A Novel and Simple PCR Walking Method for Rapid Acquisition of Long DNA Sequence Flanking a Known Site in Microbial Genome

Peng Luo · Ting Su · Chaoqun Hu ·
Chunhua Ren

Abstract Acquisition of flanking sequence adjacent to a known DNA site is an important task in microbial genome-related research. In this study, we developed a new method containing two rounds of PCR followed by cloning and sequencing. Firstly, specific primer (SP) is added into the reaction system for primary locus-specific linear amplification, and then a complex long primer (CLP) is added into the cooled reaction system for only one cycle. Amplification products from the first round of PCR are directly purified without electrophoresis, diluted, and used as the templates of the second PCR. Secondly, one long specific primer (LSP) and one long base-fixed primer (LFP) are adopted. The amplicons are purified for cloning and sequencing. The achievement of specific amplification for long flanking region mainly depends on ingenious and precise settings of PCR programs, structure design of CLP primer, adding of CLP primer after specific linear amplification, concentration ratio of CLP and SP primer, applying long primers, etc. Through this method, we successfully performed the long PCR walkings (>1.5 Kb) on *rpoB* gene of *Vibrio vulnificus*, transposon-like gene of *V. alginolyticus*, and *sto* gene of *V. cholerae*. The method provides a robust and simple strategy for rapid amplification of long unknown DNA fragments from microbes.

Keywords Flanking sequence · Long PCR ·
Walking · Microbial genome · *Vibrio vulnificus* ·
Vibrio alginolyticus · *Vibrio cholerae*

P. Luo · T. Su · C. Hu (✉) · C. Ren
Key Laboratory of Marine Bio-resources Sustainable Utilization,
CAS, South China Sea Institute of Oceanology, Chinese
Academy of Sciences, Guangzhou 510301, China
e-mail: cqhu@scsio.ac.cn

Introduction

Acquisition of flanking sequence adjacent to a known DNA site is an important task in microbial genome-related research. For instance, we explore the insert sites of bacterial transposons and the genes flanked by transposase gene [1], or obtain the full lengths of functional genes in metagenomics [2, 3]. Construction and screening of genomic libraries are generally demanded to meet this objective. However, the methods are time-consuming and laborious. Thus, convenient PCR-based methods have been developed instead of construction and screening of genome library. Some classical methods include inverse PCR [2, 4–6], PCR of various adapters (vectorett or linker) [7–11], restriction-site PCR [12–14], and TAIL-PCR [15–20], which have extensive applications to genome walking in various organisms including microorganisms.

However, these methods still have their respective merits and demerits in practice. The methods originated from intramolecular circularization or adapter-mediated PCR demand enzyme digestion of genomic DNA and ligation before PCR. Owing to unavailable information in flanking region, generally, combinations of several enzymes must be tried to increase the probability of generating convenient DNA fragments [21]. Also, high quality of genomic DNA is required by these methods. In addition, the efficiencies of intramolecular ligation efficiency and adapter ligation have been main limitations for successful construction of amplification templates [22]. Owing to these limitations, in actual operations, an extensive protocol often does not work well for the acquisition of flanking unknown DNA sequence, which entails further costs. Of these methods originated from semi-random primer PCR, TAIL-PCR is a representative method and widely used in cloning unknown flanking regions. However, to obtain a consistent PCR

product, time-consuming exploration of the precise manipulation of cycling conditions is required [23]. In addition, conventional TAIL-PCR is prone to generate small products [22], resulting in retrieves of insufficient information in one genome walking. In order to achieve higher success rate in obtaining target products with larger size, an improved method named hiTAIL-PCR has been developed [24], but the sizes of main products are still less than 1.5 kb.

Recently, a new high-throughout genome-walking method was introduced [17], which uses partially degenerate walker–adapter primer and Phi29 DNA polymerase with strand-displacement DNA synthesis activity. This leads to amplification of nonbiased whole genome into new overlapping fragments, with walker–adapter sequence attached to the 5′ end of them, using for the templates of subsequent nested PCR. However, the nonspecific amplification products that emerged in the above study could not easily be excluded by hot-start PCR or touchdown PCR [17].

To surmount the limitations inherent in the above mentioned methods and acquire flanking fragments as long as possible in a single operation process, we developed in our study a simple and efficient PCR walking method for rapid acquisition of long unknown DNA sequence flanking a known site, which just needs two steps of PCR and subsequent sequencing, without prior digestion, ligation, and tailing steps. Through this method, we successfully performed long PCR walkings on *rpoB* gene of *Vibrio vulnificus*, transposon-like gene of *V. alginolyticus*, and *sto* gene of *V. cholerae*. This method promises to be an effective approach for the isolation of long flanking DNA fragment in microbial genomes.

