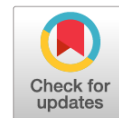


DOI: <https://doi.org/10.17816/ecogen624820>

Review



# PCR-based genome walking methods (review)

Elena S. Okulova<sup>1,2</sup>, Mikhail S. Burlakovskiy<sup>1</sup>, Ludmila A. Lutova<sup>1</sup><sup>1</sup> Saint Petersburg State University, Saint Petersburg, Russia;<sup>2</sup> All-Russian Institute of Plant Protection, Saint Petersburg, Russia

## ABSTRACT

The review discusses a range of classical and modern methods used to determine the nucleotide sequence of unknown DNA regions flanking known ones. These methods are applied to decipher the regulatory regions of genes, identify integration sites of T-DNA or viruses, and so on, in cases where the use of whole-genome sequencing is not justified. To amplify a DNA segment, a binding site for a primer must be added to the end of the unknown sequence. This can be achieved either by ligating an adapter or by annealing a degenerate primer under gentle conditions, or by looping the DNA fragment so that the target region is surrounded by known sequences. The second important task is to eliminate the inevitable products of nonspecific binding of adapters or degenerate primers, which is often resolved through multiple rounds of nested PCR. Different methods vary significantly in terms of complexity, prevalence, and the availability of required reagents.

**Keywords:** genome walking; chromosome walking; PCR; flanking sequences; integration site; T-DNA.

## To cite this article

Okulova ES, Burlakovskiy MS, Lutova LA. PCR-based genome walking methods (review). *Ecological genetics*. 2024;22(1):75–103. DOI: <https://doi.org/10.17816/ecogen624820>

Received: 19.12.2023

Accepted: 09.04.2024

Published online: 14.04.2024

DOI: <https://doi.org/10.17816/ecogen624820>

Научный обзор

# Методы «прогулки по геному» на основе ПЦР (обзор)

Е.С. Окулова<sup>1, 2</sup>, М.С. Бурлаковский<sup>1</sup>, Л.А. Лутова<sup>1</sup><sup>1</sup> Санкт-Петербургский государственный университет, Санкт-Петербург, Россия;<sup>2</sup> Всероссийский научно-исследовательский институт защиты растений, Санкт-Петербург, Россия

## АННОТАЦИЯ

В обзоре рассматривается ряд классических и современных методов, позволяющих установить нуклеотидную последовательность неизвестных участков ДНК, фланкирующих известные. Они применяются для расшифровки регуляторных областей генов, определения сайтов встраивания Т-ДНК или вирусов и т. д. в тех случаях, когда использование полногеномного секвенирования неоправданно. Для амплификации участка ДНК, к концу неизвестной последовательности необходимо добавить участок связывания для праймера; это реализуется либо путем лигирования адаптера, либо посадкой вырожденного праймера в мягких условиях, либо закольцовыванием фрагмента ДНК, чтобы изучаемый участок оказался окружен известными последовательностями. Второй важной задачей является избавление от неизбежно возникающих продуктов неспецифического связывания адаптеров либо вырожденных праймеров — чаще всего данная проблема разрешается несколькими раундами вложенной ПЦР. Разные методы существенно отличаются по трудоемкости, распространенности и доступности необходимых реактивов.

**Ключевые слова:** прогулка по геному; прогулка по хромосоме; ПЦР; фланкирующие последовательности; сайт встраивания; сайт интеграции; Т-ДНК.

## Как цитировать

Окулова Е.С., Бурлаковский М.С., Лутова Л.А. Методы «прогулки по геному» на основе ПЦР (обзор) // Экологическая генетика. 2024. Т. 22. № 1. С. 75–103. DOI: <https://doi.org/10.17816/ecogen624820>

## INTRODUCTION

The key technology of modern life sciences is undoubtedly sequencing, which enables the sequencing of biopolymers such as genomic DNA. The crown of its development, whole-genome sequencing, as the name suggests, provides comprehensive information on the genome of an organism. However, the use of such “heavy artillery” is unreasonable for many tasks due to the cost, limited accessibility, and redundancy of the output data. Most routine queries in genetics and molecular biology can still be resolved using cheap and readily available Sanger sequencing, which can process DNA fragments of up to 1,000 nucleotide pairs per read.

However, with Sanger sequencing, the challenge of providing a target DNA fragment, preferably amplified, falls on the researcher. The creation and screening of genomic DNA libraries is different in labor intensity from preparation for whole-genome sequencing; PCR-based methods are simpler, faster, and less resource-demanding. PCR requires forward and reverse primers complementary to both ends of the DNA fragment of interest. Thus, the search for mutations in already known sequences is easy, but screening for mutations in unexplored regions of DNA, even bordering the known ones, faces the “chicken and egg” problem; amplification requires a primer, which can be designed only if the sequence of the required region is known. This applies to a wide range of molecular genetics tasks related to the analysis of unknown DNA sequences, such as deciphering regulatory regions of genes; amplification of variable sequences surrounding conserved regions of genes; determination of T-DNA, virus, or transposon integration sites; filling gaps in full-genome sequencing [1], and metagenomic analysis [2, 3].

In such cases of limited genetic information, a group of methods, collectively called genome walking (GW), are used to amplify an unknown region of DNA bordering a known sequence. Therefore, primers should bind to an unknown DNA region. The classical “inverted PCR” method involves cutting DNA with restriction enzymes before self-ligating the fragments into ring structures. Thus, an unknown flanking DNA region is surrounded by a known sequence and can be amplified using a pair of primers specific to it (arranged in the opposite orientation to the classical one).

In other variants of the GW method, a short known linker/adaptor/cassette sequence is attached to an unknown DNA strand. This is accomplished either by restriction ligation or by PCR using random and semirandom primers, the 3'-end of which binds to genomic DNA

under mild annealing conditions and the 5'-end carries an adaptor sequence. For each further PCR step, a gene-specific primer and a primer complementary to the adaptor are used. Thus, all GW methods are divided into the following categories: inverted PCR, ligation-mediated PCR, and PCR with random primers [4]. The authors have added an “Other methods” section describing approaches to solving the problem of amplifying an unknown DNA sequence that are not categorized into the three proposed groups.

A common problem in GW is the need to essentially amplify a single hybrid DNA molecule arising from a fortunate event, which is usually resolved by several rounds of nested PCR. This is particularly evident in inverted PCR, in which specific conditions are required for self-ligation, including a low concentration of starting DNA. The use of random and semirandom primers inevitably generates a significant level of “noise.” In addition to the target PCR product (fragment from the gene-specific primer to the adaptor primer), an overwhelming number of byproducts (results of “counter” landing of adaptor primers on genomic DNA) are produced. Each genome walk technique includes a unique method of suppressing nonspecific amplification. Thus, the corresponding primers usually cannot be directly annealed on the initial adaptor; the binding site is only on the complement DNA strand that is synthesized by the polymerase starting from the gene-specific primer. In other embodiments, the adaptor forms a hair-pin structure that can function as a primer; the resulting DNA, due to the long self-complementary site, becomes virtually inert and does not participate in subsequent amplification reactions. To further increase the amplification efficiency of the target fragment (and decrease that of the byproducts), the annealing temperature of gene-specific primers should be higher than that of adaptor primers.

Due to the probabilistic nature of GW techniques, the success of an experiment depends on many factors, and much time is spent adapting the protocol to the subject at hand. The concentration of matrix DNA (excess can be as harmful as deficiency) and its quality (short primer-like DNA degradation products significantly increase the amount of PCR byproducts) have a significant impact. The length of the resulting target product is important for sequencing; in methods involving restriction, the length depends on the location of the enzyme recognition sites in the genome. If several products of various lengths appear during the reaction, the shortest product is preferentially amplified. PCR methods with random primers allow the size of the resulting fragments to be adjusted within certain limits. The annealing temperature

in the low-hardness/strictness/precision cycle is key when adapter primers are planted onto genomic DNA. Increasing the temperature decreases the probability of a given event and simultaneously increases the average fragment length; decreasing the temperature results in shorter fragments and increases noise. The temperature is easiest to adjust for optimal results.

## INVERSE PCR

The first method to identify an unknown DNA sequence from a known region *in vitro* was the **inverse PCR method (1988)**. This approach involves DNA cleavage by a restriction enzyme that does not have a cut site in the region to be integrated. The resulting fragments are circularized by ligation under special conditions, including a low concentration of DNA. The desired site was amplified using differently directed primers complementary to the ends of the known sequence. In this case, flanking primers to the plant region of the genome were not required (Figure 1).

The main advantage of the inverted PCR method is its high specificity because no adapters or degenerate primers are used. However, there is a lower probability of circularizing the desired fragment during this approach than that during cross-ligation. The technique also has limitations associated with the nonuniform distribution of restriction sites [5, 6].

## LIGATION-MEDIATED PCR

Several methods have been described for ligation-mediated PCR, and all involve the following steps:

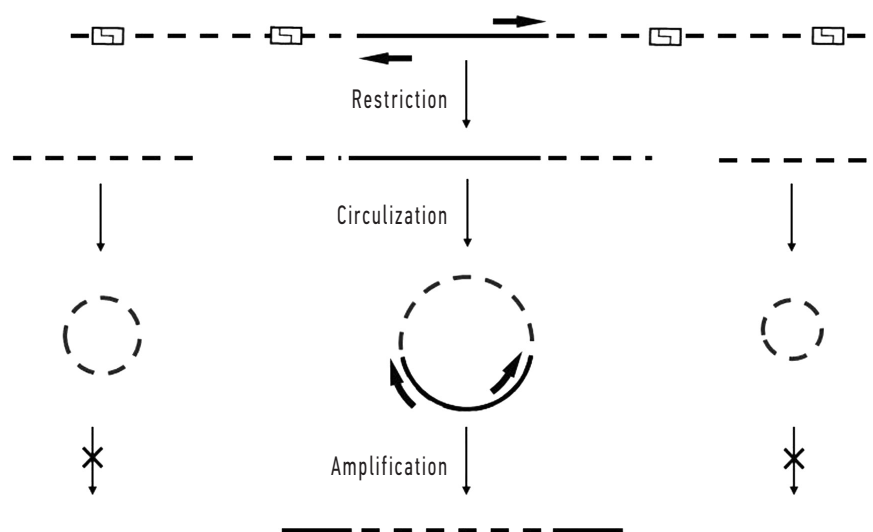
1. restriction cleavage of DNA;
2. ligation of small nucleotide sequences (adapter/linker/cassette) containing the annealing site of the adapter primer to the ends of the resulting fragments, and
3. amplification of the border region using primers specific to the target DNA site and the ligated fragment (Figure 2).

The differences between the individual methods lie in the structure of the ligation fragment and how false positives are excluded. Let us consider some techniques from this group.

Ligation-mediated PCR is extremely sensitive to the quality of the matrix DNA and requires the isolation of a high-quality product [7].

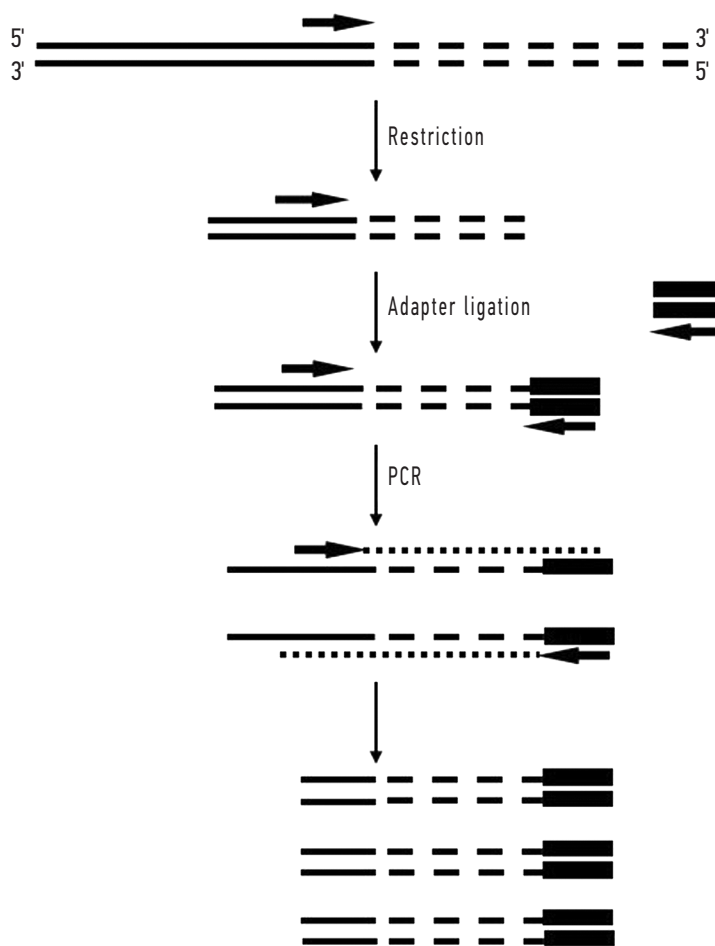
**Vectorette PCR, first used in 1990**, rapidly isolates terminal sequences from yeast artificial chromosome clones [8]. Vectorette PCR allows the amplification of DNA sequences that lie between a known primer and the nearest restriction site.

