AATACCAGAAATGTGCCCTCCCGTG
ALDOA-SSP3(5′): TGAGCTGGCAGGTTGTAGTCTCTGT
ALDOA-SSP1(3′): CCCTCGGACGATTGGACCTAGCTTG
ALDOA-SSP2(3′): GGTCTAACGGTGCCTCTCAGCCTCT
ALDOA-SSP3(3′): TCTGCCCTTCCCCATGGACGTAAGT
malQ-SSP1(5′): CTTCCTGGGTAAGCGTCAGCGTGTG
malQ-SSP2(5′): CAGCTTCGTCGGTAGATTGAACGCT
malQ-SSP3(5′): GGTGGTCAGCAGCCAGCTATATTCG
malQ-SSP1(3′): CGTCATCGCTGTATGGTGATTGGTG
malQ-SSP2(3′): CGGTGTTTACTCCTACAAAGTGCTC

malQ-SSP3(3′): GCTACATTGCCGACAGTAACAGTGC
hyg-SSP1(5′): CGGCAATTTCGATGATGCAGCTTGG
hyg-SSP2(5′): CGGGACTGTCGGGCGTACACAAATC
hyg-SSP3(5′): GACCGATGGCTGTGTAGAAGTACTC
hyg-SSP1(3′): AACTCCCCAATGTCAAGCACTTCCG
hyg-SSP2(3′): GAAACCATCGGCGCAGCTATTTACC
hyg-SSP3(3′): GAAAGCACGAGATTCTTCGCCCTCC

Solutions

1. 2.5× TBE buffer (see Recipes)
2. 0.5× TBE buffer (see Recipes)
3. 100 µM primer (see Recipes)
4. 10 µM primer (see Recipes)
5. 1.5% agarose gel (see Recipes)

Recipes

1. 2.5× TBE buffer

Reagent	Final concentration	Amount
0.5 M EDTA solution	5 mM	10 mL
Tris	225 mM	27 g
Boric acid	225 mM	13.75 g
ddH_2O	n/a	950 mL
Total	n/a	1,000 mL

Adjust the pH to 8.3 with 1 M NaOH and then replenish the solution to 1,000 mL with ddH_2O. This buffer can be stored at room temperature for three months.

2. 0.5× TBE buffer

Reagent	Final concentration	Amount
2.5× TBE buffer	0.5×	200 mL
ddH_2O	n/a	800 mL
Total	n/a	1,000 mL

This buffer can be stored at room temperature for three months.

3. 100 µM primer

Reagent	Final concentration	Quantity or Volume
Powdery primer	100 µM	n/a
1× TE buffer	1×	Volume specified in the sheet of primer synthesis
Total	n/a	Volume specified in the sheet of primer synthesis

Note: Dilute a portion of the 100 µM primer to prepare 10 µM primer and store the remaining portion at -80 °C.

4. 10 µM primer

Reagent	Final concentration	Quantity or Volume
100 µM primer	10 µM	1 µL
1× TE buffer	1×	9 µL
Total	n/a	10 µL

Note: Prepare extra volume of a 10 µM primer and divide it into multiple 1.5 mL microcentrifuge tubes; then, store the microcentrifuge tubes at -80 °C. Take one tube at a time and store it at -20 °C after use.

5. 1.5% agarose gel

Reagent	Final concentration	Quantity or Volume
Agarose	1.5%	1.5 g
0.5× TBE buffer	0.5×	100 mL

Green fluorescent nucleic acid dye (10,000×)	1×	10 μL
Total	n/a	100 mL

Laboratory supplies

1. 0.2 mL thin-wall PCR tubes (Kirgen, catalog number: KG2311)
2. 10 μL pipette tips (Sangon, catalog number: F600215)
3. 200 μL pipette tips (Sangon, catalog number: F600227)
4. 1,000 μL pipette tips (Sangon, catalog number: F630101)
5. 1.5 mL microcentrifuge tubes (Labselect, catalog number: MCT-001-150)

Equipment

1. PCR apparatus (Applied Biosystems, model: GeneAmp PCR System 2700)
2. Electrophoresis apparatus (Beijing Liuyi, model: DYY-6C)
3. Gel imaging system (Bio-Rad, model: ChemiDoc XRS+)
4. Microcentrifuge (Tiangen, model: TGear)

Software and datasets

1. Oligo 7 software (Molecular Biology Insights, Inc., USA)
2. DNASTAR Lasergene software (DNASTAR, Inc., USA)

Procedure

A. Design of primers

1. Design a POP set of three primers: primary POP (pPOP), secondary POP (sPOP), and tertiary POP (tPOP) (Figure 1).

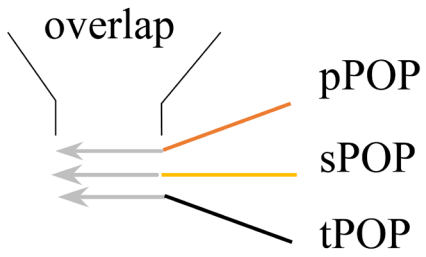


Figure 1. Interrelationship of the three primers pPOP, sPOP, and tPOP in a POP set. pPOP: Primary partially overlapping primer; sPOP: secondary partially overlapping primer; and tPOP: tertiary partially overlapping primer.