Materials and Methods

Overview of the Method

The method consists of two rounds of PCR followed by cloning and sequencing (Fig. 1). In the first round of PCR, one specific primer (SP) with the length of 20 bases and one complex long primer (CLP) are adopted. CLP primer contains a fixed 29-base oligonucleotide segment at its 5′ end, a 4-base completely degenerate site (NNNN) at its middle section, and a 7-base terminal segment (G+C ≥ 3). Every 7-base terminal segment contains one 6-base restriction enzyme digestion site, two 3-base repeat sequence, one 4-base restriction enzyme digestion site plus one purine and one pyrimidine, or two 2-base repeat plus one purine and one pyrimidine, and a T or A at the last base site. The degenerate section and 7-base terminal segment constitute an actual annealing site with DNA templates, allowing the use of a relatively low annealing temperature (40–43°C). SP primer is first added into reaction system for primary locus-specific

linear amplification at relatively high annealing temperature (58–64°C), and then CLP primer is added into quickly cooled reaction system for only one cycle. In this cycle, low denaturing temperature (80°C) is adopted instead of conventional denaturing temperature (92–95°C) to decrease the opportunity of CLP primer binding with original genomic DNA, as it is designed to anneal with newly synthesized single-strand DNA (ssDNA) and produce complementary strand.

Amplified product of the first round of PCR is directly purified without agarose gel electrophoresis to remove the unincorporated primers, diluted, and used as the template of the second PCR. In the second PCR, one 29-base long specific primer (LSP), which has identical sequence with 29 fixed bases at 5′ end of CLP primer, and one long base-fixed primer (LFP), are adopted. Both are simultaneously added into reaction system and a routine long PCR is performed. Amplicons are purified for convenient TA cloning and sequencing, and a pair of specific primers (named CSP primer) located in internal region of known sequence is used in another PCR for screening positive clones.

### Bacterial Strains and Genome DNA Extraction

Vibrio vulnificus 1.1758 was purchased from the Centre of General Microbial Culture Collection (Beijing, China). *V. cholerae* HN375 and *V. alginolyticus* E0601 were isolated in our laboratory and deposited in China Center for Type Culture Collection (Wuhan, China) with accession number CCTCC AB209168 and CCTCC AB209169. The strains were cultured overnight in alkaline peptone water (APW) at 30°C, and the genomic DNAs were extracted with the MiniBEST Bacterial Genomic DNA Extraction Kit (TaKaRa, China) as per the manufacturer's protocol.

Feasibility Verification of the Method

To confirm the feasibility of the method, one long PCR walking on *rpoB* gene of *Vibrio vulnificus* 1.1758 was first carried out. As we did not know *rpoB* sequence of the strain, specific primer rpoF1 and long specific primer rpoF2 were designed based on *rpoB* sequence of *V. vulnificus* CMCP6 (Accession number in NCBI: AE016795, 4029 bp). A randomly selected 29-base primer RL served as LFP primer in all the experiments. Ideally, 3′ end of primer T2 can closely match with two sections at *rpoB* downstream region of *V. vulnificus* 1.1758 (supposing that the 3′ end of primer T2 is randomly designed, and *V. vulnificus* 1.1758 has the identical *rpoB* sequence with *V. vulnificus* CMCP6). All the primers used in this study are summarized in Table 1.

The compositions of PCR mixture used in this study are listed in Table 2. In the first round of PCR, genomic DNA of *V. vulnificus* 1.1758 was used as the template, and the reaction was performed under the following conditions: an
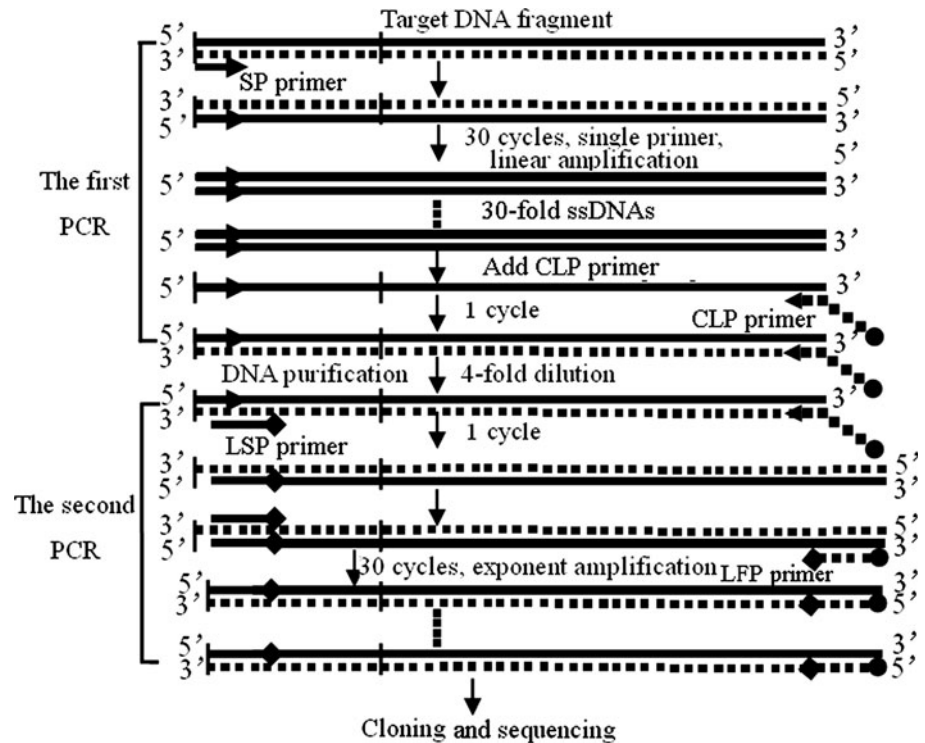
initial denaturation step of 94°C for 4 min, followed by 30 cycles of amplification; each cycle consisted of denaturation at 92°C for 20 s, annealing at 58°C for 1 min, and extension at 68°C for 4 min; then the reaction system was quickly cooled to 12°C for 10 min, and 1 μl of primer T2 was added into reaction tube; after the addition of primer T2, the reaction system was quickly skipped to one PCR cycle containing denaturation at 80°C for 30 s, annealing at 43°C for 2 min, and extension at 68°C for 4 min, followed by the final extension at 68°C for 8 min.