DNA is cleaved by a restriction enzyme to form a sticky 5'-end. A vectorette linker is ligated to the 5'-end. The target fragment is amplified by PCR using a primer specific to the target DNA and a primer specific to the vectorette. The vectorette cassettes



**Fig. 1.** Schematic representation of inverted PCR (based on E.K. Hui et al. [6]). Solid line, known DNA sequence; dashed line, unknown DNA segment; arrow, primer binding site; white rectangles, restriction sites. DNA is cleaved by a restriction enzyme that does not have a cutting site within the insert, then circularized under conditions favorable for the formation of monomeric circles and amplified. In PCR, primers complementary to the ends of the insert fragment are used in opposite directions

**Рис. 1.** Схематическое изображение принципа инвертированной ПЦР (на основе Е.К. Хуй и соавт. [6]). Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; стрелка — сайт посадки праймера; белые прямоугольники — сайты рестрикции. ДНК расщепляют рестриктазой, которая не имеет сайта разрезания в интегрируемом участке, закольцовывают в условиях, благоприятных для образования мономерных колец и амплифицируют. При ПЦР используют разнонаправленные праймеры, комплементарные концам интегрируемого фрагмента



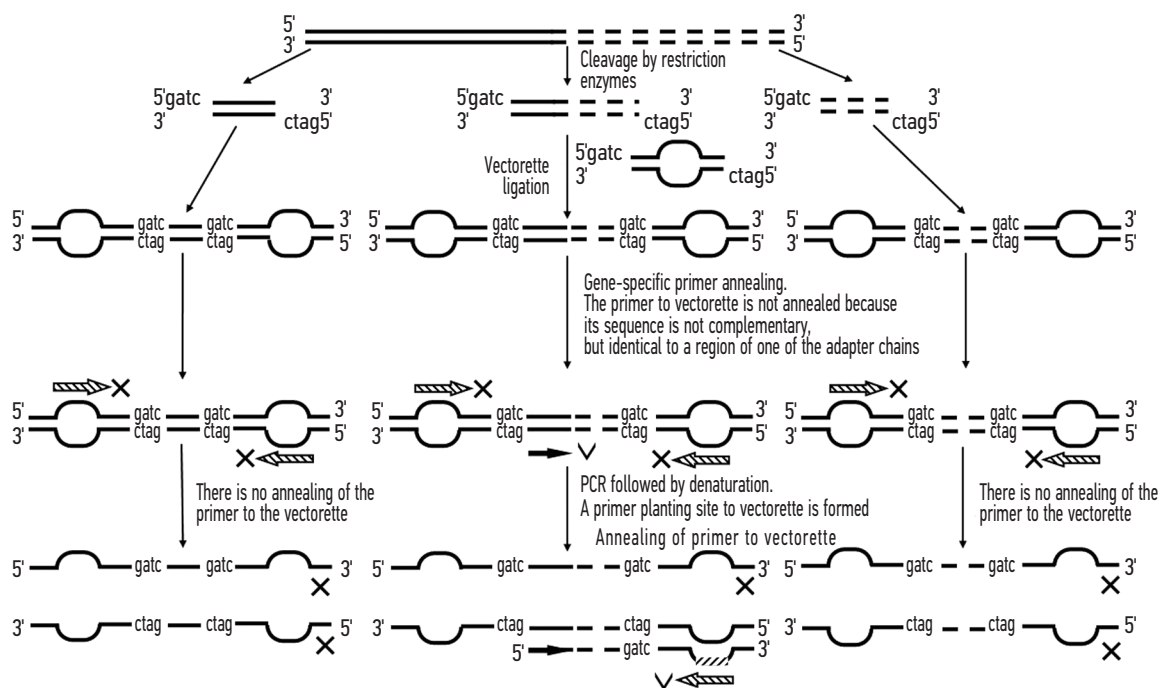
**Fig. 2.** Schematic representation of PCR mediated by ligation. Solid line, known DNA sequence; dashed line, unknown DNA segment; dotted line, amplification product; small arrow, primer binding site; black rectangles, adapters

**Рис. 2.** Схематическое изображение принципа ПЦР, опосредованной лигированием. Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; пунктирная линия — продукт амплификации; маленькая стрелка — сайт посадки праймера; черные прямоугольники — адаптер

comprise a double-stranded sequence with a central noncomplementary region and a sticky end suitable for ligation of DNA cleaved by restriction enzymes. The primer to the vectorette used in PCR has the same sequence as the mismatched portion of one of the strands; therefore, it cannot anneal and initiate elongation until its complementary strand is synthesized by the polymerase from the specific target DNA primer (Figure 3) [8].

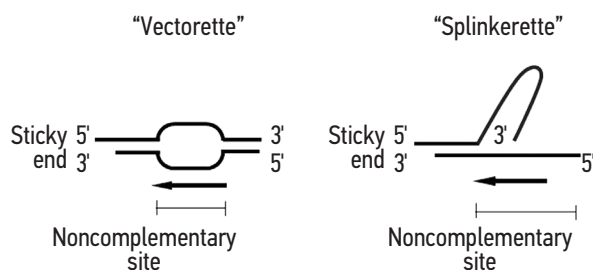
The main disadvantage of this technique is its low specificity: byproducts may be formed owing to the non-specific annealing of any primer and the end-repair reaction. This is caused by leaving unligated cassettes and restriction products with sticky 5'-ends in the solution. These ends are completed during the first cycle of the PCR reaction, and after a subsequent denaturation step, they anneal to each other, forming sufficiently stable structures that function as primers.

The mechanism of **Splinkerette PCR (1995)** is similar to that of vectorette PCR, but it allows the elimination of amplification byproducts. Instead of the central mismatched region of the vectorette DNA cassette, the splinkerette cassette includes a structure with a “hairpin” on one of the strands (Figure 4). The primer has a sequence similar to the noncomplementary region of the adapter opposite the “hairpin”; therefore, as in the case of the vectorette, it cannot anneal until the complementary strand has been synthesized. In the PCR reaction, the hairpin structure works as a primer and the polymerase starts elongation along the lower chain. The resulting “giant hairpin” is stable and functionally excluded from further reaction. Although the hairpin structure “straightens out” during denaturation, when the temperature drops, it regains its hairpin shape faster than primers can bind to it because the complementary sequences are long and have a melting point higher than that of any primer.



**Fig. 3.** Schematic representation of the “vectorette PCR” (based on E.K. Hui et al. [6]). Solid line, known DNA sequence; dashed line, unknown DNA segment; hatched arrow, primer binding site to the “vectorette”; hatched segment, DNA fragment complementary to the “vectorette” primer; black arrow, primer binding site to the target DNA. DNA is cleaved by a restriction enzyme, generating a 5'-sticky end. Then, a synthetic oligonucleotide (linker) called “vectorette” is ligated to the 5'-end. PCR amplification of the DNA fragment is performed using an internal primer specific to the target DNA and a primer specific to the “vectorette”

**Рис. 3.** Схематическое изображение принципа «vectorette PCR» (на основе Е.К. Хуй и соавт. [6]). Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; заштрихованная стрелка — сайт посадки праймера к «vectorette»; заштрихованный участок — фрагмент ДНК, комплементарный праймеру к «vectorette»; черная стрелка — сайт посадки праймера к ДНК-мишени. ДНК расщепляется рестриктазой с образованием липкого 5'-конца. Затем к 5'-концу лигируется синтетический олигонуклеотид (линкер), называемый «vectorette». ПЦР-амплификацию фрагмента ДНК проводят с использованием внутреннего праймера, специфического для ДНК-мишени, и праймера, специфического для «vectorette»



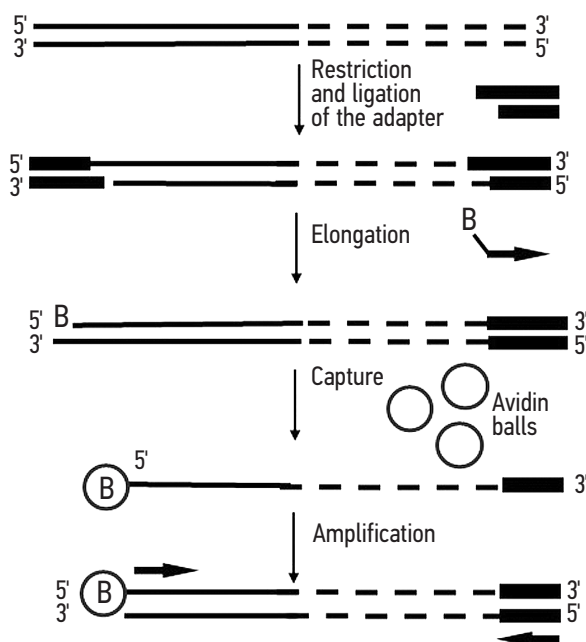
**Fig. 4.** Schematic representation of the structures of «vectorette» and «splinkerette» cassettes (based on E.K. Hui et al. [6])

**Рис. 4.** Схематическое изображение структур кассет «vectorette» и «splinkerett» (на основе Е.К. Хуй и соавт. [6])

Hence, it cannot be a “seed” for binding the ends together. Additionally, in the splinkerette system, only one of the chains can act as a nonspecific primer, while in the vectorette system, both the upstream and downstream chains can cause mis-elongation.

The advantages of splinkerette over vectorette are particularly evident in the amplification of larger fragments, wherein the formation of the target product may be hampered by high competition from artifacts arising from end-repair [5, 9].

**Capture PCR (CPCR; 1991)** also involves the use of adapters ligated to the sticky ends of genomic DNA cleaved by restriction enzymes, with the ligated constructs not restoring the restriction site. The oligonucleotide adapters lack 5'-phosphate groups, which ensures that only one oligonucleotide is covalently attached to the ends of genomic DNA fragments. The 25-nucleotide-long adapter can be used together with one of three different short complementary oligonucleotides to modify the ends generated by >30 restriction enzymes. Short



**Fig. 5.** Schematic representation of the Capture PCR (CPCR) (based on M. Lagerstrom et al. [10]). Solid line, genomic DNA; black rectangles, adapters; arrow, primer binding site; B, biotin. The first strand is synthesized using a single gene-specific biotinylated primer, enabling the fixation of this fragment on a streptavidin-coated substrate. Unlabeled DNA is removed during washing. The target fragment is then amplified with a primer to the adapter and a second specific primer

**Рис. 5.** Схематическое изображение принципа ПЦР с захватом (CPCR) (на основе M. Lagerstrom и соавт. [10]). Сплошная линия — геномная ДНК; черные прямоугольники — адаптер; стрелка — сайт посадки праймера; B — биотин. Первая цепь синтезируется на основе одного генспецифического биотинированного праймера, что позволяет зафиксировать этот фрагмент на покрытой стрептавидином подложке. Немеченная ДНК удаляется в ходе промывки. Целевой фрагмент амплифицируют с праймером к адаптору и вторым специфическим праймером

oligonucleotides have a low GC content, which ensures that they do not function as primers during PCR.

The DNA strand is elongated using a 5'-biotinylated primer complementary to a known DNA sequence. Thus, a new biotin-labeled chain was synthesized, which allows further fixation of this fragment on a streptavidin-coated solid substrate. The resulting elongation sites include the DNA fragment of interest and end with a sequence complementary to the added linker. Biotin-labeled fragments are fixed on a streptavidin-coated solid substrate. After washing, PCR is performed using the sequence bound to the carrier as the matrix. A second specific primer, complementary to the region downstream of the biotinylated oligonucleotide, was used as a primer for PCR together with the linker oligonucleotide. The resulting PCR product contains the DNA region of interest (Figure 5).

The convenience of CPCR is considerably enhanced by the use of a special substrate with streptavidin-coated magnetic beads placed in the individual wells of a titration microplate. The procedure allows simultaneous isolation of fragments from many DNA samples and minimizes the risk of contamination between reactions. This technique is very specific. Additionally, no cloning procedure is required when using solid-phase sequencing [10].

Although restriction and ligation reactions are performed simultaneously, the process can be complicated by the need for streptavidin pellets, two additional rounds of PCR, and a matrix purification procedure before sequencing, as not every laboratory practice works well with streptavidin and biotin [11].

**Extension Primer Tag Selection PCR (EPTS/LM-PCR; 2001)** is essentially an improved capture PCR technique. It is unique in that it allows the exclusion of nontarget DNA fragments using special biotinylated primers in the step preceding solid-phase PCR.

The first stage of EPTS/LM-PCR involves preparing small double-stranded fragments using restriction enzymes and forming blunt ends by amplification using a reverse biotinylated primer. Biotinylated DNA is concentrated on a spin column, and excess primers are removed. Biotinylated fragments were immobilized on magnetic beads to remove nontarget DNA and reaction components and fixed to the wall of the tube using a magnetic particle concentrator (processor).

In the second stage, solid-phase LM-PCR is performed as follows: the oligonucleotide cassette is ligated along the blunt ends of biotinylated fragments, and ligation is stopped by washing. Using subsequent alkaline

denaturation and exposure to a magnetic particle concentrator, free single-stranded nonbiotinylated DNA is separated. These fragments are exponentially PCR-amplified. The first amplification uses an external cassette primer and a primer to a known DNA sequence, while the second uses an internal cassette primer and another gene-specific primer, increasing the specificity of the method. Thus, fragments are generated that include a border with an unknown DNA sequence (Appendix 1, doi: 10.17816/ecogen624820-4207491) [12].

**The Panhandle PCR method (1992)** involves the formation of a ring structure of DNA with a long end ("panhandle") containing a boundary of known and unknown sequences (Appendix 2, doi: 10.17816/ecogen624820-4207492). The matrix is created by cleaving genomic DNA with restriction enzymes before ligation of a single-stranded oligonucleotide. The resulting fragments are dephosphorylated with alkaline phosphatase to prevent self-ligation. The phosphorylated oligonucleotide is then ligated to the 3'-end. This single-stranded oligonucleotide has two features: its 5'-end is complementary to the single-stranded ends of genomic DNA cleaved by restriction enzymes and its sequence is homologous to part of the integrated fragment.