Critical: The sequences of the three POPs are completely arbitrary, comprising a 10 nt of 3′ overlap and 15 nt of 5′ regions heterologous to each other. Each POP shows a melting temperature of 65–70 °C according to Mazars et al. [29]. A POP itself should avoid forming severe hairpin or dimer structures; meanwhile, it should avoid forming severe dimers with the paired SSP.
Note: Here, the design of the POP1 set (comprising pPOP1, sPOP1, and tPOP1) is provided as an example.

a. Open the Oligo 7 software, click *File*, then click *New Sequence* (Figure 2A); type in a 25 nt arbitrary sequence as the initial pPOP1 in the *Edit Sequence* dialog box (Figure 2B), click *Accept/Discard*, then click *Accept* (Figure 2C).

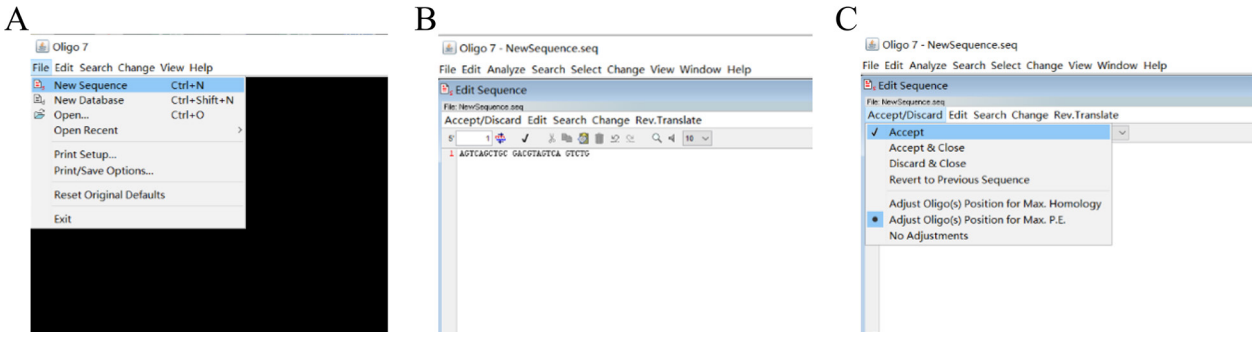


Figure 2. Screenshots showing how to enter an arbitrary DNA sequence in the software. Locations of *New Sequence* (A), *Edit Sequence* dialog box (B), and *Accept* (C) under the *File* tab.

b. Sequentially click *Analyze*, *Duplex Formation,* and *Current Oligo* (Figure 3A) to check the primer dimer (Figure 3B).

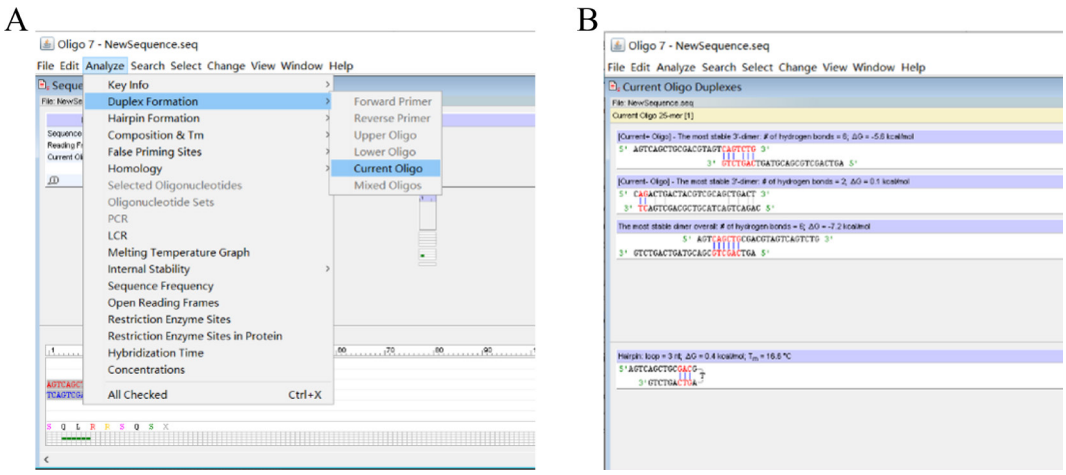


Figure 3. Screenshots showing how to check primer dimer. (A) Locations of *Duplex Formation* and *Current Oligo* under the *Analyze* tab. (B) Predicted primer dimers.

c. Sequentially click *Analyze*, *Hairpin Formation*, and *Current Oligo* (Figure 4A) to check the primer hairpin (Figure 4B).

