Chapter 15

Selecting Specific PCR Primers with MFEprimer

Wubin Qu and Chenggang Zhang

Abstract

Selecting specific primers is crucial for polymerase chain reaction (PCR). Nonspecific primers will bind to unintended genes and result in nonspecific amplicons. MFEprimer is a program for checking the specificity of PCR primers against the background DNA. In this chapter, we introduce: (1) the factors that affect the specificity of primers; (2) the principle of MFEprimer and its settings; (3) how to use the MFEprimer to examine the specificity of primers.

Key words PCR, Primer specificity, MFEprimer

1 Introduction

Designing specific primers is a key step in PCR. The forward and reverse primers act like two anchors, which determine the region of interest to be amplified. However, in real experiments, due to the presence of background DNA, which usually including genomic DNA (gDNA) and complementary DNA (cDNA), even the well-designed primers can bind to mistargeted DNA regions and result in nonspecific amplicons. The background DNA here acts as competitor of target DNA. Nowadays, with the availability of the genomic DNA sequences of most model organisms, it is now possible to check the specificity of primers against the background DNA.

MFEprimer(http://biocompute.bmi.ac.cn/CZlab/MFEprimer-2.0/) is a program to check the specificity of PCR primers against the background DNA. The first version of MFEprimer was released in 2009 [1] and the current version is 2.0 [2]. Although there are several programs available [3–6] for checking the specificity of PCR primers before our work, none of them treats primer hybrid process on the base of the biochemical reactions. These programs mainly use the sequence alignment tools such as National Center for Biotechnology Information's Basic Local Alignment Search Tool (NCBI BLAST) [7] program to find the primer binding sites and neglect the important factors which influence the specificity of

primers, such as the binding stability of the 3′-end of a primer duplex (formed by the primer and its target DNA). Other limitations of these programs include fewer organisms support and slower running speed. Therefore, we developed the MFEprimer program with several important features: (1) Mimicking the thermodynamics hybrid process of the primers and the target amplicons. (2) Using the k-mer algorithm to speed up the binding sites searching process, for example, the running time is usually 1–10 s even for the whole genomic DNA searching with only one CPU used in popular personal computer. Here, a k-mer is defined as a short DNA sequence with a length of k nucleotides. (3) Comprehensive supported organisms and support for custom database upon users' request.

In the following sections, we (1) introduce the factors that affect the specificity of primers, (2) explain the principle of MFEprimer and its settings, and (3) use an example to illustrate the usage of MFEprimer.

2   Factors That Affect the Specificity of PCR Primers

Here, we focus on the nonexperimental factors that affect the specificity of PCR primers. The nonspecific binding sites of the PCR primers are the major sources responsible for nonspecific amplifications. There are three factors which can influence the nonspecific binding sites of a primer: (1) Background DNA is the source with a large number of potential binding sites. Therefore, it is critical to know the composition of background DNA. (2) Stable binding sites (including nonspecific ones) can form stable primer duplex, which in turn have chance to attract the DNA polymerase to bind for next elongation process. (3) After DNA polymerase binds to primer duplex, only the stable 3′-end of a primer duplex can attract DNA polymerase to start the elongation process. In other words, for a given primer and background DNA, the binding stability of the whole primer determines whether a binding site would be bound by DNA polymerase, while the stability of its 3′-end (default is the last nine residues [*see* **Note 1**]) determines whether a nonspecific amplicon would be amplified.

2.1   The Background DNA

The background DNA acts as a competitor of target DNA and is the major source of the unexpected primer binding sites. Generally, the user aims to selectively amplify the specific region from the target DNA, which mixed with the background DNA. Therefore, it is obviously that we should design primers specific to target DNA and avoid nonspecific binding to background DNA. As a result, knowing background DNA is the precondition of designing specific PCR primers.

2.2 The Stability of the Primer Duplex

The annealing and extension are two key steps, which determine the anchor sites of a PCR primer. The annealing process determines which sites a primer could bind to, while the extension step decides whether the binding site could be extended.

In the annealing process of PCR, although there are plenty of DNA sites (from both target DNA and background DNA) available for binding, the primer can only bind to the sites which can form stable duplex. The binding stability of the binding site is measured by Gibbs free energy (ΔG) [8, 9]. The lower of the ΔG, the more stable is the duplex. The most stable binding site of the PCR primer is the site that has perfect match (its complement reverse sequence, shown in Fig. 1a). Some binding sites though have one or two mismatches (shown in Fig. 1b, c), still can form stable duplex with the primer [8–10]. That is because some mismatches still contribute much to the duplex stability. For instance, a G–G mismatch can contribute as much as –2.2 kcal/mol to duplex stability at 37 °C [8] (*see* **Note 2**).