The first PCR product was directly purified by PCR Cleanup Kit (Axygen, China) without agarose electrophoresis and diluted to 100 μl, to be used for the second nested PCR. PCR was initialized with a predenaturation at 94°C for 3 min, followed by 30 cycles of amplification including denaturation at 92°C for 20 s, annealing and extension at 68°C for 4 min, and then terminated by an extension at 68°C for 8 min.

The second PCR products were loaded into a 0.6% agarose gel containing eithidium bromide and electrophoresed at 100 V for 30 min followed by visualization under UV light.

Long PCR Walking into Flanking Region of Transposase-like Gene from *Vibrio alginolyticus* E0601

The amplification of long flanking region of a known DNA fragment from *Vibrio alginolyticus* E0601 was performed as a practical application of our method. We had obtained a 1928-bp DNA sequence containing predicted transposase-like gene

of *V. alginolyticus* E0601 before genome walking [25]. The sequence has no similarity with any other published DNA sequence of *V. alginolyticus* except that the part of transposase-like gene is very similar to *vpiT* gene in VPI island of *V. cholerae*. We hope to acquire the long flanking regions of the known segment in *V. alginolyticus* E0601 genome through the method proposed in this study.

Specific primer VapF1 and long specific primer vapF2 were designed based on the known 1928-bp DNA fragment of the strain. Besides, six long complex primers (NRL1–NRL6) were adopted as CLP primers. The conditions for the first PCR for this study were identical to those of above mentioned first PCR of *V. vulnificus* except that the low stringent annealing temperature was reduced from 43 to 40°C and extension time was enhanced from 4 min to 6 min.

The first PCR products were purified and diluted according to the above mentioned approach, to be used for the templates of the second PCR. The conditions of the second PCR were identical with the second PCR of *V. vulnificus* described above except that the extension time was enhanced from 4 to 6 min. The second PCR products were electrophoresed and photographed under the aforementioned condition.

The PCR Walking into Unknown Flanking DNA Region of *sto* Gene in *Vibrio cholerae* HN375

Previously, we acquired a 325-bp sequence containing heat-stable enterotoxin gene (*sto*) from *V. cholerae* HN375

Table 1 Primers used in the study

Primers	Types of primers	Sequence 5′–3′	Length
rpoF1	SP	TGGTTTACTCTTATACCGAG	20
VapF1	SP	GTTCTATCCGACCATAGATG	20
ST-F1	SP	CTATTCATTGCATTAATGTT	20
rpoF2	LSP	TATACCGAGAAAAAGCGCATCCGTAAGGA	29
VapF2	LSP	TGGGCTAACCTCGGATAGACAGGATTACT	29
ST-F2	LSP	CAGTGCAGCAACCACAACAAATTGAAAGC	29
RL	LFP	TAGTCATGCCATGTCAGTTCTAGCATCTG	29
T2	CLP	RL-NNNNCGTCTTC	40
NRL1	CLP	RL-NNNNCTGCTGA	40
NRL2	CLP	RL-NNNNGACGACT	40
NRL3	CLP	RL-NNNNCAGCAGT	40
NRL4	CLP	RL-NNNNGTCGTCA	40
NRL5	CLP	RL-NNNNGCTCTCA	40
NRL6	CLP	RL-NNNNGGATCCA	40
NRL7	CLP	RL-NNNNCAGCTGA	40
NRL8	CLP	RL-NNNNGACGTCA	40
NRL9	CLP	RL-NNNNCCTAGGA	40
NRL10	CLP	RL-NNNNTCGTACA	40
rpoF3	CSP	ATCGAACAGGATCCTGAAGG	20
rpoR3	CSP	GGTTGCCAAGGTGGTCGATAT	21
VapF3	CSP	GAATTACAAGCTCGATCTTCAGC	23
VapR3	CSP	AGGCCAACCTATATCTTCTAATC	23
ST-F3	CSP	GTGAAAATGAAGAATGCCCAT	21
ST-R3	CSP	TAGTGKTATGCGTTGCCAGC	20