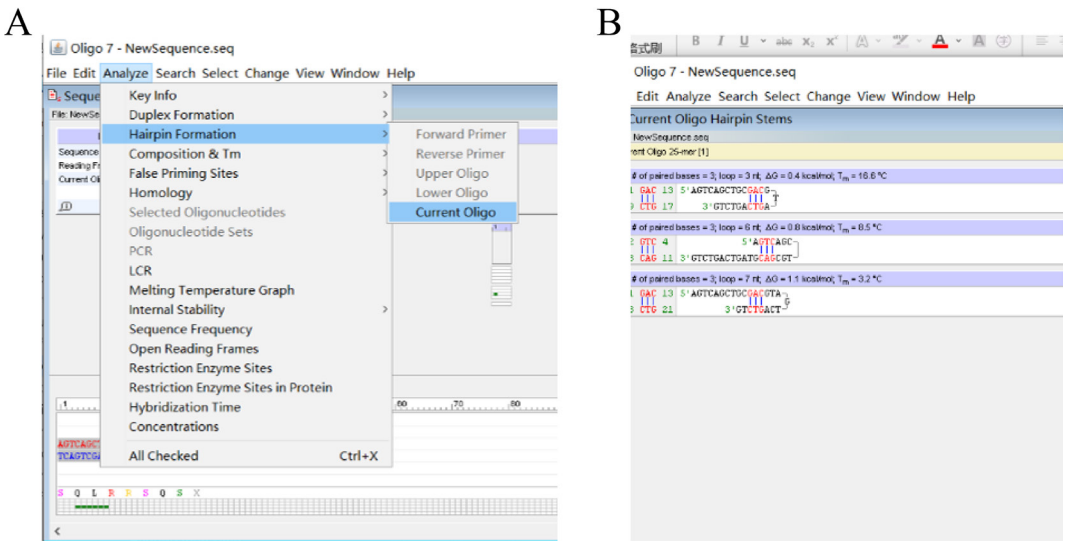


Figure 4. Screenshots showing how to check primer hairpin. (A) Locations of the *Hairpin Formation* and *Current Oligo* under the *Analyze* tab. (B) Predicted primer hairpins.

Note: Optimize the sequence of this initial pPOP1 as follows if it is unsatisfactory due to forming a severe primer dimer(s) or hairpin(s).

d. Click *Edit* and *Entire Sequence* (Figure 5A) to return to the *Edit Sequence* dialog box (Figure 2B); edit the sequence based on the above analysis results, click *Accept/Discard*, and click *Accept* (Figure 2C). Then, minimize this dialog box to show the dialog box shown in Figure 3A.

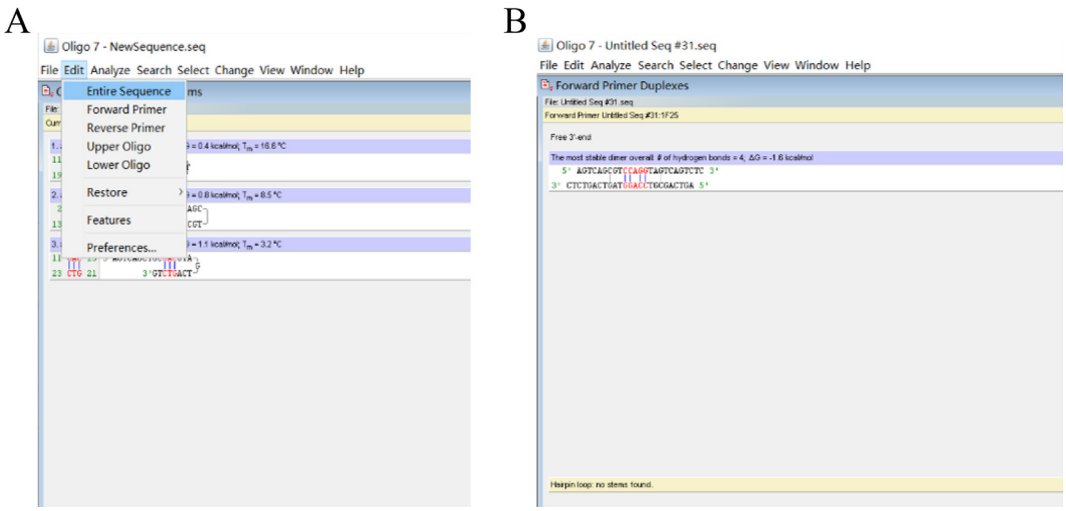


Figure 5. Screenshots showing how to optimize the sequence of the initial pPOP1. (A) Screenshot showing how to return to the *Edit Sequence* dialog box. (B) Predicted primer hairpins.

e. Repeat steps A1b and c to evaluate the edited pPOP1, focusing on the formation of primer dimer and hairpin, until a satisfactory pPOP1 (Figure 5B) is obtained.
Note: The pPOP1 shown in Figure 5B is satisfactory, as it only forms an acceptable primer dimer and does not form any hairpin.
f. Fix the 3′ part (10 nt) of this satisfactory pPOP1 as the overlap.
g. Attach an arbitrary 15 nt oligo to the 5′ end of this overlap, creating an initial sPOP1.
h. Evaluate this sPOP1 using the aforementioned method (namely steps from A1a–e) until a satisfactory sPOP1 is obtained.
i. Design a satisfactory tPOP1 similarly to designing the sPOP1.
Critical: Design more than one POP set, so as to perform parallel POP-PCRs in a genome walking cycle. A severe primer dimer should be avoided between any two POPs in a POP set.
Note: The four POP sets provided by this protocol are universal to genome walking of any species.

2. Select a set of three SSPs (outmost SSP1, middle SSP2, and innermost SSP3) from a known DNA along the direction of 5′ to 3′.
Note: Here, the design of gadA-SSP1(5') is provided as an example.
a. Open the Oligo 7 software, click *File* and then click *Open* to input the known DNA sequence of *gadA* gene (Figure 6A).