Denaturation and intrachain annealing are performed to form a circular structure with a long end. An unknown sequence appears in the structure of the circle. A fragment containing an unknown site is amplified using two pairs of nested primers [13].

**The Boomerang PCR method (1995)** is named after manner in which the polymerase starts elongating a chain from a specific primer annealing site, forms a loop around another chain, and eventually returns to the original fragment on the DNA to form a new landing site for the same primer. The loop is formed by adapters designed with self-complementary ends and a noncomplementary middle part. This method allows the use of different variants of adapters that, depending on the restriction enzyme, can be designed to ligate to both sticky and blunt ends of cleaved DNA (Appendix 3, doi: 10.17816/ecogen624820-4207493). The amplification of target fragments by PCR leads to the generation of a specific product containing a boundary of known and unknown DNA sequence, which can then be cloned and sequenced [14].

However, this method is susceptible to producing many false positives, as each case of primer binding to a nonspecific site generates a fragment capable of amplification on par with the target fragment. This issue can be addressed by employing several rounds of nested PCR with multiple sequenced primers.

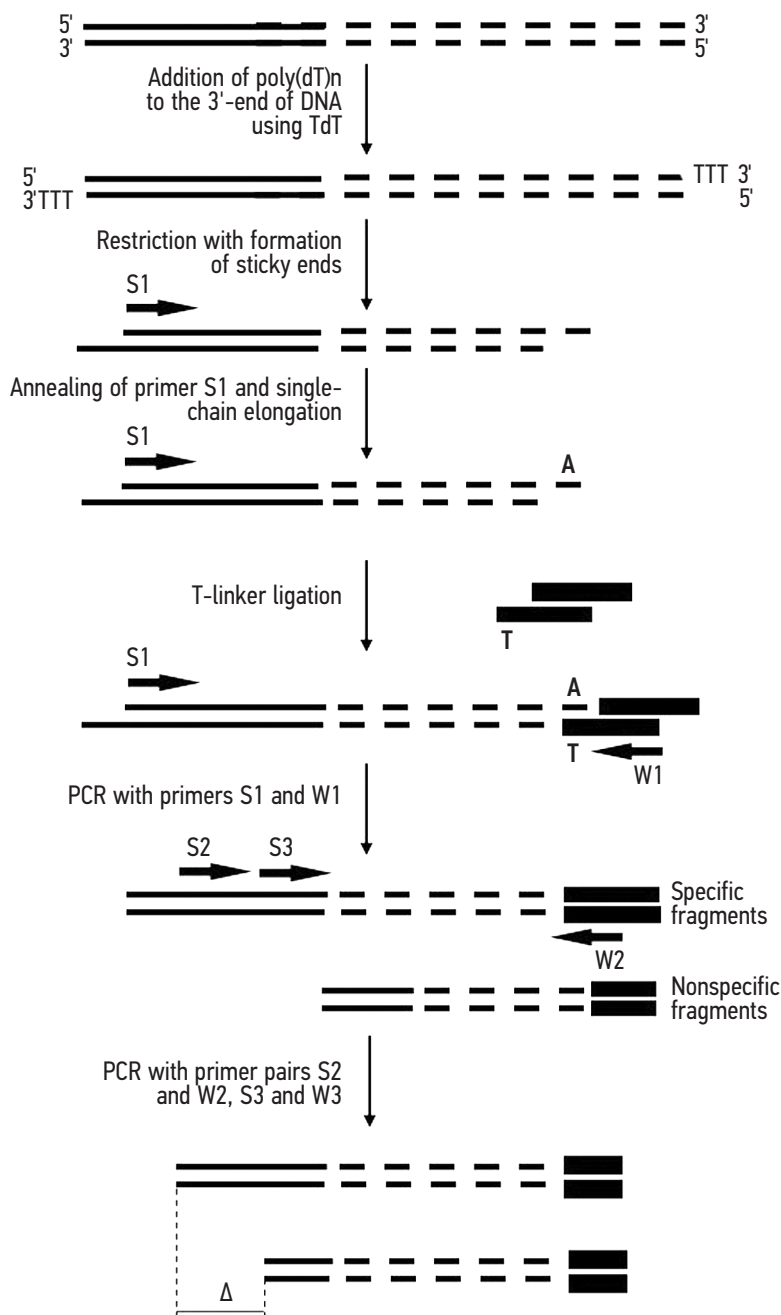
**T-linker-specific ligation PCR (T-linker PCR; 2003)** involves PCR with T-linker ligation to walk chromosomes or genes. In the first step, poly(dT)<sub>n</sub> is added to the 3'-ends of genomic DNA molecules using terminal deoxynucleotidyl transferase (TdT). This addition is necessary to prevent nonspecific binding to adapters later on. DNA with poly(dT)<sub>n</sub>-sequenced DNA is cleaved with restriction enzymes to form 3'-protruding ends without A. The second step involves chain elongation of the target molecule using specific primer S1 and formation of a "tail" with A at the 3'-end using TaqDNA polymerase. Using T4 DNA ligase, a T-linker is ligated at the 3'-end of the target molecule. This is followed by two rounds of nested PCR to amplify the target molecule; the first round uses the external primer pair S1 and W1, while the second round uses primer pairs S2-W2 and S3-W2. Specifically amplified molecules are identified based on the length differences ( $\Delta$  S2-S3) between the S2-W2 and S3-W2 primer products (Figure 6) [15].

**Adapter ligation PCR (2007)** involves the addition of a special double-stranded adapter with long and short arms to the sticky ends of restriction fragments. The amino group at the 3'-end of the short arm prevents amplification of the DNA chain during PCR. The sequence of the short arm of the adapter does not contain the annealing site of the adapter primer; however, the longer arm contains a 22-nucleotide sequence that matches it exactly. The second primer corresponds to a known DNA sequence; from it, a complementary strand containing the landing site of the adapter-specific primer is synthesized. This produces fragments with the two necessary primer annealing sites (Appendix 4, doi: 10.17816/ecogen624820-4207494) [16].

Another noteworthy method is **PCR with restriction site elongation (RSE-PCR; 2010)**. In its application, it does not involve the ligation of any sequence; instead, an adapter primer is used, which allows lengthening of the restriction site (Appendix 5, doi: 10.17816/ecogen624820-4207495) [17].

When using **template blocking PCR (2010)** to reduce nonspecific amplification, the 3'-ends of genomic DNA fragments cleaved by restriction enzymes are blocked by dideoxynucleoside triphosphate (ddNTP) and ligated with properly designed cassettes without a phosphate group at the 5'-end. This approach prevents the amplification of a fragment that does not contain the target sequence resulting from restriction and cassette ligation. The modified cassette-flanked genomic DNA fragments are used as a matrix for amplifying the target gene with a gene-specific primer and a cassette primer (Appendix 6, doi: 10.17816/ecogen624820-4207496) [18].





**Fig. 6.** Schematic representation of the T-linker PCR (based on Y. Yuanxin et al. [15]). Solid line, known DNA sequence; dashed line, unknown DNA segment; arrow, primer binding site; black rectangles, linker; S1, S2, and S3, specific primers binding to the known sequence of the target molecule; W1 and W2, walking primers binding to the T-linker sequence; A, A-“tail” of the target molecule; T, T-nucleotide of the T-linker;  $\Delta$ , presumed difference in amplification products with specific primers S2 and S3 in separate reactions of the second cycle

**Рис. 6.** Схематическое изображение принципа T-linker ПЦР (на основе Y. Yuanxin и соавт. [15]). Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; стрелка — сайт посадки праймера; черные прямоугольники — линкер; S1, S2 и S3 — специфические праймеры, связывающиеся с известной последовательностью молекулы-мишени; W1 и W2 — шагающие праймеры, связывающиеся с последовательностью T-линкера; A — A-«хвост» молекулы-мишени; T — T-нуклеотид T-линкера;  $\Delta$  — предполагаемая разница в продуктах амплификации со специфическими праймерами S2 и S3 в разделенных реакциях второго цикла

**Single specific primer PCR (SSP-PCR; 1989)**, also known as unidirectional GW, can also be included in the current group of methods. Although a cloning vector is used instead of linkers in this case, this method is

fundamentally similar to adapter ligation-mediated PCR methods.

To amplify a fragment containing the boundary of known and unknown DNA sequences, information about

only a small section of DNA is needed to select a gene-specific primer.

The method includes the following steps:

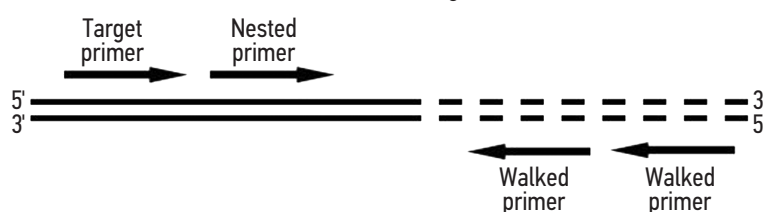
- restriction of chromosomal DNA by one or a combination of enzymes;
- ligation of fragments into any cloning vector,
- and amplification of the specific ligated fragment using one primer specific to a known region and a second primer that hybridizes with the vector.

Enough fragments are ligated to the vector DNA, and the vector-specific primer will hybridize with all fragments. A gene-specific primer is used to select target products, allowing exponential accumulation of the desired fragments containing the primer's annealing site (Appendix 7, doi: 10.17816/ecogen624820-4207497). In contrast to standard methods that work with genomic libraries, in SSP-PCR, PCR is performed soon after transformation without separating bacterial clones into separate cultures carrying individual copies of plasmids, which saves time.

SSP-PCR allows a random combination of restriction enzymes, enabling the amplification of DNA for which no restriction site information is available [19].

## PCR WITH RANDOM PRIMERS

The group of PCR techniques that use random primers involves the use of sets of nonspecific primers that can randomly sit on a DNA molecule during PCR. This category does not require complex DNA manipulation before or after PCR. Typically, these methods involve alternating cycles of high and low annealing temperatures. Primary PCR uses a cycle with mild conditions to increase the likelihood of planting "walking" primers on an unknown DNA sequence. This is followed by 2–3 rounds of nested PCR under stringent conditions for efficient annealing of site-specific primers. These cycles are performed to exponentially amplify the target fragments and eliminate false results (Figure 7). The differences in the methods of this group lie in the structure of the "walking" primers and in the number and manner of alternating cycles under different annealing conditions.



**Fig. 7.** The general scheme of primer placement in PCR with random primers. Solid line, known DNA sequence; dashed line, unknown DNA segment; arrows, primer binding sites

**Рис. 7.** Общая схема расположения праймеров при ПЦР со случайными праймерами. Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; стрелки — сайты посадки праймеров

Some methods involve self-ligation of the investigated DNA sequence and formation of circular structures (e.g., UFW method). The main limitation of these methods is the excessive accumulation of nontarget DNA products because of nonspecific annealing of a random primer [1].

**Targeted gene walking PCR (1991)** was one of the first techniques to be developed. This is based on the observation that a primer can initiate elongation on either an unknown or a specific target sequence that has only partial homology with the 3'-end.

This approach uses three types of primers: "target" primers, which hybridize with a specific known target sequence; "internal"  $^{32}\text{P}$ -labeled detection primers, which are located a short distance inward relative to the "target" PCR primers; and "walking" primers, for hybridization with unknown sequences.

The general protocol for "PCR with genome walking" comprises three consecutive steps:

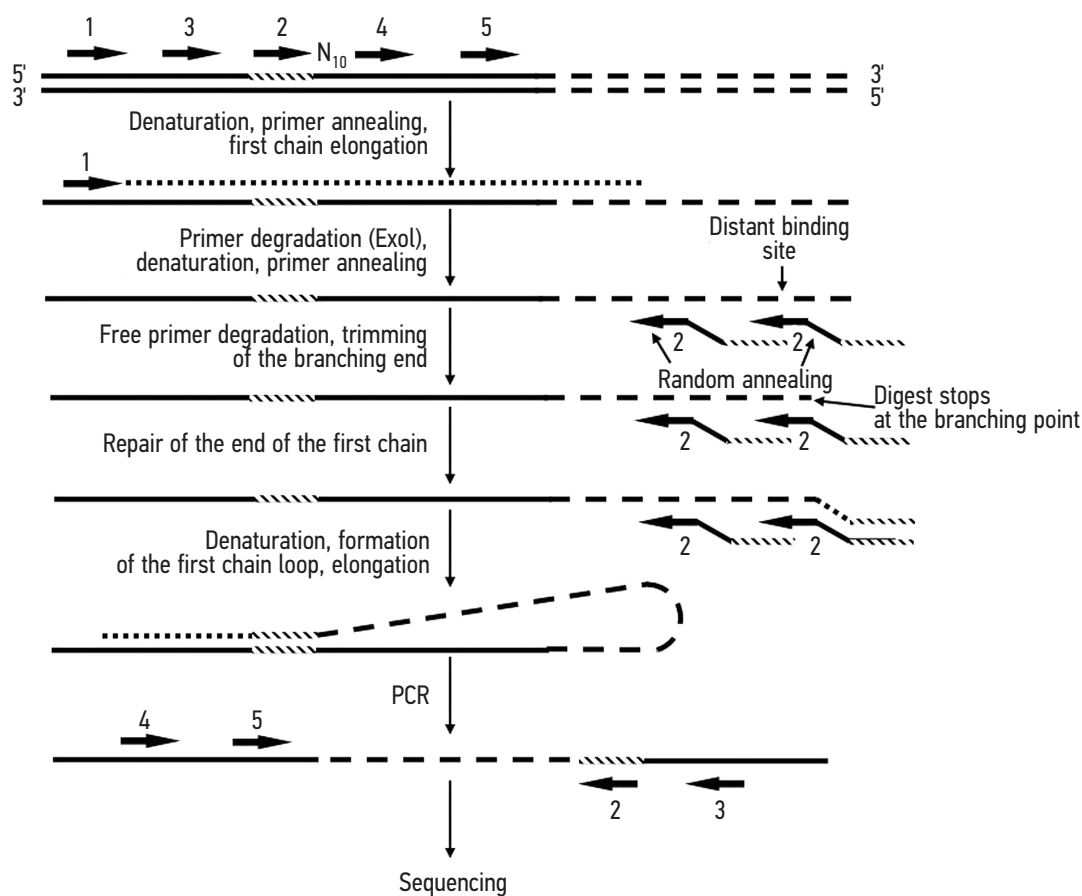
1) a series of PCR reactions with identical components in each tube (including the "target" primer) except for different "walk" primers;

2) an aliquot from each PCR reaction is used to select target fragments using nested internal  $^{32}\text{P}$ -labeled primers and primers to a known sequence. This procedure identifies fragments that contain sections of "target" DNA, and

3) the labeled band is excised from the gel, reamplified, and directly sequenced.