The 5′ end of each of the primers T2 and NRL1–NRL10 has identical sequence with the primer RL. *SP* specific primer, *LSP* long specific primer, *CLP* complex long primer, *LFP* long base-fixed primer, *CSP* primers used for clone screening

Table 2 Compositions of PCR systems

Reaction	Grouping	Type of primers	PCR walking					
			V. vulnificus		*V. aliginolyticus*		*V. cholerae*	
			Primer name	Usage	Primer name	usage	Primer name	Usage
First PCR		SP	rpoF1	A	VapF1	A	ST-F1	A
		CLP	T2	B	NRL*x*	B	NRL*y*	B
Second PCR	Test group	LSP	rpoF2	C	VapF2	C	ST-F2	C
		LFP	RL	C	RL	C	RL	C
	Control group 1	LFP	RL	B	RL	B	RL	B
	Control group 2	LSP	rpoF2	B	VapF2	B	ST-F2	B
PCR for clone screening		CSP	rpoF3	C	VapF3	C	ST-F3	C
		CSP	rpoR3	C	VapR3	C	ST-R3	C

A: 0.5 µl (1 µM); B: 1 µl (10 µM); C: 0.5 µl (10 µM); NRL*x*: NRL1, NRL2, NRL3, NRL4, NRL5, or NRL6; NRL*y*: NRL1, NRL2, NRL3, NRL4, NRL5, NRL6, NRL7, NRL8, NRL9, or NRL10. Except for primers, other components in PCR system contain 10× PCR buffer 2.5 µl, dNTP 2 µl, 0.25 µl LA DNA polymerase (5 U/µl), and ddH$_2$O to 25 µl. LA DNA polymerase: LA Taq (TaKaRa, China), dNTP: dNTP mixture (2.5 mM each)

and sequence analysis indicated that the flanking region of *sto* gene in *V. cholerae* HN375 were likely different from those in other *V. cholerae* strains. Therefore, we hope to study the characteristics of downstream sequence of *sto* gene in *V. cholerae* HN375, and amplification of flanking region of this gene was performed as another practical application of our method.

Specific primer ST-F1 and long specific primer ST-F2 were designed on the basis of 325-bp known DNA segment. Beside the aforementioned primers (NRL1–NRL6),

four new CLP primers (NRL7–NRL10) were also used. The two PCR amplifications were performed using the methods described above.

Cloning and Sequencing

To further verify that the aforementioned PCR products in test groups were authentically primed by primer RL and LSPs, we excised some bands that only occurred in the electrophoresis gels of test groups, purified them with Agarose Gel DNA Recovery Kit (TaKaRa, China), and cloned them into pMD18-T vector (TaKaRa, China). Pairs of internal specific primers (CSP primers, see Table 1) anchoring with known DNA regions were used for the identification of positive clones. The bacterial strains containing recombined plasmid were directly sent for plasmid extraction and subsequent sequencing (in Invitrogen Co., Ltd., with Applied Biosystems 3730 Automatic Sequencer). The sequences were spliced and deposited in the GenBank database (HM622073, EU787499, and GU598214).

Results

PCR Walking in *rpoB* Gene of *Vibrio vulnificus* 1.1758

As expected, two fragments with the sizes of 3044 and 4055 bp were obtained in the second round of PCR. As shown in Fig. 2, three obvious bands, A, B, and C in lane 1 were produced in the second amplification of *rpoB* gene from *V. vulnificus* 1.1758 using this method; but no bands occurred in lanes 2 and 3. The sizes of bands A and B were consistent with the size of predicted two bands, revealing that they should be the bands generated by specific amplification. No band occurred in lanes 2 and 3, showing that single primer rpoF2 or RL had little possibility to generate nonspecific amplification. However,
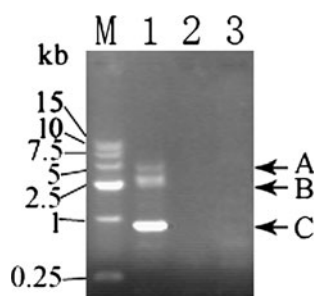
Fig. 2 PCR walking of *rpoB* gene from *Vibrio vulnificus* 1.1758. *Lane M*: 15 Kb ladder. *Lane 1*: amplified products from the second round of PCR using primers rpoF2 and RL, *lanes 2 and 3*: there were no amplified products from the second round of PCR using primer RL or rpoF2 alone. *A*, *B*, and *C* represent three obvious bands in *lane 1*