Figure 6. Screenshots showing how to select a satisfactory gadA-SSP1(5′). (A) Screenshot showing how to input a known DNA sequence. (B) Screenshot showing how to select an oligo with defined length. (C) Predicted primer dimers and hairpin.

b. Click *Change* and *Current Oligo Length* to define the length (for example 25 nt) of the oligo to be analyzed; move the mouse to a position of the input sequence (the 25 nt following the mouse is automatically shaded) and then click to select this shaded 25 nt (Figure 6B).
c. Evaluate the selected 25 nt oligo according to the method (Figures 3 and 4) for evaluating the pPOP1.
Note: If the current oligo is unsatisfactory due to forming a severe primer dimer(s) or hairpin(s), re-select and evaluate a new oligo by repeating steps A2a–c until a satisfactory gadA-SSP1(5′) (Figure 6C) is obtained.
d. Sequentially design the nested gadA-SSP2(5′) and gadA-SSP3(5′) like designing the gadA-SSP1(5′).

Critical: Each SSP shows a melting temperature of 65–70 °C [29] and should avoid forming severe hairpin or dimer structure.

B. POP-PCR procedure

As shown in Figure 7, a POP-PCR set contains three rounds of nested amplifications. Primary POP-PCR is driven by pPOP and SSP1; secondary POP-PCR is driven by sPOP and SSP2; and tertiary POP-PCR is driven by tPOP and SSP3.

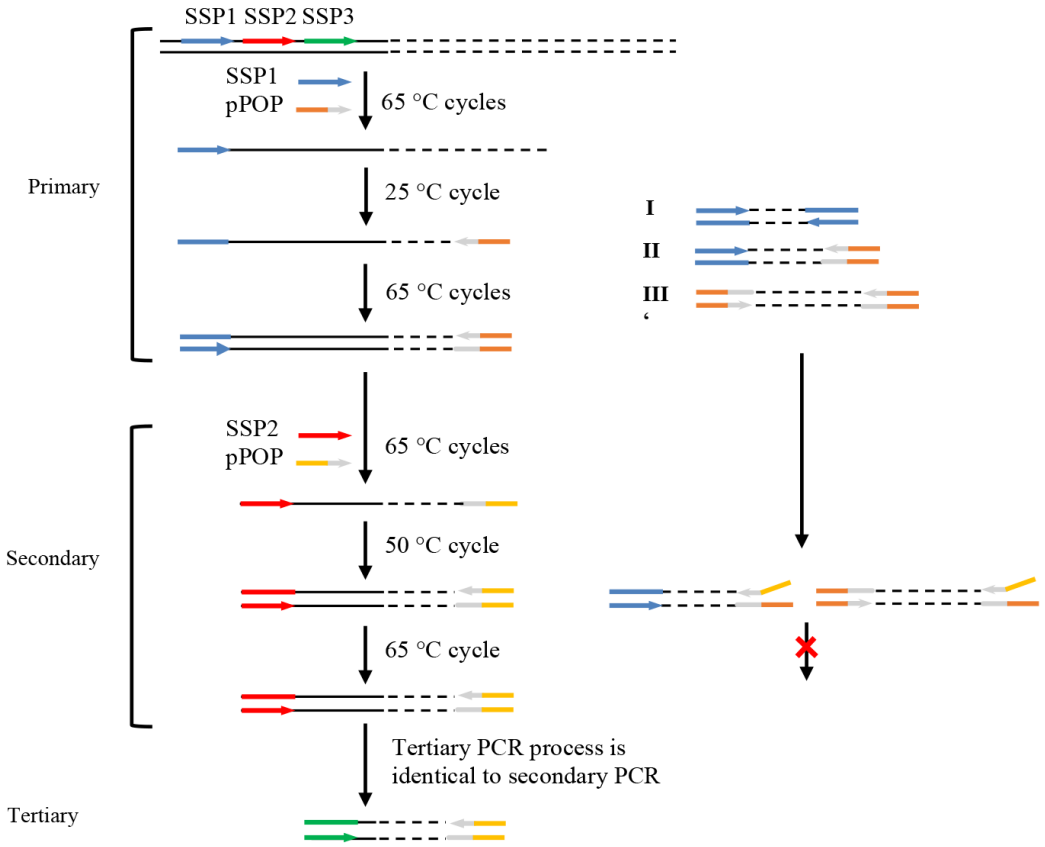


Figure 7. Schematic diagram of POP-PCR. Thin solid line: known sequence; thin dotted line: unknown sequence; arrows: primers; thick lines: primer complements. SSP: sequence-specific primer, pPOP: primary partially overlapping primer, sPOP: secondary partially overlapping primer, and tPOP: tertiary partially overlapping primer.

Note: The primary 25 °C cycle facilitates pPOP partially annealing with a site(s) on the unknown flank and elongating toward the known region to synthesize a target DNA, thus mediating the so-called genome walking. The secondary or tertiary 50 °C cycle facilitates the POP partially annealing with the previous POP site through the 3′ overlap. In each PCR, the target single-stranded DNA formed in that 25 °C or 50 °C cycle can be converted into a double strand by the SSP in the following 65 °C cycle and thus can be exponentially amplified by the remaining 65 °C cycles. However, non-target single-stranded DNA formed in that 25 °C or 50 °C cycle cannot be further processed in the following 65 °C cycles, as it lacks a perfect binding site for any primer.