To increase the frequency of positive results, a whole series of "walking" primers (at least 20) is commonly used. Although reactions with "walking" primers are performed in parallel, the process itself can be quite labor-intensive [20].

**Restriction site PCR (RSPCR; 1993)** uses sequences of oligonucleotides (RSOs) that are specific to the region of DNA containing the restriction recognition site by restrictionases instead of "walking" primers (Appendix 8, doi: 10.17816/ecogen624820-4207498). Restriction site sequences occur in all organisms and are repeated frequently enough that, theoretically, there will always be a restriction site (starting at any site) within the PCR range. However, the use of RSOs could pose a problem



**Fig. 8.** Schematic representation of the UFW method (based on K.W. Myrick and W.M. Gelbart [21]). Solid line, known DNA sequence; dashed line, unknown DNA segment; dotted line, amplification product; short arrows with numbers, UFW primers, numbered in the order of use  
**Рис. 8.** Схематическое изображение принципа метода UFW (на основе K.W. Myrick и W.M. Gelbart [21]). Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; пунктирная линия — продукт амплификации; короткие стрелки с числами — праймеры UFW, пронумерованы в порядке использования

because they are not unique sequences and are not sufficiently specific for PCR. Nested PCR is performed to increase specificity. Additionally, the product of this PCR is sequenced by genomic amplification with transcript sequencing (GAWTS) using another internal primer to a known sequence [11].

The peculiarity of the **universal fast walking (UFW; 2002)** method involves a 5–6 h series of reactions performed in a single tube in a single thermocycler program (Figure 8).

The method begins with the synthesis of the first chain, followed by the destruction of primer #1 by an exonuclease. The strands are then denatured and annealed with primer #2, which contains a random site at the 3'- and 5'-ends that has a complementary segment to a known DNA sequence.

The second exonuclease simultaneously destroys the unbound primer and digests the first chain at the starting point of the random 3'-end (at the "branching" point) of the furthest bound primer #2. Sequence formation at

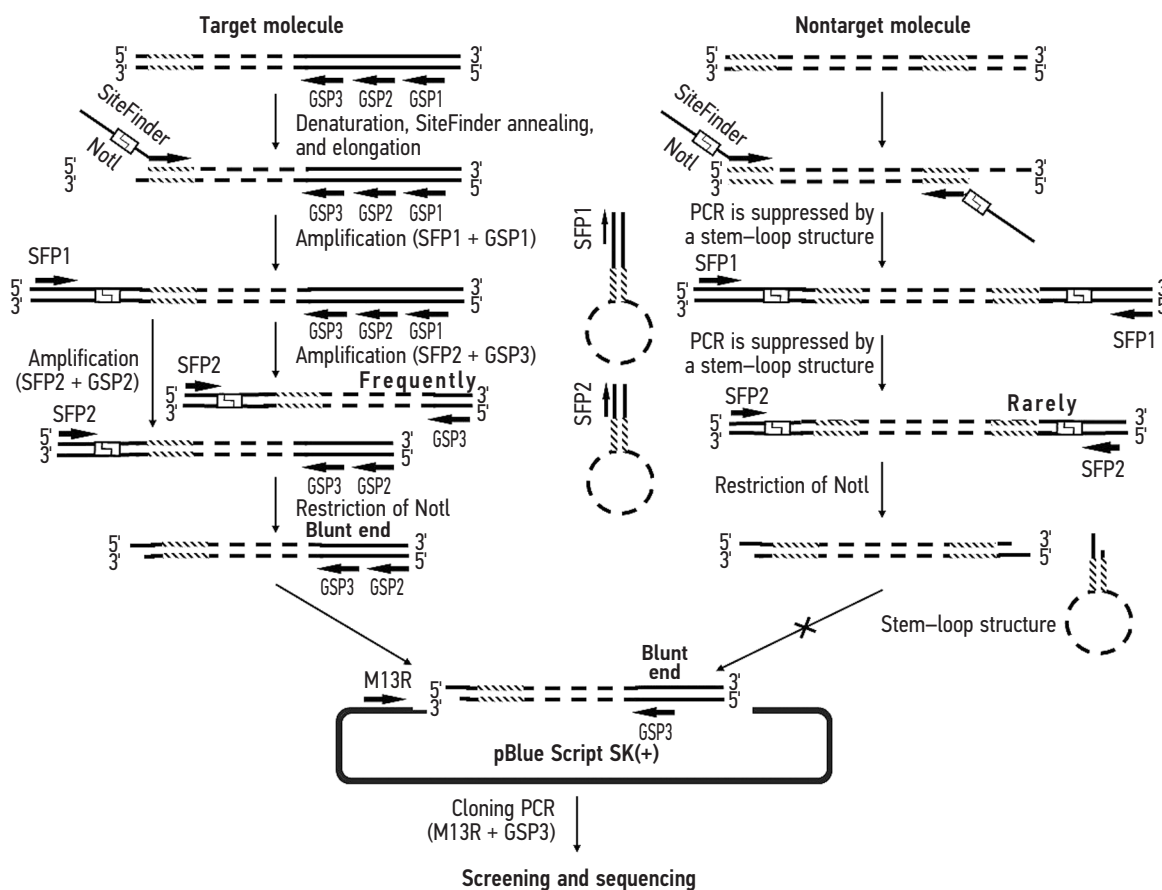
the 3'-end of the first chain occurs at the expense of the 5' (nonrandom) part of primer #2. The chains are denatured again.

The first chain forms a loop by in-chain annealing between the complementary annealing site of primer #2 and its copy at the other end of the sequence. Next comes the elongation of this chain after the loop along itself, producing a product containing known sequences at the edges and an unknown region in the center.

Nested primers #3 and #4, whose annealing sites are located on both sides of the annealing site of primer #2, allow obtaining a specific amplicon containing the boundaries of known and unknown sequences [21].

The range of action of this method is directly related to polymerase capabilities.

**The SiteFinding method (Site Finding-PCR; 2005)** uses "false" SiteFinder primers with a known sequence of four nucleotides at the 3'-end, containing a rare restriction site for NotI in PCR, which facilitates cloning with commonly used vectors: 5'-...GCGGCCGCGCNNNNNNNNNGCCT3'



**Fig. 9.** Schematic representation of the SiteFinding-PCR method principle (based on G. Tan et al. [22]). Solid line: known DNA sequence; dashed line: unknown DNA segment; arrow: primer binding site; white rectangle: restriction site; GSP: gene-specific primers, SFP: SiteFinding primers

**Рис. 9.** Схематическое изображение принципа метода ПЦР SiteFinding (на основе G. Тан и соавт. [22]). Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; стрелка — сайт посадки праймера; белый прямоугольник — сайт рестрикции; GSP — генспецифические праймеры, SFP — праймеры SiteFinding

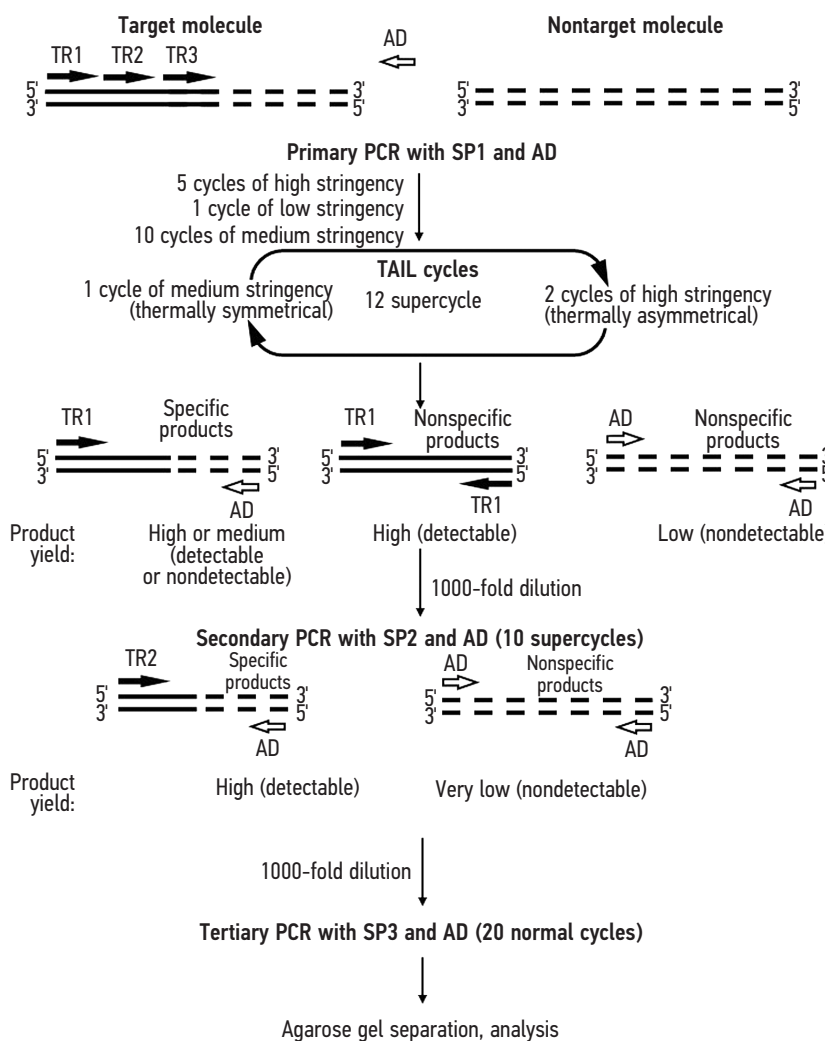
and 5'-...GCGGCCGCGCNNNNNNNNNGCGC3'. A nested primer (SFP1 and SFP2) and three gene-specific primers (GSP) are also used (Figure 9).

First, the SiteFinding reaction is performed with low-temperature annealing to form the restriction site NotI. In a desired embodiment, this reaction produces a fragment comprising the boundary of known and unknown DNA with NotI restriction sites and the primer landing sites SFP1 and GSP1. Second, nested PCR follows: the target DNA is exponentially amplified using nested PCR with GSP and primers SFP1 and SFP2. As the primer planting process is random, it is possible to form untargeted double-stranded products containing SFP planting sites at both ends. Further amplification is suppressed due to the formation of stem-loop structures due to the presence of inverted end repeats (this sequence does not further participate in the reaction). Third, the SFP-GSP amplification product undergoes NotI cleavage and is purified by electrophoretic separation on an agarose gel.

This results in target molecules with one sticky (in the restriction site) and one blunt end, facilitating the ligation of such fragments into the linearized pBluescript SK(+) vector. Restriction products with two sticky ends or in a stem-loop structure cannot be introduced into the vector.

Fourth, this fragment is cloned. An internal specific primer (GSP3) is used to screen clones, ensuring that only those clones containing the specific product are selected because only target molecules have a complementary site for GSP3 [22].

**The thermal asymmetric interlaced PCR method (TAIL PCR; 1995)** is widely used. It involves the use of a set of nested sequence-specific primers (TR1, TR2, and TR3) along with short arbitrary degenerate (AD) primers (15–16 bps) with a low melting point and varying degrees of degeneracy. By alternating the annealing temperature between high (62°C–68°C) and low (44°C), the relative amplification efficiency of specific and nonspecific products can be thermally controlled.



**Fig. 10.** Schematic representation of the TAIL-PCR method (based on Y.-G. Liu et al. [23]). Solid line, known DNA sequence; dashed line, unknown DNA segment; arrow, primer binding site; TR1, TR2, TR3, nested primers complementary to the known sequence; AD, short arbitrary degenerate primers (15–16 bp) with low melting temperature and varying degrees of degeneracy. Alternating annealing temperatures from high (62 to 68 °C) in high stringency cycles to low (44 °C) in low stringency cycles thermally controls the relative efficiency of amplification of specific and nonspecific products

**Рис. 10.** Схематическое изображение принципа метода TAIL-PCR (на основе Y.-G. Liu и соавт. [23]). Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; стрелка — сайт посадки праймера; TR1, TR2, TR3 — вложенные праймеры, комплементарные известной последовательности; AD — короткие произвольные вырожденные праймеры (15–16 п. н.) с низкой температурой плавления и различной степенью вырожденности. Чередую температуру отжига от высокой (62 до 68 °C) в циклах высокой точности до низкой (44 °C) в циклах низкой точности, термически контролируется относительная эффективность амплификации специфических и неспецифических продуктов

TAIL-PCR reactions are performed in sequential order: primary, secondary, and tertiary, in which the product of one reaction is used as a matrix for the next reaction (Figure 10).