band C with the length of approximate 900 bp occurred in lane 1, while no bands existed in corresponding horizontal position in lanes 2 and 3, which indicated that band C was indeed generated by specific amplification. The bands A, B, and C were excised, reamplified, and cloned for sequencing. The sequencing results showed that the three bands were specific amplification products primed by the primers rpoF2/RL as both ends of every acquired sequence were identical with the sequences of the primers rpoF2 and RL, respectively. The section of seven fixed bases (CGTCTTC) at 3′ end of the primer T2 had only one mismatch site (the third base, T) with the upstream segment of acquired *rpoB* gene sequence from *V. vulnificus* 1.1758, which led to a 865-bp band (band C) in the second PCR as well. The sequences were assembled, and a complete sequence was deposited in GenBank database (HM622073). The results obtained in this study documented the feasibility of the new method for PCR walking into long flanking unknown genomic region.

PCR Walking into Long Flanking Region of Known DNA Fragment from *Vibrio alginolyticus* E0601

Figure 3 shows the typical amplification result from the second round of PCR walking for *Vibrio alginolyticus* E0601. A band at 1.0–2.5 Kb (lane 1) and a band of about 4.0 Kb (lane 5) were obtained by this method using primers NRL1 and NRL5, respectively. No bands were produced in the control groups in second round PCR using LFP primer (RL) or LSP primer suggesting that nontarget products primed by the LFP primer or LSP primer alone did not appear and the bands in lane 1 and lane 5 were most likely produced by the amplification with both primer RL and long specific primer VapF2. Sole band in the lane 1 or the lane 5 suggested that the primers NRL1 and NRL5 could anchor a flanking DNA site in the extension range of the first PCR with 6-min extension time, while other primers (NRL2, NRL3, and NRL4) could not. The use of the primer NRL6 did not lead to any PCR product either in the test group or in the control group (data not shown).

The band in lane 5 (Fig. 3) was excised and purified for cloning and sequencing. One complete sequence of 3820 bp in length was obtained by successive sequencing and sequence splicing, and the sequence (remove exogenous adapter sequence and vector sequences) has been deposited in GenBank database (EU787499). Two terminal portions of the complete sequence contain the sequences of VapF2 and NRL5, individually, and the upstream of complete sequence also contains the known DNA region (992 bp). The results demonstrated that the DNA fragment in lane 5 was the true extension of the known region using the primers vapF2 and RL, and our method could acquire nearly 4000-bp flanking fragment by two rounds of PCR.
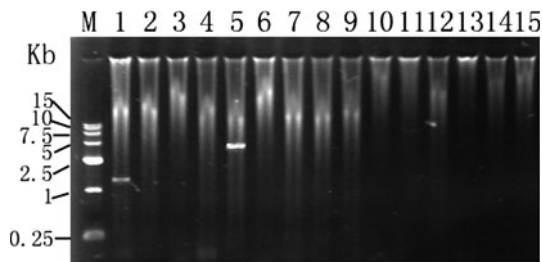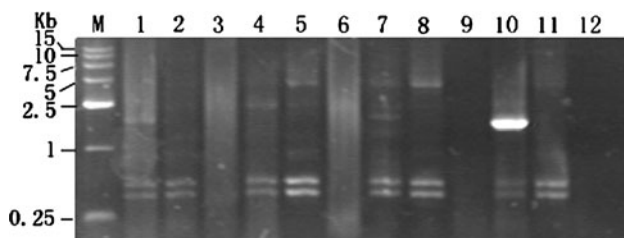
Fig. 3 The typical electrophoresis result for long PCR walking into flanking DNA region of *Vibrio alginolyticus* E0601. *Lane M*: 15-Kb ladder, *lanes 1–5*: amplified products from the second round of PCR using the primers VapF2 and RL, *lanes 6–10*: amplified products from the second round of PCR using primer RL alone, served as control group 1, *lanes 11–15*: amplified products from the second round of PCR using the primer vapF2 alone, served as control group 2

PCR Walking into Flanking Region of Acquired *sto* Gene from *Vibrio cholerae* HN375

Figure 4 shows the typical amplification result from the second round of PCR walking into *sto* flanking region of *V. cholerae* HN375 using the CLP primers, NRL7–NRL10. A clear band of approximatively 1600 bp in length from test group (lane 10) was detected when using primer NRL10 in the first round of PCR. Though there were two weak and short fragments (between 250 and 1000 bp) in the lanes of the test groups, the identical amplified fragments in size also existed in the control group 1, where only the primer RL was used in the second round of PCR. Thus, through size comparison of PCR fragment between the test and control groups, these weak and short fragments was identified to be nonspecific amplification caused by the primer RL alone. Using primers NRL1–NRL6 did not produce any specific amplification product except that two weak and short fragments were produced, which were consistent with the nonspecific fragments in other test groups (data not shown).