1. Primary POP-PCR

a. Pipette primary POP-PCR components (Table 1) into a 0.2 mL PCR tube.

Table 1. Primary POP-PCR mix

Reagent	Final concentration	Amount (μL)
Genomic DNA	Microbe, 0.2–2 ng/μL; plant or human, 2–20 ng/μL	1
LA Taq polymerase (5 U/μL)	0.05 U/μL	0.5
pPOP (10 μM)	0.2 μM	1
SSP1 (10 μM)	0.2 μM	1
10× LA PCR buffer II (Mg^{2+} plus)	1×	5
dNTP mixture (2.5 mM each)	0.4 mM each	8
ddH_2O	n/a	33.5
Total	n/a	50

b. Completely mix the components with a pipette.

c. Centrifuge at 1500 g for 10–20 s at 4 °C with a microcentrifuge to gather the mixture.

d. Run the amplification in the PCR apparatus (Table 2).

Table 2. Primary POP-PCR cycling conditions

Step	Temp. (°C)	Duration	Cycle
Initial denaturation	94	1 min	1
Initial denaturation	98	1 min	1
Denaturation	94	30 s	5
Annealing	65	1 min	
Extension	72	2 min	
Denaturation	94	30 s	
Annealing	25	1 min	1
Extension	72	2 min	
Denaturation	94	20 s	
Annealing	65	1 min	30
Extension	72	2 min	
Final extension	72	5 min	1
Hold	4	forever	

e. Put the PCR product on ice.

f. Take 1 μL of the product as the template for secondary POP-PCR.

g. Store the remaining product at -20 °C for future assays.

2. Secondary POP-PCR

a. Pipette secondary POP-PCR components (Table 3) into a 0.2 mL PCR tube.

Table 3. Secondary POP-PCR mix

Reagent	Final concentration	Amount (μL)
Primary PCR product	n/a	1
LA Taq polymerase (5 U/μL)	0.05 U/μL	0.5
sPOP (10 μM)	0.2 μM	1
SSP2 (10 μM)	0.2 μM	1
10× LA PCR buffer II (Mg^{2+} plus)	1×	5
dNTP mixture (2.5 mM each)	0.4 mM each	8
ddH_2O	n/a	33.5
Total	n/a	50

Critical: Dilute primary POP-PCR product 10–1,000 fold if necessary.

b. Completely mix the components with a pipette.

c. Centrifuge at 1500 g for 10–20 s at 4 °C with a microcentrifuge to gather the mixture.

d. Run the amplification in the PCR apparatus (Table 4).

Table 4. Secondary POP-PCR cycling conditions

Step	Temp. (°C)	Duration	Cycle
Denaturation	94	30 s	5
Annealing	65	1 min	
Extension	72	2 min	
Denaturation	94	30 s	
Annealing	50	1 min	1
Extension	72	2 min	
Denaturation	94	30 s	
Annealing	65	1 min	30
Extension	72	2 min	
Final extension	72	5 min	1
Hold	4	forever	

e. Put the PCR product on ice.
f. Take 1 μL of the product as the template of tertiary POP-PCR.
g. Store the remaining product at -20 °C for future assays.

3. Tertiary POP-PCR
a. Pipette tertiary POP-PCR components (Table 5) into a 0.2 mL PCR tube.

Table 5. Tertiary POP-PCR mix

Reagent	Final concentration	Amount (μL)
Secondary PCR product	n/a	1
LA Taq polymerase (5 U/μL)	0.05 U/μL	0.5
tPOP (10 μM)	0.2 μM	1
SSP3 (10 μM)	0.2 μM	1
10× LA PCR buffer II (Mg^{2+} plus)	1×	5
dNTP mixture (2.5 mM each)	0.4 mM each	8
ddH_2O	n/a	33.5
Total	n/a	50

Critical: Dilute secondary POP-PCR product 10–1,000 fold if necessary.

b. Completely mix the components with a pipette.
c. Centrifuge at 1500 g for 10–20 s at 4 °C with a microcentrifuge to gather the mixture.
d. Run the amplification in the PCR apparatus (Table 6).

Table 6. Tertiary fork PCR cycling conditions

Step	Temp. (°C)	Duration	Cycle
Denaturation	94	30 s	5
Annealing	65	1 min	
Extension	72	2 min	
Denaturation	94	30 s	
Annealing	50	1 min	1
Extension	72	2 min	
Denaturation	94	30 s	
Annealing	65	1 min	30
Extension	72	2 min	
Final extension	72	5 min	1
Hold	4	forever	

e. Store the PCR product at -20 °C for future assays.

C. Gel electrophoresis

1. Completely mix 5 μL of POP-PCR product and 1 μL of 6× loading buffer.
2. Load the mixture into a 1.5% agarose gel supplemented with 1× green fluorescent nucleic acid dye.
3. Set the electrophoresis apparatus to a voltage of 150 V (the distance between the two electrodes is 30 cm).
4. Check the PCR product using the ChemiDoc XRS+ imaging system after approximately 25 min of electrophoresis (Figure 8).