The primary reaction is performed in six reaction tubes containing genomic DNA, primers complementary to a known sequence, and various degenerate primers. The primary PCR product is used as a matrix for the secondary reaction. In the secondary, similar components are used except that primer TR1 is replaced with primer TR2, which is located further away from the 5'-end. The tertiary reaction uses the PCR product of the secondary reaction

as a matrix and a third nested primer (TR3) with the same six AD primers used in the primary and secondary reactions.

Thus, TAIL-PCR requires 12 PCR reactions to identify a relatively small site. False-positive results are also possible due to amplification artifacts during PCR [23, 24].

**Fusion primer and nested integrated PCR (FPNI-PCR; 2011)** involves the use of two sets of primers: 1) primers specific to a known DNA sequence and 2) fusion primers (composite primers, FAD, and FP), one segment of which is a degenerate sequence (AD) and the other is complementary to the primers of the

following steps. The combination of fusion primers of a known sequence and a degenerate sequence distinguishes this method from TAIL-PCR (Appendix 9, doi: 10.17816/ecogen624820-4207506).

In the first step, a large volume of a mixture of DNA, a gene-specific primer (SP1) designed for a region of the genome with known sequence, and a combination of nine AD fusion primers is prepared. Amplifications involve 3–6 repetitions of two cycles with harsh annealing conditions followed by a cycle with milder conditions. Theoretically, single-stranded PCR products from a gene-specific primer are generated during cycles with high annealing temperatures, while double-stranded products using FP primers (FP) are obtained during cycles with milder conditions (the melting point of FP primers is significantly lower than SP). At this stage, many nonspecific products may be formed.

In the second and third steps, nested PCR is performed using target-specific primers (SP2 and SP3, respectively) and FP-specific primers (FSP1 and FSP2, respectively). These steps represent PCR with a high annealing temperature to promote selective amplification of target sequences. Additionally, the large hairpin formed in some nonspecific products also contributes in preventing their further amplification (PCR suppression).

Thus, nonspecific products obtained in the first step of FPNI-PCR are not amplified in the second and third steps and are significantly diluted in the final mixture [25].

**Partially overlapping primer-based PCR (POP-PCR; 2015)** is another GW technique. It uses a set of relatively long POP primers that partially overlap at the 3'-end: POP1 for primary PCR, POP2 for secondary PCR, and POP3 for tertiary PCR. The primers are randomly sequenced and contain identical 3'-ends of 10 bp and heterologous 5'-ends of 15 bp in length. This partially overlapping design ensures that POP primers anneal to each other's complementary sites only at relatively low temperatures. Nested primers are used as GSP. A total of three rounds of PCR (primary, secondary, and tertiary) are performed. Each round comprises three annealing stages: stage 1, five highly harsh cycles (at 65°C); stage 2, one lowly harsh cycle (at 25°C) and moderately harsh cycle (at 50°C); and stage 3, 30 highly harsh cycles (at 65°C). Each subsequent round uses the products of the previous round as the matrix. The products of the last round of PCR contain the most specific fragments that include the boundary of known and unknown DNA sequences (Figure 11).

POP-PCR is a fairly efficient method but it requires numerous degenerate "walking" primers, which complicates the experimental operations [26].

**Stepwise partially overlapping primer-based PCR (SWPOP-PCR; 2018)** is an improved version of POP-PCR method. The main disadvantage of POP-PCR is the need to use a separate POP primer in each round of PCR, which complicates the experiment and leads to amplification errors.

The key to SWPOP-PCR is the development of an improved set of overlapping primers in which the 3'-end (10 bp) of the subsequent SWPOP primer is identical to the 5'-end of the previous primer. Therefore, annealing between the SWPOP primer and its partially complementary site (the previous SWPOP site) occurs only at relatively low temperatures (Appendix 10, doi: 10.17816/ecogen624820-4207507) [4].

**Levano-Garcia et al. (2005)** proposed a method for mapping transposon integration sites by touchdown PCR using a pair of primers, one of which is a hybrid consensus-degenerated oligonucleotide and the other is a sequence-specific primer that anneals only from one of the chains of the inserted marker gene. The sequence of the gene-specific primer is designed to allow annealing of the 3'-end at the known DNA site occurs close to the target unknown genomic sequence, preferably 40–60 bp away from the boundary.

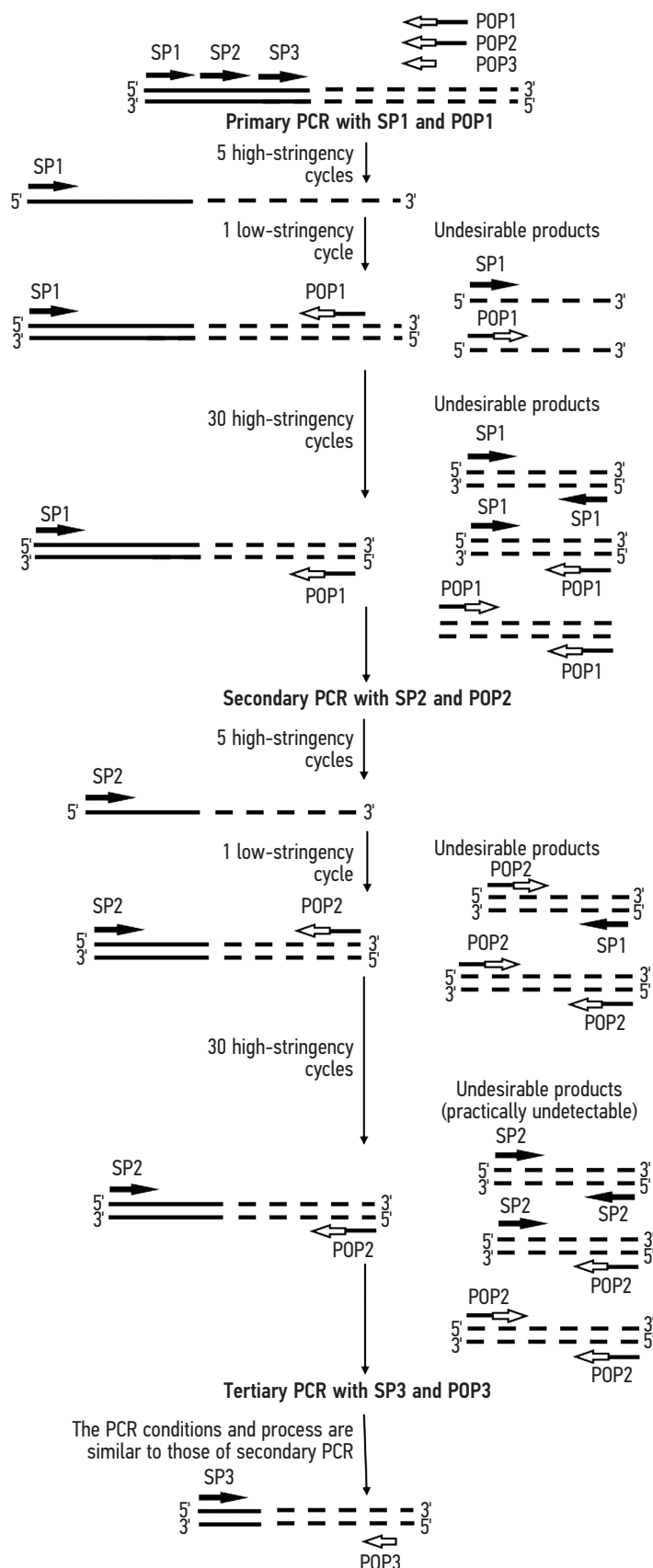
Hybrid primers are designed according to the CODEHOP method [27] and can be any 25–43-mer oligonucleotide having a nondegenerated consensus sequence at the 5'-end (13–31 bp) followed by a 10-bp sequence with degenerate bases at various positions within this segment, followed by 2 bp at the 3'-end that are nondegenerated [28].

**The palindromic sequence targeted PCR (PST-PCR; 2021) method** involves targeting walking primers to palindromic sequences arbitrarily present in natural DNA matrices.

PST-PCR involves two rounds of PCR. The first round uses a combination of a single sequence-specific primer (SSP) and a semirandom primer targeting a palindromic sequence (PST). The second round uses a combination of one or two universal primers: one is annealed with a 5'-tail attached to a SSP, and the other is annealed with another 5'-tail attached to a PST primer (Appendix 11, doi: 10.17816/ecogen624820-4207508).

The key advantage of PST-PCR is the convenience of using a single universal primer with unchanged sequence in GW processes using various matrices [7].

**A feature of PCR with single-long primer and randomly-amplified polymorphic DNA primer PCR (SLRA PCR; 2019)** is the use of a single-long (30–35 bp) SLP primer with an annealing site on a known sequence. SLP-PCR allows amplification of a known region of the



**Fig. 11.** Schematic representation of the POP-PCR (based on H. Li et al. [26]). Solid line, known DNA sequence; dashed line, unknown DNA segment; arrow, primer annealing site

**Рис. 11.** Схематическое изображение принципа метода POP-PCR (на основе H. Li и соавт. [26]). Сплошная линия — известная последовательность ДНК; штриховая линия — неизвестный участок ДНК; стрелка — сайт посадки праймера

genome toward an unknown sequence boundary. A series of nested PCRs with three gene-specific GSP primers (26–28 bps) and randomly amplifying polymorphic DNA with RAPD primers (10 bp) are used to screen out nonspecific fragments. The novelty of the approach lies in the use of long primers and using the same high temperature for annealing and elongation during nested PCRs (Appendix 12, doi: 10.17816/ecogen624820-4207509) [29].

**Fusion primer-driven racket structure formation PCR (Fusion primer-driven racket PCR, FPR-PCR; 2022)** uses a trifunctional fusion primer fused to two sequence-specific fragments generated during primary PCR. A trifunctional primer mediates walking, selective amplification, and intrachain annealing of the target DNA.

FPR-PCR involves two rounds of PCR (primary and secondary). Primary PCR begins with five moderately harsh cycles (at 55°C) that allow the 3'-end of the fusion primer (FP) to hybridize with only its annealing site in the known region (SSP1) to increase the number of copies of the first chain. A subsequent single mildly harsh cycle (at 25°C) helps the FP to partially anneal in the unknown region of this chain and elongate toward the known region to produce the target second chain, forming its inverted repeat. The target DNA is then exponentially amplified in the subsequent harsh (at 65°C) cycles. Partially single-stranded DNA undergoes intra-stranded annealing between the SSP3 site and its inverted repeat. This results in a racket-shaped structure in which the unknown single-chain region is bounded by a known double-chain "handle."

Secondary PCR is a nested PCR with primers SSP2 and SSP4 for exponential amplification of the target DNA. Any untargeted DNA cannot be amplified due to lack of an ideal binding site with these two primers. Target DNA becomes the main product of the reactions (Appendix 13, doi: 10.17816/ecogen624820-4207510) [1].

**Wristwatch PCR (2022)** and POP-PCR use a set of partially overlapping WWP primers. Their difference lies in the structure of the oligonucleotides used: the wristwatch primers have 5'- and 3'-overlapping regions and a heterologous interval, thus, annealing produces a structure resembling the shape of a wristwatch or a bubble. Annealing of primers in this way is possible at sufficiently low temperatures (40°C). WWP primers also have a high melting point (60°C–65°C) and uniform distribution of the four bases (A, T, C, and G). The sequence of each WWP is randomized.

Wristwatch PCR consists of three nested (primary, secondary, and tertiary) PCRs, each using a different

WWP primer and gene-specific GSP primers to amplify the boundary of known and unknown DNA sequence (Appendix 14, doi: 10.17816/ecogen624820-4207511) [30].

## OTHER METHODS

Methods that use an approach different from the previous two were included in this group.

In 1999, a restriction-independent method for cloning segments of genomic DNA beyond known sequences was proposed that includes, as a first step, the elongation of a gene-specific primer, a method commonly used in RNA studies. This method represents the first step of cDNA synthesis and is commonly known as "5'-rapid amplification of cDNA ends" (5'-RACE) [31].

The method does not depend on cutting or mapping with restriction enzymes. It relies on the ability of terminal transferase to attach a chain of cytosines to the 3'-end of DNA. In the initial step, single-stranded DNA covering the flanking region is created by linear amplification with a single primer in a known region. A homo-oligomeric cytosine "tail" is added using terminal transferase. The elongated fragments are then amplified by PCR with a nested gene-specific primer in a known region. A homo-oligomeric polyguanine primer complementary to the cytosine tail in the unknown region is used as the reverse. It contains an "ATAT" sequence at the 5'-end, which allows further ligase-independent incorporation of the obtained DNA molecules into the T-vector by TA-cloning [32] and sequencing (Appendix 15, doi: 10.17816/ecogen624820-4207512). This method was first successfully applied to the bacterial genome [33]. Later, it was used to adapt the 5'-RACE-based approach to mapping the boundary of known and unknown sequence for DNA of eukaryotic organisms [34, 35].