Fig. 4 The typical electrophoresis result for PCR walking into unknown DNA region of *Vibrio cholerae* HN375. *Lane M*: 15 Kb ladder, *lanes 1, 4, 7, and 10*: amplified products from the second round of PCR using the primers ST-F2 and RL, *lanes 2, 5, 8, and 11*: amplified products from the second round of PCR using the primer RL alone, served as control group 1, *lanes 3, 6, 9, and 12*: amplified products from the second round of PCR using the primer ST-F2 alone, served as control group 2

The clear band in lane 10 was excised and purified for cloning and sequencing. A 1632-bp sequence was obtained, and a splicing 1798-bp sequence containing part of the previously known 325-bp sequence was deposited in GenBank database (GU598214). 1632-bp cloning sequence contained the sequences of primers ST-F2 and NRL10, and the part of the 325-bp known sequence. The sequencing results strongly demonstrated that the DNA fragment in lane 10 was the true extension of the known *sto* region caused by primers ST-F2 and RL, and the reliability of our method was confirmed once again.

Discussion

Given that the construction and screening of genomic libraries are too laborious and time-consuming, researchers have been working on rapid and simple methods for acquiring flanking sequences. Although TAIL-PCR, as a representative method, has been widely used toward this aim in plants and other organisms including microbes [15–20], it does not always work well in actual application. The often-encountered problems include: (1) The method tends to produce undesired smaller products; (2) Nonspecific products (primed by SP or AD primer alone) still exist after three rounds of PCR; (3) It seems that specific products appeared in the second round of PCR, but they disappeared in the last round of PCR; (4) The procedure of TAIL-PCR is so complicated and the reiterative exploration of optimizing procedure are often needed for a good result. In the following improved method, named hiTAIL-PCR, the authors modified the PCR procedure of TAIL-PCR and improved the design of degenerate primers to decrease the production of smaller products, but some small molecular products still appeared in the figures offered by other authors [24]. In other aspects, authors did not do more modification compared with the primary TAIL-PCR. Thus, the encountered problems in conventional TAIL-PCR may appear in hiTAIL-PCR in the same way. In effect, we think the flaws of TAIL-PCR (hiTAIL-PCR) arise from the following factors: (1) High degeneracy at the 3′ end of AD (LAD) primers will result in too many anchoring sites with genomic DNA in small extension range. (2) Before the new amplification, the products are not purified to remove the superfluous SP or AD primers. (3) So long as low-stringency annealing exists in multiple cycle process, there are enough chances for exponential amplification of nonspecific products.

In order to overcome the existing problems and the flaws in TAIL-PCR and other walking methods, we developed the method proposed in this study. In this study, we demonstrated the specificity and the efficiency of the new method by the feasibility verification, the practical

applications, and the sequencing inspection. The achievement of long PCR walking depends on the following key measures:

(1) After linear amplification with SP primer, one CLP primer is added into reaction system for only one low stringency cycle. CLP primer has no use for the primary linear amplification, and the subsequent addition of CLP primer largely decreases the possibility of nonspecific amplification triggered by this primer. In other methods, such as TAIL-PCR [15], restriction-site PCR [12], and multiple strand-displacement PCR [21], adding arbitrary primers at the very start and the existence of multiple low stringency cycles result in nontarget products primed by arbitrary primers or specific primers.

(2) In only one low stringency cycle, low denaturing temperature (80°C) causes only a few parts of primary genomic DNA to attain a completely molten form, which largely decreases the binding of CLP primer and SP primer with primary genomic DNA. Meanwhile, newly synthesized specific ssDNA have higher copies than primary genomic DNA (nearly 30-fold increase) and are not complementary with each other. Therefore, under low denaturing temperature, they have more chances to be anchored by CLP primer.

(3) Concentration control of primers: the ratio of SP primer to CLP primer is 1 to 10. The amount of SP primer only maintains efficient linear amplification while superfluous CLP primer has sufficient chances to bind with the newly synthesized ssDNA in the last low stringency cycle.

(4) Annealing temperature in the last stringency cycle is another factor related to the specificity and the number of amplified DNA bands (target and nontarget). A much low annealing temperature will lead to too many binding sites of CLP primer. On the one hand, it facilitates generating a spot of nonspecific DNA segments that have identical end sequence (two new ssDNA primed by walker-adapter primer are complementary); on the other hand, it gives rise to the occurrence of some short specific DNA segments, which will suppress amplification of long PCR products. In the method of this study, annealing with temperature range of 40–43°C ensures that 3′ end of CLP primer (7 fixed bases plus 4 degenerate bases) can closely match with the templates.

(5) Design of CLP primer: seven fixed bases at 3′ end of CLP primer ensure it utmost has only one binding site in sufficiently large range, which facilitates generating long amplification and decreasing the number of amplified bands. Also, the structure of the 3′ ends of CLP primers allows using a lower annealing temperature (<40°C) and increasing mismatch sites to enlarge the universality of available CLP primers.