2.3 The 3′-End Stability of the Primer Duplex

After the annealing process, the DNA polymerase will bind to the 3′ end of the primer duplex, and then synthesize a new DNA strand complementary to the DNA template strand by adding dNTPs that are complementary to the template in 5′–3′ direction. Therefore, a stable 3′-end of a primer is essential for elongation, but not for the 5′ end [10, 11]. Therefore, the 5′- and 3′-ends of a primer have different scoring weights [10]. An unstable 3′-end (e.g., mismatches in 3′-end, shown in Fig. 1c) will stop the elongation process, while the 5′-end allows an addition of a tagging sequence or mismatches and did not affect the 3′-end elongation (shown in Fig. 1b). However, a high stable 3′-end will also contribute substantially to nonspecific extension in PCR reactions [1, 11] (*see* **Note 1**).
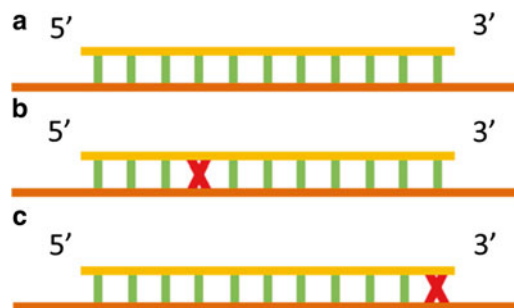

Fig. 1 Primer binding diagrams. Primer binds perfectly (**a**), with one mismatch near the 5′ terminal (**b**) and with one mismatch at the 3′ terminal (**c**). *Yellow bar*: the PCR primer. *Long red bar*: the DNA template sequence. *Red cross*: mismatch region

3 MFEprimer

3.1 Principle

MFEprimer is a fast and thermodynamics-based PCR primer specificity-checking program [1, 2]. The genomic DNA (gDNA) and complementary DNA (cDNA) were used as background DNAs for analysis. The PCR primer binding sites were searched against the whole background DNA with a *k*-mer-based algorithm and were sorted by the stability of the primer duplex. The searching results include both target amplicon and nonspecific potential amplicons. The main MFEprimer workflow consists of the below four steps:

1. Find the stable binding sites of the 3′-end of both forward primer and reverse primer among the target and background DNA sequences. In default, the stability of the last nine residues at the 3′-end of a primer was used to represent the primer 3′-end stability (*see* **Note 3**). The Nearest-Neighbor (NN) model [12] was used to calculate the stability (*see* **Note 4**). The unstable binding sites of the 3′-end will be filtered out directly and will not go through the following steps because only the primer duplex with stable 3′-end has the ability to start the extension reaction (*see* **Note 1**).

2. Calculating the binding stability of the entire primer sequence. The more stable the primer–template duplex, the higher the probability of the binding reaction.

3. Running a virtual PCR analysis and filtering out the unreasonable amplicons, such as the amplicon with 100,000,000 bp in size.

4. Output the predicted PCR amplicons. All the amplicons were sorted by the joint binding stability of both forward and reverse primers. The joint binding stability is the geometric mean of the binding stabilities of forward primer and reverse primer. The amplicons include not only the target amplicon but also the potential nonspecific amplicons.

3.2 Settings

Figure 3 shows the main features of the MFEprimer-2.0 home page. There are four sections in the home page:

1. *Background database selection*: This section is mandatory. Currently, there are 27 species databases available and some of them are added upon the request from the users, such as the dog's genomics DNA database [13].

2. *Input the primer sequences* (5′ → 3′): This section is also mandatory and the primers should be in 5′–3′ direction. Multiple primers are allowed and will be considered as either forward primer or reverse primer.

3. *Results filter settings*: This section is optional. Usually, MFEprimer-2.0 will output many predicted amplicons, including the ones that probably won't be amplified in real PCR. A very long list of amplicons will be useless, and users will still don't know how to evaluate the specificity of the primers. Results

filter settings can filter out these low probability amplicons. For example, in most cases, there are some predicted amplicons which have very low T_m values (*see* **Note 3**) and these amplicons would not be amplified in real PCR due to the T_m difference between these amplicons and the target amplicon. In this situation, the low T_m amplicons should be filtered out and let users focus on the amplicons that will be probably amplified (*see* **Note 5**).