The method of rolling circle amplification of genomic templates for inverse PCR (RCA-GIP; 2010) is based on creating circular fragments of genomic DNA with subsequent amplification using DNA polymerase  $\phi$ 29, characterized by high processivity without the need for linker sequences.

The DNA is cleaved by various restriction enzymes in different tubes, then the fragments are joined by T4 ligase to produce circular DNA. The circular DNA is then amplified using DNA polymerase  $\phi$ 29 and hexamer primers. The primers are annealed on the matrix and elongated. When polymerase  $\phi$ 29 encounters a double-stranded stretch of DNA, it displaces the second strand and continues elongation. The substituted newly synthesized site serves as a planting site for new primers and becomes the matrix. A structure resembling a branching



tree is formed, in which synthesis occurs on each branch. This produces many linear concatemers suitable for the inverse PCR matrix that can be easily amplified, sequenced, or cloned, allowing simultaneous mapping of the 3'- and 5'-unknown ends of a virtually unlimited number of genomic sequences (Figure 12) [36].

4SEE (2020) is based on next-generation sequencing and allows identification of the location of numerous transgenic insertions in the genome in a single application and to characterize the complex chromosomal rearrangements associated with this event. Usually, full-genome sequencing does not consider situations where heterologous insertions compromise the integrity of both the plant DNA and the integrating fragment itself, which makes identification of the insertion site difficult. The 4SEE approach relies on the chromosome conformation capture (3C) method [37], which uses only molecular methods to map chromosome contact zones and uses the principle of ligation of DNA molecules converged in space.

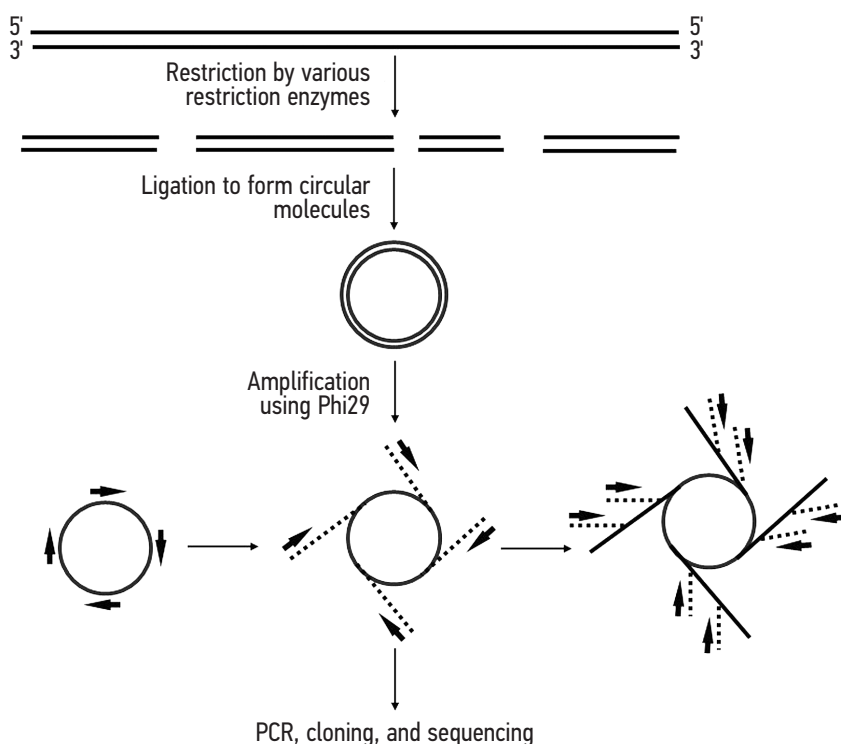
The first step involves treating the cells with formaldehyde to preserve the native structure of the nucleus. This fixes protein–protein, protein–DNA, DNA–DNA, and other interactions through the formation of covalent bonds between closely related organic molecules.

In the second step, DNA fragmented after crosslink unfolding using restriction enzyme treatment, followed by ligation of DNA fragments under high dilution conditions to increase the likelihood of autoligation of the molecules.

DNA ligation products are further processed to form ring structures by a second restriction cleavage. Unknown DNA bound to a known sequence is amplified using inverted PCR before it is sequenced. Sequencing results are compared with DNA bibliographies, and the results are subjected to bioinformatics analysis. The polymeric properties of the genome provide a high frequency of contacts between linearly neighboring sequences. Thus, by calculating the relative enrichment of a library of specific genome regions ligated to each other, we can infer the probability of interactions between these regions in the three-dimensional core space (Figure 13) [38].

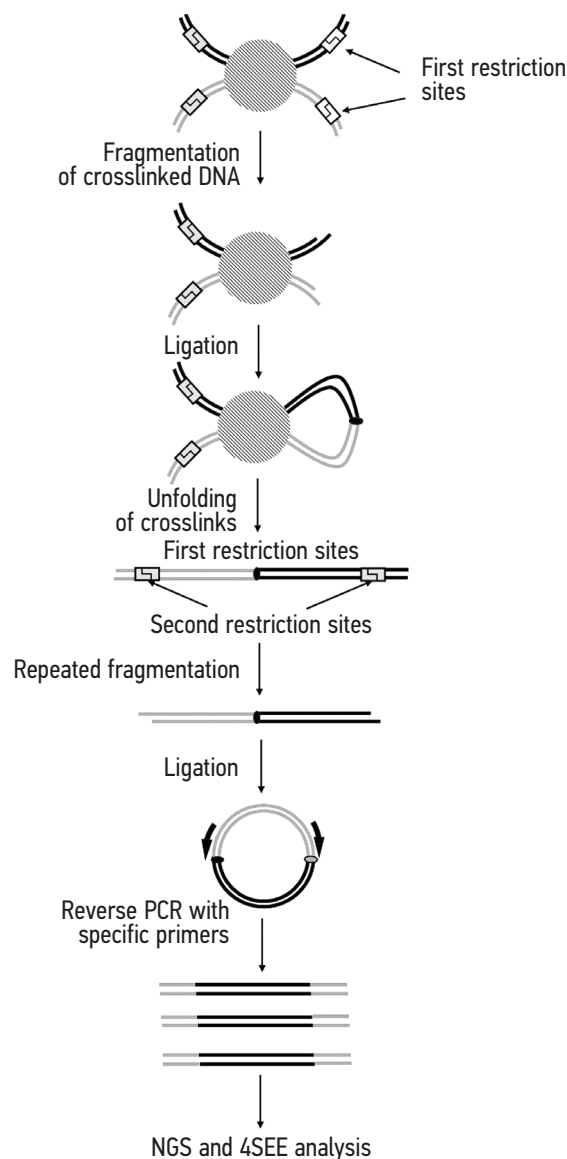
## CONCLUSION

Several GW methods based on different mechanisms have been proposed over the past decades. All the methods are functional; thus, a researcher wishing to decipher an unknown DNA sequence flanking a known sequence can choose a method based on the available reagents and skills.



**Fig. 12.** Schematic representation of the RCA-GIP method (based on A. Tsaftaris et al. [36]). Black line, genomic DNA; arrows, random hexameric primers; dotted lines, copies of concatemers

**Рис. 12.** Схематическое изображение принципа RCA-GIP (на основе А. Tsaftaris и соавт. [36]). Черная линия — геномная ДНК; стрелки — случайные гексамерные праймеры; пунктирные линии — копии конкатемеров



**Fig. 13.** Schematic representation of the 4SEE principle. Shaded circle, DNA crosslinking region with formaldehyde; gray line, known DNA sequence; black line, unknown DNA segment; white and gray rectangles: restriction sites

**Рис. 13.** Схематическое изображение принципа 4SEE. Заштрихованный круг — область сшивки ДНК формальдегидом; серая линия — последовательность известной ДНК; черная линия — неизвестный участок ДНК; белые и серые прямоугольники — сайты рестрикции

Theoretically, an ideal GW system should include only commonly used methods (e.g., PCR) without modifications (e.g., adding reagents during PCR), use only “normal” reagents (standard polymerases, primers up to 50 bps long without radioactive and other tags); be held in one phase, and have samples immediately suitable for sequencing. None of the currently existing systems fully satisfies all the requirements, but many modern techniques do. For example, PCR-based methods require only primers (this is one of the cheapest and most readily available reagents) and a thermocycling amplifier, the basic equipment of a genetic laboratory. The only disadvantage is the need to perform multiple stages of nested PCR with different primers.

Methods that increase labor intensity include sequence ligation procedures and working with reagents that are rarely used (e.g., biotin and streptavidin). However, an advantage of using primers with biotin is that they allow immediate use of the product in solid-phase sequencing.

Owing to the probabilistic nature of GW methods, researchers usually need to adapt the chosen protocol to their subject. The most significant factors here are the quality and concentration of matrix DNA and, in the case of PCR with random primers, the annealing temperature at the low-hardness stage of the cycle.

Although these methods are now increasingly being replaced by full-genome sequencing, many of them remain relevant due to their economic feasibility.

**Table.** Review of methods enabling amplification of an unknown DNA sequence flanking a known one  
**Таблица.** Обзор методов, позволяющих амплифицировать неизвестную последовательность ДНК, фланкирующую известную

Name of Method	Year	Method content	Targeting	Exclusion of nonspecific fragments	Labor intensity										
					restriction	special sequence ligation	ligation into a cloning vector	special primer structure	biotin or isotope tag	washing or cleaning step	looping phase	special reagents	labor-intensive PCR or sequencing procedure		
Inverse PCR [5]	1988	Restriction and circularizing of fragments	Amplification of the desired site using multi-directional primers complementary to the ends of the known sequence	Highly specific	+	-	-	-	-	-	+	-	-	-	-
Vectorette PCR [8]	1990	Restriction and ligation of the linker	Use of primers complementary to regions of known DNA sequence and cassette	Specific structure of linker (bubble) and primer to it	+	+	-	-	-	-	-	-	-	-	-
Splinkerette PCR [9]	1995	Restriction and ligation of the linker	Use of primers complementary to regions of known DNA sequence and cassette	Specific structure of linker (hairpin) and primer to it	+	+	-	-	-	-	-	-	-	-	-
Capture PCR [10]	1991	Linker restriction and ligation and the use of biotin-tagged primers	Using a primer complementary to a region of a known sequence	Use of biotinylated primers	+	+	-	-	+	+	-	-	-	-	-
EPTS/LM-PCR [12]	2001	Restriction and amplification with reverse biotinylated primers, ligation of oligonucleotide cassette, and nested solid-phase PCR	Amplification of restriction products using reverse biotinylated primers	Two-step nested PCR with external and internal cassette primers	+	+	-	-	+	+	-	-	-	-	-

Table (continued) / Продолжение таблицы

Name of Method	Year	Method content	Targeting	Exclusion of nonspecific fragments	Labor intensity									
					restriction	special sequence ligation	ligation into a cloning vector	special primer structure	biotin or isotope tag	washing or cleaning step	looping phase	special reagents	labor-intensive PCR or sequencing procedure	
"Panhandle" PCR [13]	1992	Restriction and ligation of a single-stranded phosphorylated oligonucleotide to create a circular structure	Using a primer complementary to a region of a known sequence	Dephosphorylation of restriction fragments by alkaline phosphatase	+	+	-	-	-	-	+	-	-	-
"Boomerang" PCR [14]	1995	Restriction and ligation of adapters looping two complementary DNA strands	Using a primer complementary to a region of a known sequence	Once a new chain has been created around the adapter sequence, elongation of this chain over the complementary matrix proceeds, and the primer binding site is synthesized again	+	+	-	-	-	-	-	-	-	-
T-linker PCR [15]	2003	PCR with T-linker ligation to walk chromosomes or genes	Use of primers complementary to regions of known DNA sequence and the T-linker	Two-step nested PCR	+	+	-	-	-	-	-	+	-	-
PCR with adapter ligation [16]	2007	Restriction and ligation of long- and short arm double-stranded adapters	Applying a primer to a known sequence and a primer corresponding to the long sequence of the long arm of the adapter	Blocking of the 3'-end of the short arm with an amino group	+	+	-	-	-	-	-	-	-	-

Table (continued) / Продолжение таблицы

Name of Method	Year	Method content	Targeting	Exclusion of nonspecific fragments	Labor intensity											
					restriction	special sequence ligation	ligation into a cloning vector	special primer structure	biotin or isotope tag	washing or cleaning step	looping phase	special reagents	labor-intensive PCR or sequencing procedure			
RSE-PCR [17]	2010	Restriction and the use of an adaptor primer allowing restriction site elongation	Use of primers complementary to a region of known sequence and an extended restriction site	Two-step nested PCR	+	-	-	-	+	-	-	-	-	-	-	
Template-blocking PCR [18]	2010	Restriction and ligation of the cassette	Use of primers complementary to regions of known DNA sequence and cassette	Blocking the 3'-ends of genomic DNA fragments cleaved by restriction enzymes with dideoxynucleoside triphosphate	+	+	-	-	-	-	-	-	-	+	-	
SSP-PCR [19]	1989	Restriction and ligation of restriction products into vector followed by amplification	Use of primers complementary to known sequence and vector regions	Use of gene-specific primer	+	-	+	-	-	-	-	-	-	-	-	-
PCR with targeted genome walk [20]	1991	Use of three types of primers for PCR: "target," internal, <sup>32</sup> P-tagged, and walking primers. Reamplification of the tagged fragment	Use of a targeted primer that hybridizes with a specific known target sequence	Nested PCR using internal <sup>32</sup> P-tagged primers	-	-	-	-	+	-	-	-	-	-	-	-