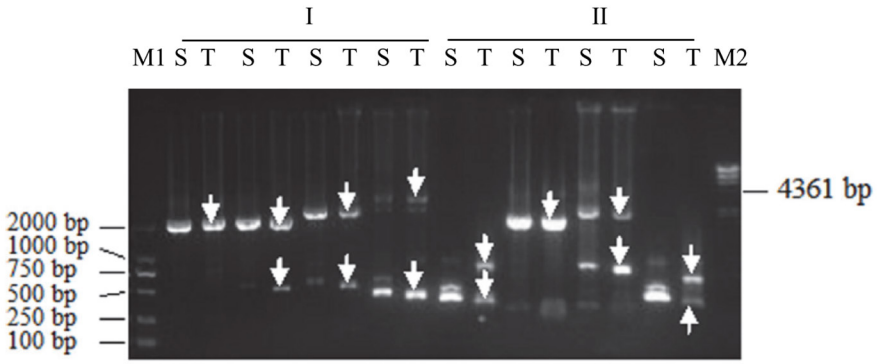


Figure 8. Genome walking of the *gadA* locus in *Levilactobacillus brevis* NCL912 I: walking into the 5′ region of *gadA*, and II: walking into the 3′ region of *gadA*. Each walking comprised four parallel POP-PCR sets: POP1, POP2, POP3, and POP4. The white arrows indicate the target bands. Lanes Ss: secondary PCR products; lanes Ts: tertiary PCR products; lane M1: DL2000 DNA marker; lane M2: λ-Hind III digest DNA marker.

D. Recovery of PCR product

1. Completely mix 40 μL of secondary/tertiary POP-PCR product and 8 μL of 6× loading buffer.
2. Load the mixture into a 1.5% agarose gel supplemented with 1× green fluorescent nucleic acid dye.
3. Set the electrophoresis apparatus to a voltage of 150 V (the distance between the two electrodes is 30 cm).
4. Visualize the PCR product using the ChemiDoc XRS+ imaging system. Subsequently, cut out clear DNA band(s) using a knife.
5. Extract the DNA band(s) from the cut gel using the DiaSpin DNA Gel Extraction kit.
6. Confirm the extracted DNA band(s) with 1.5% agarose gel electrophoresis.

E. DNA sequencing

Sequence the obtained DNA band(s) at Sangon Biotech Co., Ltd (Shanghai, China).

Data analysis

Use the *By Clustal W Method* function in MegAlign software to analyze the POP-PCR product.
1. Open the MegAlign software, click *File* and *Enter Sequences* (Figure 9A) to input a POP-PCR product and the corresponding known DNA segment (Figure 9B).
2. Click *Align* and *By Clustal W Method* (Figure 10A) to output the alignment result (Figure 10B).

Figure 9. Screenshots showing how to input DNA sequences. (A) Screenshot of the MegAlign software showing the location of *Enter Sequences* under the *File* tab. (B) Input DNA sequences.

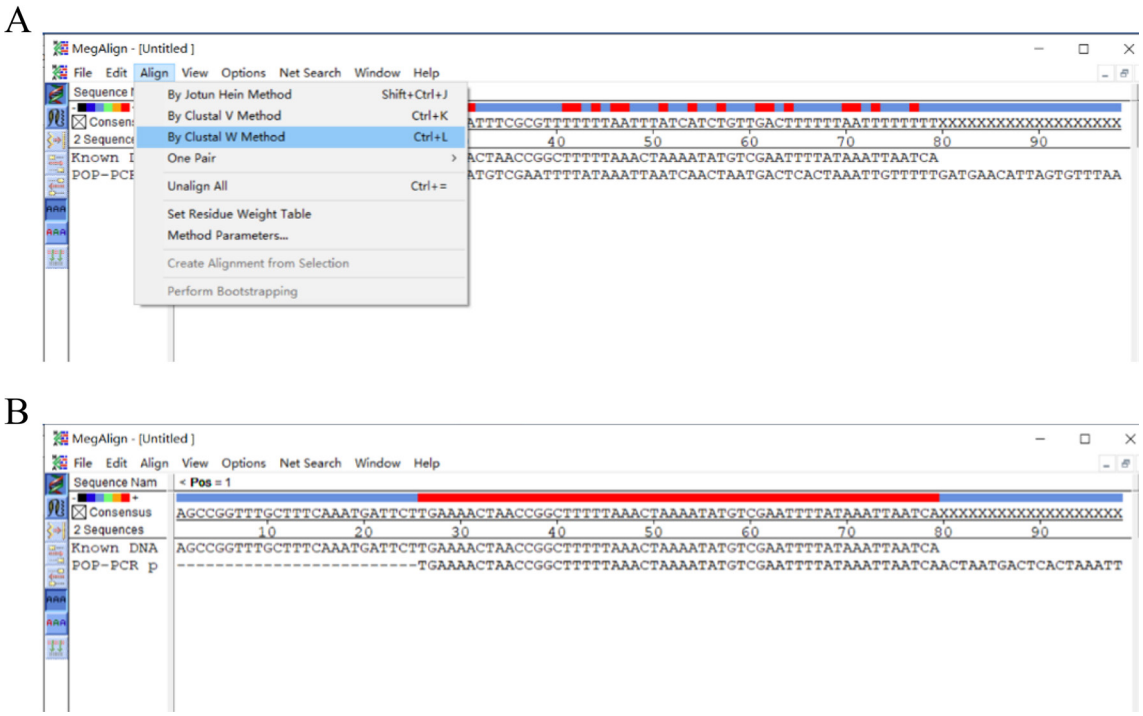


Figure 10. Screenshots showing how to align the input sequences. (A) Screenshot of the MegAlign software showing the location of *By Clustal W Method* under the *Align* tab. (B) Alignment result.