(6) A pair of long primers is used in the second PCR (nested PCR), which increases the annealing temperature and enhances the specificity.

(7) The first amplification products are directly purified without electrophoresis steps to remove unincorporated primers, and diluted PCR products are used as templates to decrease nonspecific amplification derived from primary genomic DNA and the interference of superfluous SP or CLP primer to the second round of PCR. These key measures also embody the differences and the improvements of our method compared with other methods.

In general, there are two main nonspecific products in hemispecific PCR: one primed by specific primer alone (type I) and the other primed by arbitrary primer (type II) [15]. The former can be eliminated simply by successive reactions with long specific primers [15]. In this method of our study, a series of measures mentioned above (including nested PCR with long primers) can greatly decrease type I nonspecific products and type II nonspecific products. In addition, a very simple strategy is adopted in the method to distinguish two types of nonspecific products, that is, every test group (using LSP primer and LFP primer) is performed with two controls in the second round of PCR. In the first control, only LSP primer is used in amplification for checking type I nonspecific products, and in the second control, only LFP primer is used in amplification for checking type II nonspecific products. Those bands in the test group, which are consistent with the bands in any control, will be identified to be type I or type II nonspecific bands and will be excluded. In the three examples of this study, PCR walkings into flanking regions did not bring about type I nonspecific products (lane 3 in Fig. 2, lanes 11–15 in Fig. 3, and lanes 3, 6, 9, and 12 in Fig. 4). In other methods of hemispecific PCR walking (e.g., TAIL-PCR and restriction-site PCR), as random primers take part in multiple low stringent cycles, they have sufficient chances to trigger single primer amplification. Therefore, type II products should be considered as the most possible source of nonspecific products. Through the above measures in our method, only small amounts of type II products may appear. In the cases of this study, no nonspecific type II products were found in the walkings for *rpoB* gene of *V. vulnificus* and transposase-like gene from *V. alginolyticus* (lane 2 in Fig. 2, and lane 6–10 in Fig. 3). Though some type II products were generated in the walking for *sto* gene of *V. cholerae* (judging from the bands in lanes 2, 5, 8, and 11 in Fig. 4), they could be easily excluded through abovementioned strategy.

Compared to other PCR walking methods (especially to TAIL-PCR), our method has the following advantages: Firstly, this method is much simpler and time saving. Our method just contains two rounds of PCR (61 cycles). Supposing we hope to get 3-Kb fragment with a 3-min extension time (theoretically, 1-Kb extension per minute), our method at least saves 40% of the time consumed compared with TAIL-PCR. Secondly, the method has high specificity, and there are only a few possibilities for generating nontarget products in the second PCR through a series of above measures. Thirdly, the method can give rise to the production of long DNA fragment (>3.5 Kb) while TAIL-PCR has difficulty in acquiring such long fragment due to the inherent limitation of TAIL-PCR procedure (too many steps in the procedure). Supposing we set the extension time to be at 5 min, the process of TAIL-PCR (hiTAIL-PCR) would become very long, and moreover, it poses a rigorous challenge to the activity of DNA polymerase. Finally, the procedure of our method is simple and relatively fixed, and it does not need laborious exploration for the optimal procedure parameters. In order to get a good positive result, we just need to change the annealing temperature in sole low-stringency step (increase or decrease the matching degree of CLP primers with templates) in the first PCR or increase the design of CLP primers. Moreover, application of the method has no limitation in species as 3′ end sequence of CLP primer can be at random. Although there were 10 CLP primers used, researchers can design new CLP primers according to their special demand. Therefore, the method proposed in this study provides a robust and simple strategy for rapid amplification to obtain long unknown DNA fragments, and greatly decreases the appearance of nontarget products despite using partly random primers.