Table (continued) / Продолжение таблицы

Name of Method	Year	Method content	Targeting	Exclusion of nonspecific fragments	Labor intensity																
					restriction	special sequence ligation	ligation into a cloning vector	special primer structure	biotin or isotope tag	washing or cleaning step	looping phase	special reagents	labor-intensive PCR or sequencing procedure								
RSPCR [11]	1993	Use of primers that are specific to the section of DNA that contains the restriction site recognized by restriction enzymes	Use of primers complementary to known sequence regions and specific to restriction sites recognized by restriction enzymes	Nested PCR, sequencing by genomic amplification with transcript sequencing (GAWTS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
UFW [21]	2002	A 5-6 h series of reactions performed in a single tube on a single thermocycler program using exonucleases that degrade primers	Use of a semirandom primer that contains a random site at the 3'-end and the 5'-end with a complementary fragment to a known DNA sequence. Loop formation due to intrachain annealing of the nonrandom portion of the primer. Use of nested primers to obtain specific amplicons	Destruction of extra primers by exonucleases	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-
SiteFinding-PCR [22]	2005	PCR using "false" primers with a known sequence of 4 bp at the 3'-end.	A low-temperature ligation reaction is performed, resulting in the possible amplification of the target site	Cleavage of "target" sites by restriction enzyme and ligation into a vector with further cloning	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-

Table (continued) / Продолжение таблицы

Name of Method	Year	Method content	Targeting	Exclusion of nonspecific fragments	Labor intensity										
					restriction	special sequence ligation	ligation into a cloning vector	special primer structure	biotin or isotope tag	washing or cleaning step	looping phase	special reagents	labor-intensive PCR or sequencing procedure		
TAIL-PCR [23]	1995	Using a set of nested primers specific to a known sequence along with short arbitrary degenerate primers with low melting points and varying degrees of degeneracy	Use of primers complementary to regions of known DNA sequence and the T-linker and degenerated primers	By alternating the annealing temperature, the relative efficiency of amplification of specific and nonspecific products can be controlled	-	-	-	+	-	-	-	-	-	+	
FPNI-PCR [25]	2011	PCR with fusion primers, one site representing a degenerate sequence and the other site corresponding to a specific sequence of the genome	Use of primers complementary to regions of known DNA sequence and fusion primers	Nested PCR	-	-	-	+	-	-	-	-	-	-	+
POP-PCR [26]	2015	PCR with partially overlapping primers at the 3'-end	Use of primers complementary to regions of known DNA sequence and partially overlapping primers	Three-step nested PCR	-	-	-	+	-	-	-	-	-	-	-

Table (continued) / Продолжение таблицы

Name of Method	Year	Method content	Targeting	Exclusion of nonspecific fragments	Labor intensity										
					restriction	special sequence ligation	ligation into a cloning vector	special primer structure	biotin or isotope tag	washing or cleaning step	looping phase	special reagents	labor-intensive PCR or sequencing procedure		
SWPOP-PCR [4]	2018	PCR with partially overlapping primers at the 3'-end	Use of primers complementary to regions of known DNA sequence and partially overlapping primers in which the 3'-end of the subsequent SWPOP primer is identical to the 5'-end of the previous primer	Three-step nested PCR	-	-	-	+	-	-	-	-	-	-	-
Levano-Garcia method [28]	2005	Touchdown PCR with a hybrid consensus-degenerated primer	Use of gene-specific and hybrid consensus-degenerated primers	Touchdown PCR	-	-	-	+	-	-	-	-	-	-	+
PST-PCR [7]	2021	PCR with targeting of walking primers to palindromic sequences	Use of gene-specific and semirandom primers targeting a palindromic sequence	PCR using primers annealed with the 5'-tails of primers from the previous PCR	-	-	-	+	-	-	-	-	-	-	-
SLRA PCR [29]	2019	PCR using a single-long (30–35 bps) primer with an annealing site on a known sequence	Using a single-long primer with an annealing site on a known genome region of the unknown sequence boundary	Nested PCR series with gene-specific and RAPD primers	-	-	-	+	-	-	-	-	-	-	-



Table (continued) / Окончание таблицы

Name of Method	Year	Method content	Targeting	Exclusion of nonspecific fragments	Labor intensity													
					restriction	special sequence ligation	ligation into a cloning vector	special primer structure	blotin or isotope tag	washing or cleaning step	looping phase	special reagents	labor-intensive PCR or sequencing procedure					
FPR-PCR [1]	2022	PCR using a trifunctional fusion primer with the formation of a racket-like structure	A specific DNA fragment enables intrachain annealing of the fusion primer to form a racket-like structure	Nested PCR	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Wristwatch PCR [30]	2022	PCR with partially overlapping primers at the 5'- and 3'-ends with a noncomplementary interval in the middle, which creates a wristwatch-like structure upon annealing	Use of primers complementary to regions of known DNA sequence and partially overlapping primers	Three-step nested PCR	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5'-RACE based method [33]	1999	Amplification of single-chain fragments with the addition of a homo-oligomeric cytosine tail	Application of a primer to a known sequence along with a polyguanine primer	Nested PCR	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RCA-GIP [36]	2010	Restriction and circularizing of fragments followed by amplification using DNA polymerase $\phi 29$ and random hexamer primers	Amplification using DNA polymerase $\phi 29$ and primers, formation of several linear concatemers	Use of polymerase $\phi 29$ , which is characterized by increased processivity and a low error rate	+	-	-	-	-	-	-	-	-	-	-	-	-	-
4SEE [38]	2020	Relying on the chromosome conformation capture method (3C)	Next-generation sequencing		-	-	-	-	-	-	-	-	-	-	-	-	-	-

## ADDITIONAL INFORMATION

**Authors' contribution.** All authors contributed to the development of the article's concept, read, and approved the final version before publication. Personal contribution of the authors: L.A. Lutova — problem statement, drawing conclusions; E.S. Okulova — literature analysis, writing the main text, tabular and graphical presentation of results; M.S. Burlakovskiy — assistance in describing methods.

**Funding source.** This study was not supported by any external sources of funding.

**Competing interests.** The authors declare that they have no competing interests.

### Additional materials.

Supplement 1. Schematic representation of the EPTS/LM-PCR method principle. doi: 10.17816/ecogen624820-4207491

Supplement 2. Schematic representation of the "Panhandle"-PCR method principle. doi: 10.17816/ecogen624820-4207492

Supplement 3. Schematic representation of the "boomerang"-PCR method principle. doi: 10.17816/ecogen624820-4207493

Supplement 4. Schematic representation of the PCR principle with adapter ligation. doi: 10.17816/ecogen624820-4207494

Supplement 5. Schematic representation of the PCR principle with restriction site elongation. doi: 10.17816/ecogen624820-4207495

Supplement 6. Schematic representation of the Template-blocking PCR principle. doi: 10.17816/ecogen624820-4207496

Supplement 7. Schematic representation of the SSP-PCR principle. doi: 10.17816/ecogen624820-4207497

Supplement 8. Schematic representation of the PCR principle with a restriction site. doi: 10.17816/ecogen624820-4207498

Supplement 9. Schematic representation of the FPNI-PCR method principle. doi: 10.17816/ecogen624820-4207506

Supplement 10. Schematic representation of the SWPOP-PCR method principle. doi: 10.17816/ecogen624820-4207507

Supplement 11. Schematic representation of the PST-PCR method principle. doi: 10.17816/ecogen624820-4207508

Supplement 12. Schematic representation of the SLRA-PCR method principle. doi: 10.17816/ecogen624820-4207509

Supplement 13. Schematic representation of the FPR-PCR method principle. doi: 10.17816/ecogen624820-4207510

Supplement 14. Schematic representation of the wristwatch PCR method principle. doi: 10.17816/ecogen624820-4207511

Supplement 15. Schematic representation of the restriction-independent method for cloning genomic DNA segments beyond known sequences. doi: 10.17816/ecogen624820-4207512

## REFERENCES

- Pei J, Sun T, Wang L, et al. Fusion primer driven racket PCR: A novel tool for genome walking. *Front Genet.* 2022;13:969840. doi: 10.3389/fgene.2022.969840
- Uchiyama T, Watanabe K. Improved inverse PCR scheme for metagenome walking. *Biotechniques.* 2006;41(2):183–188. doi: 10.2144/000112210
- Kotik M. Novel genes retrieved from environmental DNA by polymerase chain reaction: Current genome-walking techniques for future metagenome applications. *J Biotechnol.* 2009;144(2):75–82. doi: 10.1016/j.jbiotec.2009.08.013
- Chang K, Wang Q, Shi X, et al. Stepwise partially overlapping primer-based PCR for genome walking. *AMB Express.* 2018;8(1):77. doi: 10.1186/s13568-018-0610-7
- Ochman H, Gerber AS, Hartl DL. Genetic applications of an inverse polymerase chain reaction. *Genetics.* 1988;120(3):621–623. doi: 10.1093/genetics/120.3.621
- Hui EK-W, Wang P-C, Lo SJ. Strategies for cloning unknown cellular flanking DNA sequences from foreign integrants. *Cell Mol Life Sci.* 1998;54(12):1403–1411. doi: 10.1007/s000180050262
- Kalendar R, Shustov AV, Schulman AH. Palindromic sequence-targeted (PST) PCR, version 2: An advanced method for high-throughput targeted gene characterization and transposon display. *Front Plant Sci.* 2021;12:691940. doi: 10.3389/fpls.2021.691940
- Riley J, Butler R, Ogilvie D, et al. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* 1990;18(10):2887–2890. doi: 10.1093/nar/18.10.2887
- Devon RS, Porteous DJ, Brookes AJ. Splinkerettes — improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res.* 1995;23(9):1644–1645. doi: 10.1093/nar/23.9.1644
- Lagerstrom M, Parik J, Malmgren H, et al. Capture PCR: efficient amplification of DNA fragments adjacent to a known sequence in human and YAC DNA. *Genome Res.* 1991;1(2):111–119. doi: 10.1101/gr.1.2.111
- Sarkar G, Turner RT, Bolander ME. Restriction-site PCR: a direct method of unknown sequence retrieval adjacent to a known locus by using universal primers. *Genome Res.* 1993;2(4):318–322. doi: 10.1101/gr.2.4.318
- Schmidt M, Hoffmann G, Wissler M, et al. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. *Hum Gene Ther.* 2001;12(7):743–749. doi: 10.1089/104303401750148649
- Jones DH, Winistorfer SC. Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA. *Nucleic Acids Res.* 1992;20(3):595–600. doi: 10.1093/nar/20.3.595
- Hengen PN. Vectorette, splinkerette and boomerang DNA amplification. *Trends Biochem Sci.* 1995;20(9):372–373. doi: 10.1016/s0968-0004(00)89079-9
- Yuanxin Y, Chengcai A, Li L, et al. T-linker-specific ligation PCR (T-linker PCR): an advanced PCR technique for chromosome walking or for isolation of tagged DNA ends. *Nucleic Acids Res.* 2003;31(12):e68. doi: 10.1093/nar/gng068
- O'Malley RC, Alonso JM, Kim CJ, et al. An adapter ligation-mediated PCR method for high-throughput mapping of T-DNA inserts in the *Arabidopsis* genome. *Nat Protoc.* 2007;2:2910–2917. doi: 10.1038/nprot.2007.425
- Ji J, Braam J. Restriction site extension PCR: A novel method for high-throughput characterization of tagged DNA fragments and genome walking. *PLoS ONE.* 2010;5(5):10577. doi: 10.1371/journal.pone.0010577