Note: The walking is considered successful if the SSP3-sided part of the POP-PCR product overlaps the known DNA (an example is shown in Figure 10B).

Validation of protocol

This protocol or parts of it has been used and validated in the following research articles:
• Li et al. [1] Partially Overlapping Primer-Based PCR for Genome Walking. *Plos One* (Figure 2).

Our POP-PCR has also been validated by several other studies. For instance, Zhang et al. [22] used the POP-PCR to successfully clone the tubulin gene and its promoter region and terminator region in *Tribonema minus*. Yuan et al. [23] used the POP-PCR to obtain the 3′ flanking sequences of the alpha-tubulin gene in *Haematococcus pluvialis* strains. Wada et al. [24] used the POP-PCR method to identify the insertion site of T-DNA in *Arabidopsis* T3 6-5. The results indicated that the T-DNA was integrated into *Arabidopsis* chromosome 3 at position 1 334 923.

General notes and troubleshooting

General notes

1. The developed POP-PCR comprises three rounds of nested PCR amplifications. However, secondary amplifications can generally release a positive result.
2. Before tertiary POP-PCR, we suggest checking secondary POP-PCR products using agarose gel electrophoresis to know if a clear DNA band(s) appears. If a clear DNA band(s) appears, it is generally not necessary to do tertiary PCR.
3. Like the other PCR-based genome walking schemes, the current POP-PCR scheme also suffers from the issue of multiple DNA bands. However, if multiple DNA bands appear, only the largest band needs to be sequenced.
4. Simultaneously performing parallel POP-PCRs will improve the success and efficiency of genome walking.
5. The current POP-PCR protocol is applicable to genome walking of any species.

Troubleshooting

Problem 1: No clear DNA band(s) is obtained after two or even three rounds of PCRs.
Possible cause: 1) In primary PCR, target amplification is weak; meanwhile, non-target amplification is strong. 2) The annealing temperature (50 °C) of the relaxed cycle in secondary/tertiary PCR is too high.
Solution: 1) Dilute the primary PCR product 10–10,000 times and then use 1 μL of each dilution as the template in the next POP-PCR. Afterward, tertiary PCRs are performed using the secondary PCR products as templates, respectively. 2) Lower the annealing temperature of the relaxed cycle. If still no clear DNA(s) appears in any secondary/tertiary POP-PCR, redesign the SSP set.

Problem 2: A POP-PCR product cannot be directly sequenced.
Possible cause: The non-target background interferes with the sequencing.
Solution: Clone the clear DNA band and then sequence.

Problem 3: Clear DNA band(s) is not a wanted product.
Possible cause: Genomic DNA may be contaminated.
Solution: Re-extract genomic DNA and ensure it is not contaminated; perform PCR amplification in a clean experimental area; ensure that the consumables are sterile.

Acknowledgments

This study was supported by the Major Discipline Academic and Technical Leaders Training Program of Jiangxi Province (grant No. 20225BCJ22023), China. This POP-PCR-based genome walking protocol has been originally described and validated in *Plos One* [1].

Competing interests

The authors declare no competing interests.

Received: September 13, 2024; Accepted: December 02, 2024; Available online: December 17, 2024; Published: February 05, 2025

References

1. Li, H., Ding, D., Cao, Y., Yu, B., Guo, L. and Liu, X. (2015). Partially overlapping primer-based PCR for genome walking. *PLoS One.* 10(3): e0120139.