4. *Experimental settings*: This section is optional. These settings are used by the nearest-neighbor (NN) model to calculate the melting temperature and Gibbs free energy. Different values of these parameters will result in different T_m values. We recommended the users to specify real PCR conditions to get the most accurate estimation of the T_m value.

4   How to Use MFEprimer to Check the Specificity of Primers

MFEprimer can check most types of primers, including conventional PCR primers, real-time PCR primers, degenerate primers, and multiplex PCR primers. Besides the primers, users should also provide the accurate background DNA, which is the DNA from the user's own sample. For example, if the user wishes to amplify a region of a human gene, the human gDNA and cDNA should be selected as the background DNA. However, if the gDNA was completely digested when preparing the sample during the experiment, then only the cDNA sequence database should be selected as the background DNA (*see* **Note 6**).

4.1 Preparing Primers

We have shown an example PCR primer pair, 5′-GACCAGTAGA CCTCGGCGAA-3′ and 5′-ACCTTGGCAAGTGCTCCTCT-3′, to amplify and distinguish the two transcript variants (GenBank No. NM_010923 and NM_180960) of the mouse Nnat (GeneID: 18111) gene. The gene structure of these two transcript variants and the locations of the two primers are shown in Fig. 2 (*see* **Note 7**). And the figure is from the VizPrimer [14] web site (http://biocompute.bmi.ac.cn/CZlab/VizPrimer).
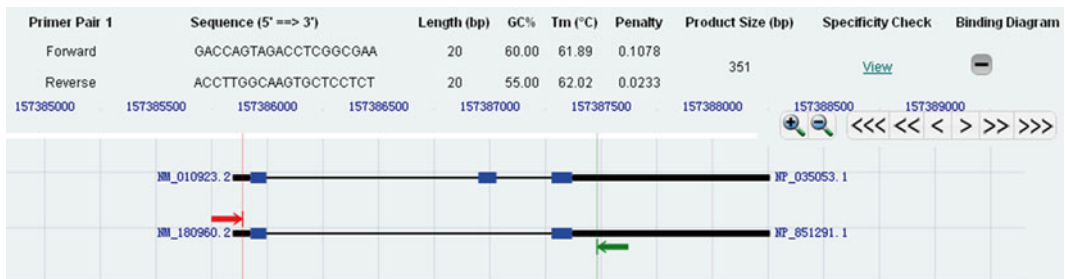

Fig. 2 The gene structure of two transcript variants of the mouse Nnat gene and the locations of the two primers

Fig. 3 Settings in the home page of MFEprimer-2.0

Fig. 4 The "Query" section of the MFEprimer-2.0 result page

4.2 MFEprimer Settings

As Fig. 3 shows, we first selected the "Mouse—RNA & Genomic" as the background DNA to avoid the sample from being potentially contaminated with the genomic DNA. Secondly, the primer sequences were written in $5' \rightarrow 3'$ direction. In the "Results filter settings" section, we set more strict cutoff values for T_m "From" = 10 and primer pair coverage (PPC) "From" = 10. For other parameters we use the default value. Finally, we can click the "Run" button to launch the MFEprimer program.

4.3 The MFEprimer Results

The first line of the result page begins with the title "MFEprimer-2.0 Report," followed by several shortcut links for each section of the result (shown in Fig. 4). Usually, the result page is long and these links can guide users to the right section quickly. The MFEprimer-2.0 result page contains six sections below.

1. *Query*: Fig. 4 shows the list of the input primer sequences with general information such as size, GC content, and T_m value annotation.

2. *Histogram of size, T_m, and ΔG of x potential amplicons*: This section provides histograms of the size, ΔG, and T_m when the number of predicted amplicons >10. The users can see what the worst-case scenario is for their primers and decide how to proceed (shown in Fig. 5).

3. *Descriptions of x potential amplicons*: Fig. 6 shows the brief description of all the potential amplicons predicted by MFEprimer-2.0. The amplicons in red are the ones that could be probably amplified in PCR.

4. *Amplicon details*: Information on the hybridization details for each predicted amplicon and the amplicon sequence. Figure 7 shows the first amplicon.

5. *Parameters*: The top part of Fig. 8, the parameters used during the evaluation process.