18. Bae J-H, Sohn J-H. Template-blocking PCR: An advanced PCR technique for genome walking. *Anal Biochem.* 2010;398(1):112–116. doi: 10.1016/j.ab.2009.11.003
19. Shyamala V, Ames GF. Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR. *Gene.* 1989;84(1):1–8. doi: 10.1016/0378-1119(89)90132-7
20. Parker JD, Rabinovitch PS, Burmer GC. Targeted gene walking polymerase chain reaction. *Nucleic Acids Res.* 1991;19(11):3055–3060. doi: 10.1093/nar/19.11.3055
21. Myrick KV, Gelbart WM. Universal fast walking for direct and versatile determination of flanking sequence. *Gene.* 2002;284(1–2):125–131. doi: 10.1016/s0378-1119(02)00384-0
22. Tan G, Gao Y, Shi M, et al. SiteFinding-PCR: a simple and efficient PCR method for chromosome walking. *Nucleic Acids Res.* 2005;33(13):e122. doi: 10.1093/nar/gni124
23. Liu Y-G, Whittier RF. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics.* 1995;25(3):674–681. doi: 10.1016/0888-7543(95)80010-j
24. Zeng T, Zhang D, Li Y, et al. Identification of genomic insertion and flanking sequences of the transgenic drought-tolerant maize line “SbSNAC1-382” using the single-molecule real-time (SMRT) sequencing method. *PLoS One.* 2020;15(4):e0226455. doi: 10.1371/journal.pone.0226455
25. Wang Z, Ye S, Li J, et al. Fusion primer and nested integrated PCR (FPNI-PCR): a new high-efficiency strategy for rapid chromosome walking or flanking sequence cloning. *BMC Biotechnol.* 2011;11:109. doi: 10.1186/1472-6750-11-109
26. Li H, Ding D, Cao Y, et al. Partially overlapping primer-based PCR for genome walking. *PLoS One.* 2015;10(3):120139. doi: 10.1371/journal.pone.0120139
27. Rose TM, Schultz ER, Henikoff JG, et al. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res.* 1998;26(7):1628–1635. doi: 10.1093/nar/26.7.1628
28. Levano-Garcia J, Verjovski-Almeida S, da Silva AC. Mapping transposon insertion sites by touchdown PCR and hybrid degenerate primers. *Biotechniques.* 2005;38(2):225–229. doi: 10.2144/05382ST03
29. Li F, Fu C, Li Q. A simple genome walking strategy to isolate unknown genomic regions using long primer and RAPD primer. *Iran J Biotechnol.* 2019;17(2): e2183. doi: 10.21859/ijb.2183
30. Wang L, Jia M, Li Z, et al. Wristwatch PCR: A versatile and efficient genome walking strategy. *Front Bioeng Biotechnol.* 2022;10:792848. doi: 10.3389/fbioe.2022.792848
31. Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *PNAS USA.* 1988;85(23):8998–9002. doi: 10.1073/pnas.85.23.8998
32. Zhou MY, Gomez-Sanchez CE. Universal TA cloning. *Curr Issues Mol Biol.* 2000;2(1):1–7.
33. Rudi K, Fossheim T, Jakobsen KS. Restriction cutting independent method for cloning genomic DNA segments outside the boundaries of known sequences. *Biotechniques.* 1999;27(6):1170–1172. doi: 10.2144/99276st03
34. Spalinskas R, Van den Bulcke M, Van den Eede G, Milcamps A. LT-RADE: An efficient user-friendly genome walking method applied to the molecular characterization of the insertion site of genetically modified maize MON810 and rice LLRICE62. *Food Anal Methods.* 2013;6:705–713. doi: 10.1007/s12161-012-9438-y
35. Leoni C, Gallerani R, Ceci LR. A genome walking strategy for the identification of eukaryotic nucleotide sequences adjacent to known regions. *BioTechniques.* 2008;44(2):229–235. doi: 10.2144/000112680
36. Tsafaris A, Pasentzis K, Argiriou A. Rolling circle amplification of genomic templates for inverse PCR (RCA-GIP): a method for 5'- and 3'-genome walking without anchoring. *Biotechnol Lett.* 2010;32:157–161. doi: 10.1007/s10529-009-0128-9
37. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science.* 2002;295(5558):1306–1311. doi: 10.1126/science.1067799
38. Krispil R, Tannenbaum M, Sarusi-Portuguez A, et al. The position and complex genomic architecture of plant T-DNA insertions revealed by 4SEE. *Int J Mol Sci.* 2020;21:2373. doi: 10.3390/ijms21072373

## СПИСОК ЛИТЕРАТУРЫ

1. Pei J, Sun T, Wang L, et al. Fusion primer driven racket PCR: A novel tool for genome walking // *Front Genet.* 2022. Vol. 13. ID 969840. doi: 10.3389/fgene.2022.969840
2. Uchiyama T, Watanabe K. Improved inverse PCR scheme for metagenome walking // *Biotechniques.* 2006. Vol. 41, N. 2. P. 183–188. doi: 10.2144/000112210
3. Kotik M. Novel genes retrieved from environmental DNA by polymerase chain reaction: Current genome-walking techniques for future metagenome applications // *J Biotechnol.* 2009. Vol. 144, N. 2. P. 75–82. doi: 10.1016/j.jbiotec.2009.08.013
4. Chang K, Wang Q, Shi X, et al. Stepwise partially overlapping primer-based PCR for genome walking // *AMB Express.* 2018. Vol. 8, N. 1. ID 77. doi: 10.1186/s13568-018-0610-7
5. Ochman H, Gerber A.S., Hartl D.L. Genetic applications of an inverse polymerase chain reaction // *Genetics.* 1988. Vol. 120, N. 3. P. 621–623. doi: 10.1093/genetics/120.3.621
6. Hui E.K.-W., Wang P.-C., Lo S.J. Strategies for cloning unknown cellular flanking DNA sequences from foreign inter-
- grants // *Cell Mol Life Sci.* 1998. Vol. 54, N. 12. P. 1403–1411. doi: 10.1007/s000180050262
7. Kalendar R., Shustov A.V., Schulman A.H. Palindromic sequence-targeted (PST) PCR, version 2: An advanced method for high-throughput targeted gene characterization and transposon display // *Front Plant Sci.* 2021. Vol. 12. ID 691940. doi: 10.3389/fpls.2021.691940
8. Riley J., Butler R., Ogilvie D., et al. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones // *Nucleic Acids Res.* 1990. Vol. 18, N. 10. P. 2887–2890. doi: 10.1093/nar/18.10.2887
9. Devon R.S., Porteous D.J., Brookes A.J. Splinkerettes — improved vectorettes for greater efficiency in PCR walking // *Nucleic Acids Res.* 1995. Vol. 23, N. 9. P. 1644–1645. doi: 10.1093/nar/23.9.1644
10. Lagerstrom M., Parik J., Malmgren H., et al. Capture PCR: efficient amplification of DNA fragments adjacent to a known sequence in human and YAC DNA // *Genome Res.* 1991. Vol. 1, N. 2. P. 111–119. doi: 10.1101/gr.1.2.111

11. Sarkar G., Turner R.T., Bolander M.E. Restriction-site PCR: a direct method of unknown sequence retrieval adjacent to a known locus by using universal primers // *Genome Res.* 1993. Vol. 2, N. 4. P. 318–322. doi: 10.1101/gr.2.4.318
12. Schmidt M., Hoffmann G., Wissler M., et al. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples // *Hum Gene Ther.* 2001. Vol. 12, N. 7. P. 743–749. doi: 10.1089/104303401750148649
13. Jones D.H., Winistorfer S.C. Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA // *Nucleic Acids Res.* 1992. Vol. 20, N. 3. P. 595–600. doi: 10.1093/nar/20.3.595
14. Hengen P.N. Vectorette, splinkerette and boomerang DNA amplification // *Trends Biochem Sci.* 1995. Vol. 20, N. 9. P. 372–373. doi: 10.1016/s0968-0004(00)89079-9
15. Yuanxin Y., Chengcai A., Li L., et al. T-linker-specific ligation PCR (T-linker PCR): an advanced PCR technique for chromosome walking or for isolation of tagged DNA ends // *Nucleic Acids Res.* 2003. Vol. 31, N. 12. P. e68. doi: 10.1093/nar/gng068
16. O'Malley R.C., Alonso J.M., Kim C.J., et al. An adapter ligation-mediated PCR method for high-throughput mapping of T-DNA inserts in the *Arabidopsis* genome // *Nat Protoc.* 2007. Vol. 2. P. 2910–2917. doi: 10.1038/nprot.2007.425
17. Ji J., Braam J. Restriction site extension PCR: A novel method for high-throughput characterization of tagged DNA fragments and genome walking // *PLoS ONE.* 2010. Vol. 5, N. 5. ID 10577. doi: 10.1371/journal.pone.0010577
18. Bae J.-H., Sohn J.-H. Template-blocking PCR: An advanced PCR technique for genome walking // *Anal Biochem.* 2010. Vol. 398, N. 1. P. 112–116. doi: 10.1016/j.ab.2009.11.003
19. Shyamala V., Ames G.F. Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR // *Gene.* 1989. Vol. 84, N. 1. P. 1–8. doi: 10.1016/0378-1119(89)90132-7
20. Parker J.D., Rabinovitch P.S., Burmer G.C. Targeted gene walking polymerase chain reaction // *Nucleic Acids Res.* 1991. Vol. 19, N. 11. P. 3055–3060. doi: 10.1093/nar/19.11.3055
21. Myrick K.V., Gelbart W.M. Universal fast walking for direct and versatile determination of flanking sequence // *Gene.* 2002. Vol. 284, N. 1–2. P. 125–131. doi: 10.1016/s0378-1119(02)00384-0
22. Tan G., Gao Y., Shi M., et al. SiteFinding-PCR: a simple and efficient PCR method for chromosome walking // *Nucleic Acids Res.* 2005. Vol. 33, N. 13. P. e122. doi: 10.1093/nar/gni124
23. Liu Y.-G., Whittier R.F. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking // *Genomics.* 1995. Vol. 25, N. 3. P. 674–681. doi: 10.1016/0888-7543(95)80010-j
24. Zeng T., Zhang D., Li Y., et al. Identification of genomic insertion and flanking sequences of the transgenic drought-tolerant maize line “SbSNAC1-382” using the single-molecule real-time (SMRT) sequencing method // *PLoS One.* 2020. Vol. 15, N. 4. ID e0226455. doi: 10.1371/journal.pone.0226455
25. Wang Z., Ye S., Li J., et al. Fusion primer and nested integrated PCR (FPNI-PCR): a new high-efficiency strategy for rapid chromo-  
some walking or flanking sequence cloning // *BMC Biotechnol.* 2011. Vol. 11. ID 109. doi: 10.1186/1472-6750-11-109
26. Li H., Ding D., Cao Y., et al. Partially overlapping primer-based PCR for genome walking // *PLoS One.* 2015. Vol. 10, N. 3. ID 120139. doi: 10.1371/journal.pone.0120139
27. Rose T.M., Schultz E.R., Henikoff J.G., et al. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences // *Nucleic Acids Res.* 1998. Vol. 26, N. 7. P. 1628–1635. doi: 10.1093/nar/26.7.1628
28. Levano-Garcia J., Verjovski-Almeida S., da Silva A.C. Mapping transposon insertion sites by touchdown PCR and hybrid degenerate primers // *Biotechniques.* 2005. Vol. 38, N. 2. P. 225–229. doi: 10.2144/05382ST03
29. Li F., Fu C., Li Q. A simple genome walking strategy to isolate unknown genomic regions using long primer and RAPD primer // *Iran J Biotechnol.* 2019. Vol. 17, N. 2. ID e2183. doi: 10.21859/ijb.2183
30. Wang L., Jia M., Li Z., et al. Wristwatch PCR: A versatile and efficient genome walking strategy // *Front Bioeng Biotechnol.* 2022. Vol. 10. ID 792848. doi: 10.3389/fbioe.2022.792848
31. Frohman M.A., Dush M.K., Martin G.R. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer // *PNAS USA.* 1988. Vol. 85, N. 23. P. 8998–9002. doi: 10.1073/pnas.85.23.8998
32. Zhou M.Y., Gomez-Sanchez C.E. Universal TA cloning // *Curr Issues Mol Biol.* 2000. Vol. 2, N. 1. P. 1–7.
33. Rudi K., Fossheim T., Jakobsen K.S. Restriction cutting independent method for cloning genomic DNA segments outside the boundaries of known sequences // *Biotechniques.* 1999. Vol. 27, N. 6. P. 1170–1172. doi: 10.2144/99276st03
34. Spalinskas R., Van den Bulcke M., Van den Eede G., Milcamps A. LT-RADE: An efficient user-friendly genome walking method applied to the molecular characterization of the insertion site of genetically modified maize MON810 and rice LLRICE62 // *Food Anal Methods.* 2013. Vol. 6. P. 705–713. doi: 10.1007/s12161-012-9438-y
35. Leoni C., Gallerani R., Ceci L.R. A genome walking strategy for the identification of eukaryotic nucleotide sequences adjacent to known regions // *BioTechniques.* 2008. Vol. 44, N. 2. P. 229–235. doi: 10.2144/000112680
36. Tsaftaris A., Pasentzis K., Argiriou A. Rolling circle amplification of genomic templates for inverse PCR (RCA-GIP): a method for 5'- and 3'-genome walking without anchoring // *Biotechnol Lett.* 2010. Vol. 32. P. 157–161. doi: 10.1007/s10529-009-0128-9
37. Dekker J., Rippe K., Dekker M., Kleckner N. Capturing chromosome conformation // *Science.* 2002. Vol. 295, N. 5558. P. 1306–1311. doi: 10.1126/science.1067799
38. Krispil R., Tannenbaum M., Sarusi-Portuguez A., et al. The position and complex genomic architecture of plant T-DNA insertions revealed by 4SEE // *Int J Mol Sci.* 2020. Vol. 21. ID 2373. doi: 10.3390/ijms21072373

## AUTHORS' INFO

**\*Elena S. Okulova**; address: 7–9 Universitetskaya emb.,  
Saint Petersburg, 199034, Russia; ORCID: 0009-0001-7349-8925;  
eLibrary SPIN: 7166-0090; e-mail: elenaok.advert@gmail.com

**Mikhail S. Burlakovsky**, Cand. Sci. (Biology);  
ORCID: 0000-0001-6694-0423; eLibrary SPIN: 3679-0860;  
e-mail: burmish@yandex.ru

**Ludmila A. Lutova**, Dr. Sci. (Biology), Professor;  
ORCID: 0000-0001-6125-0757; Scopus Author ID: 6603722721;  
eLibrary SPIN: 3685-7136; e-mail: la.lutova@gmail.com

## ОБ АВТОРАХ

**\*Елена Сергеевна Окулова**; адрес: Россия,  
199034, Санкт-Петербург, Университетская наб., д. 7–9;  
ORCID: 0009-0001-7349-8925; eLibrary SPIN: 7166-0090;  
e-mail: elenaok.advert@gmail.com

**Михаил Сергеевич Бурлаковский**, канд. биол. наук;  
ORCID: 0000-0001-6694-0423; eLibrary SPIN: 3679-0860;  
e-mail: burmish@yandex.ru

**Людмила Алексеевна Лутова**, д-р биол. наук, профессор;  
ORCID: 0000-0001-6125-0757; Scopus Author ID: 6603722721;  
eLibrary SPIN: 3685-7136; e-mail: la.lutova@gmail.com

\* Corresponding author / Автор, ответственный за переписку