References

1. Kwon, Y. M., & Ricke, S. C. (2000). Efficient amplification of multiple transposon-flanking sequences. *Journal of Microbiol Methods, 41*, 195–199.
2. Uchiyama, T., & Watanabe, K. (2006). Improved inverse PCR scheme for metagenome walking. *Biotechniques, 41*, 183–188.
3. Morimoto, S., & Fujii, T. (2009). A new approach to retrieve full lengths of functional genes from soil by PCR-DGGE and metagenome walking. *Applied Microbiology and Biotechnology, 83*, 389–396.
4. Ochman, H., Gerber, A. S., & Harlt, D. J. (1988). Genetic applications of an inverse polymersase chain reaction. *Genetics, 120*, 621–623.
5. Huang, G., Zhang, L., & Birch, R. G. (2000). Rapid amplification and cloning of Tn5 flanking fragments by inverse PCR. *Letters in Applied Microbiology, 31*, 149–153.
6. Martin, M. C., Fernandez, M., Linares, D. M., & Alvarez, M. A. (2005). Sequencing, characterization and transcriptional analysis of the histidine decarboxylase operon of *Lactobacillus buchneri*. *Microbiology, 151*, 1219–1228.
7. Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., et al. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Research, 18*, 2887–2890.
8. Rosenthal, A., & Jones, D. C. (1990). Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. *Nucleic Acids Research, 18*, 3095–3096.
9. Lukyanov, K. A., Matz, M. V., Bogdanova, E. A., Gurskaya, N. G., & Lukyanov, S. A. (1996). Molecule by molecule PCR amplification of complex DNA mixtures for direct sequencing: An approach to in vitro cloning. *Nucleic Acids Research, 24*, 2194–2195.
10. Ali, A., Mahmud, Z. H., Morris, G. J., Sozhamannan, S., & Johnson, J. A. (2000). Sequence Analysis of Tn*phoA* insertion sites in *Vibrio cholerae* mutants defective in rugose polysaccharide production. *Infection and Immunity, 68*, 6857–6864.
11. Nthangeni, M. B., Ramagoma, F., Tlou, M. G., & Litthauer, D. (2005). Development of a versatile cassette for directional genome walking using cassette ligation-mediated PCR and its application in the cloning of complete lipolytic genes from *Bacillus* species. *Journal of Microbiol Methods, 61*, 225–234.
12. Sarker, G., Turner, R. T., & Bolander, M. E. (1993). Restriction-site PCR: A direct method of unknown sequence retrieval adjacent to a known locus by using universal primers. *PCR Methods and Applications, 2*, 318–322.
13. Puskas, L. G., Fartmann, B., & Bottka, S. (1994). Restricted PCR: Amplification of an individual sequence flanked by a highly repetitive element from total human DNA. *Nucleic Acids Research, 22*, 3251–3252.
14. Jiang, Y., Pei, J. J., Song, X., & Shao, W. L. (2007). Restriction site-dependent PCR: An efficient technique for fast cloning of new genes of microorganisms. *DNA Research, 14*, 283–290.
15. Liu, Y. G., & Whittler, R. F. (1995). Thermal asymmetric interlaced PCR: Automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genome, 25*, 674–681.
16. Michiels, A., Tucker, M., Ende, W. V. D., & Laere, A. V. (2003). Chromosomal walking of flanking regions from short known sequences in GC-Rich plant genomic DNA. *Plant Molecular Biology Reporter, 21*, 295–302.
17. Motomura, M., Chihaya, N., Shinozawa, T., Hamasaki, T., & Yabe, K. (2003). Cloning and characterization of the O-methyltransferase I gene (*dmtA*) from *Aspergillus parasiticus* associated with the conversions of demethylsterigmatocystin to sterigmatocystin and dihydrodemethylsterigmatocystin to dihydrosterigmatocystin in aflatoxin biosynthesis. *Applied and Environmental Microbiology, 65*, 4987–4994.
18. Patil, P. B., & Sonti, R. V. (2004). Variation suggestive of horizontal gene transfer at a lipopolysaccharide (lps) biosynthetic locus in *Xanthomonas oryzae* pv. *oryzae*, the bacterial leaf blight pathogen of rice. *BMC Microbiology, 4*, 40.
19. Wang, P., Sun, Y., Li, X., Zhang, L., Li, W., & Wang, Y. (2004). Rapid isolation and functional analysis of promoter sequences of the nitrate reductase gene from *Chlorella ellipsoidea*. *Journal of Applied Phycology, 16*, 11–16.
20. Huang, H. Q., Wang, G. Z., Zhao, Y. Y., Shi, P., Luo, H., & Yao, P. (2010). Direct and efficient cloning of full-length genes from environmental DNA by RT-qPCR and modified TAIL-PCR. *Applied Microbiology and Biotechnology, 87*, 1141–1149.
21. Reddy, P. S., Mahanty, S., Kaul, T., Nair, S., Sopory, S. K., & Reddy, M. K. (2008). A high-throughput genome-walking

method and its use for cloning unknown flanking sequences. *Analytical Biochemistry, 381*, 248–253.

22. Satyanarayana, K. V., Chandrashekar, A., & Ravishankar, G. A. (2006). Evaluation of PCR-based methods for isolating flanking regions of genes. *Molecular Biotechnology, 32*, 111–116.

23. Kuromon, T., Hirayama, T., Kiyosue, Y., Takabe, H., Mizukado, S., Sakurai, T., et al. (2004). A collection of 11800 single-copy Ds transposon insertion lines in *Arabidopsis*. *The Plant Journal, 37*, 897–905.

24. Liu, Y. G., & Chen, Y. L. (2007). High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *BioTechniques, 43*, 649–656.

25. Luo, P., & Hu, C. Q. (2008). Analysis on molecular biological characteristic of the gene and its flanking sequences, similar with transposase in *Vibrio cholerae* pathogenicity island, among *V. alginolyticus* strains. *Acta Microbiologica Sinica, 48*, 1367–1372.