2. Myrick, K. V. and Gelbart, W. M. (2002). Universal Fast Walking for direct and versatile determination of flanking sequence. *Gene.* 284: 125–131.
3. Tian, B., Wu, H., Wang, R., Chen, H. and Li, H. (2024). N_7-ended walker PCR: An efficient genome-walking tool. *Biochem Genet.* doi.org/10.1007/s10528-024-10896-1.
4. Fraiture, M. A., Papazova, N. and Roosens, N. H. (2021). DNA walking strategy to identify unauthorized genetically modified bacteria in microbial fermentation products. *Int J Food Microbiol.* 337: 108913.
5. Wang, R., Gu, Y., Chen, H., Tian, B. and Li, H. (2025). Uracil base PCR implemented for reliable DNA walking. *Anal Biochem.* 696: 115697.
6. Li, H., Lin, Z., Guo, X., Pan, Z., Pan, H. and Wang, D. (2024). Primer extension refractory PCR: an efficient and reliable genome walking method. *Mol Genet Genomics.* 299(1): 27.
7. Guo, X., Zhu, Y., Pan, Z., Pan, H. and Li, H. (2024). Single primer site-specific nested PCR for accurate and rapid genome-walking. *J Microbiol Methods.* 220: 106926.
8. Kotik, M. (2009). Novel genes retrieved from environmental DNA by polymerase chain reaction: Current genome-walking techniques for future metagenome applications. *J Biotechnol.* 144(2): 75–82.
9. Leoni, C., Volpicella, M., De Leo, F., Gallerani, R. and Ceci, L. R. (2011). Genome walking in eukaryotes.*FEBS J.* 278(21): 3953–3977.
10. Wei, C., Lin, Z., Pei, J., Pan, H. and Li, H. (2023). Semi-site-specific primer PCR: A simple but reliable genome-walking tool. *Curr Issues Mol Biol.* 45(1): 512–523.
11. Sun, T., Jia, M., Wang, L., Li, Z., Lin, Z., Wei, C., Pei, J. and Li, H. (2022). DAR-PCR: a new tool for efficient retrieval of unknown flanking genomic DNA. *AMB Express.* 12(1): 131.
12. Pei, J., Sun, T., Wang, L., Pan, Z., Guo, X. and Li, H. (2022). Fusion primer driven racket PCR: A novel tool for genome walking. *Front Genet.* 13: e969840.
13. Kalendar, R., Shustov, A. V. and Schulman, A. H. (2021). Palindromic sequence-targeted (PST) PCR, version 2: An advanced method for high-throughput targeted gene characterization and transposon display. *Front Plant Sci.* 12: e691940.
14. Kalendar, R., Shustov, A. V., Seppänen, M. M., Schulman, A. H. and Stoddard, F. L. (2019). Palindromic sequence-targeted (PST) PCR: a rapid and efficient method for high-throughput gene characterization and genome walking. *Sci Rep.* 9(1): 17707.
15. Evangelene Christy, S. M. and Arun, V. (2024). Isolation of actin regulatory region from medicinal plants by thermal asymmetric interlaced PCR (TAIL PCR) and its bioinformatic analysis. *Braz J Bot.* 47(1): 67–78.
16. Chang, K., Wang, Q., Shi, X., Wang, S., Wu, H., Nie, L. and Li, H. (2018). Stepwise partially overlapping primer-based PCR for genome walking. *AMB Express.* 8(1): 77.
17. Wang, L., Jia, M., Li, Z., Liu, X., Sun, T., Pei, J., Wei, C., Lin, Z. and Li, H. (2023). Protocol to access unknown flanking DNA sequences using Wristwatch-PCR for genome-walking. *STAR Protoc.* 4(1): 102037.
18. Wang, L., Jia, M., Li, Z., Liu, X., Sun, T., Pei, J., Wei, C., Lin, Z. and Li, H. (2022). Wristwatch PCR: A versatile and efficient genome walking strategy. *Front Bioeng Biotechnol.* 10: e792848.
19. Pan, H., Guo, X., Pan, Z., Wang, R., Tian, B. and Li, H. (2023). Fork PCR: a universal and efficient genome-walking tool. *Front Microbiol.* 14: e1265580.
20. Chen, H., Wei, C., Lin, Z., Pei, J., Pan, H. and Li, H. (2024). Protocol to retrieve unknown flanking DNA sequences using semi-site-specific PCR-based genome walking. *STAR Protoc.* 5(1): 102864.
21. Lin, Z., Wei, C., Pei, J. and Li, H. (2023). Bridging PCR: An efficient and reliable scheme implemented for genome-walking. *Curr Issues Mol Biol.* 45(1): 501–511.
22. Zhang, Y., Wang, H., Yang, R., Wang, L., Yang, G. and Liu, T. (2020). Genetic transformation of *Tribonema minus*, a eukaryotic filamentous oleaginous yellow-green alga. *Int J Mol Sci.* 21(6): 2106.
23. Yuan, G., Xu, X., Zhang, W., Zhang, W., Cui, Y., Qin, S. and Liu, T. (2019). Biolistic transformation of *Haematococcus pluvialis* with constructs based on the flanking sequences of its endogenous alpha tubulin gene. *Front Microbiol.* 10: e01749.
24. Wada, N., Kazuki, Y., Kazuki, K., Inoue, T., Fukui, K. and Oshimura, M. (2016). Maintenance and function of a plant chromosome in human cells. *ACS Synth Biol.* 6(2): 301–310.
25. Li, H., Gao, D., Cao, Y. and Xu, H. (2008). A high γ-aminobutyric acid-producing *Lactobacillus brevis* isolated from Chinese traditional paocai. *Ann Microbiol.* 58(4): 649–653.
26. Li, H., Li, W., Liu, X. and Cao, Y. (2013). *gadA* gene locus in *Lactobacillus brevis* NCL912 and its expression during fed-batch fermentation. *FEMS Microbiol Lett.* 349(2): 108–116.
27. Wang, Q., Liu, X., Fu, J., Wang, S., Chen, Y., Chang, K. and Li, H. (2018). Substrate sustained release-based high efficacy biosynthesis of GABA by *Lactobacillus brevis* NCL912. *Microb Cell Fact.* 17(1): 80.

28. Gustafson, S., Proper, J. A., Bowie, E. and Sommer, S. S. (1987). Parameters affecting the yield of DNA from human blood. *Anal Biochem.* 165(2): 294–299.
29. Mazars, G. R., Moyret, C., Jeanteur, P. and Theillet, C. G. (1991). Direct sequencing by thermal asymmetric PCR. *Nucleic Acids Res.* 19(17): 4783–4783.