6. *Citation*: The bottom part of Fig. 8, these papers report the MFEprimer program.
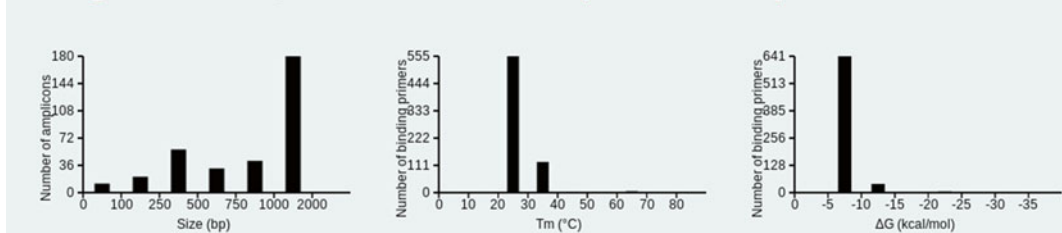

Fig. 5 The "Histogram of size, T_m, and ΔG of x potential amplicons" section of MFEprimer-2.0 result page

Descriptions of **344** potential amplicons

What's this?   How to explain the result?

ID		Accession	Fp x Rp	Size (bp)	PPC (%)	Fp Tm (°C)	Rp Tm (°C)	Fp ΔG (kcal/mol)	Rp ΔG (kcal/mol)	Fp 3' ΔG (kcal/mol)	Rp 3' ΔG (kcal/mol)
☐	1	NC_000068.6	Seq1 x Seq2	1632	100.0	61.9	62.0	-22.9	-22.6	-4.0	-2.9
☐	2	NM_010923.2	Seq1 x Seq2	432	100.0	61.9	62.0	-22.9	-22.6	-4.0	-2.9
☐	3	NM_180960.2	Seq1 x Seq2	351	100.0	61.9	62.0	-22.9	-22.6	-4.0	-2.9
☐	4	NC_000068.6	Seq2 x Seq2	92	-36.0	38.3	38.3	-11.6	-11.6	-2.9	-2.9
☐	5	NC_000067.5	Seq2 x Seq2	1083	-30.3	35.2	35.2	-10.8	-10.8	-2.9	-2.9
☐	6	NC_000067.5	Seq2 x Seq2	321	-26.5	30.7	33.0	-9.7	-10.1	-2.9	-2.9
☐	7	NC_000079.5	Seq2 x Seq2	302	-26.5	30.7	33.2	-9.7	-10.2	-2.9	-2.9
☐	8	NC_000074.5	Seq2 x Seq2	1874	-25.6	35.2	30.7	-10.8	-9.7	-2.9	-2.9
☐	9	NC_000070.5	Seq2 x Seq2	1110	-25.6	30.7	35.2	-9.7	-10.8	-2.9	-2.9
☐	10	NC_000069.5	Seq2 x Seq2	332	-25.6	35.2	30.7	-10.8	-9.7	-2.9	-2.9
☐	11	NC_000069.5	Seq2 x Seq2	252	-25.6	35.2	30.7	-10.8	-9.7	-2.9	-2.9

Fig. 6 The part of "Descriptions of x potential amplicons" section of the MFEprimer-2.0 result page. This figure is reproduced from (MFEprimer-2.0: A fast thermodynamics-based program for checking PCR primer specificity 2012) with permission from Oxford University Press [2]

Amplicon details

What's this?   How to explain the result?

☐ 1: Seq1 + Seq2 ==> gi|149338249|ref|NC_000068.6| Mus musculus strain C57BL/6J chromosome 2, MGSCv37 C57BL/6J

PPC = 100.0%, Size = 1632 (bp), GC content = 52.7%
Fp: Tm = 61.9 (°C), ΔG = -22.9 (kcal/mol), 3'ΔG = -4.0 (kcal/mol)
Rp: Tm = 62.0 (°C), ΔG = -22.6 (kcal/mol), 3'ΔG = -2.9 (kcal/mol)
Binding sites: 157385870(20/20) ... 157387501(20/20)

```
>>>Seq1
      1                 20
5' GACCAGTAGACCTCGGCGAA 3'
   ::::::::::::::::::::                                                    157387501
5' GACCAGTAGACCTCGGCGAAcccttgctctcgaccaccca...gtcccctgtgttccctcgtcAGAGGAGCACTTGCCAAGGT 3'
   157385870                                            ::::::::::::::::::::
                                                     3' AGAGGAGCACTTGCCAAGGT 5'
                                                        20                 1
                                                                    Seq2<<<
```

>Amp_1 Seq1 + Seq2 ==> gi|149338249|ref|NC_000068.6| Mus musculus strain C57BL/6J chromosome 2, MGSCv37 C57BL/6J
```
GACCAGTAGACCTCGGCGAAcccttgctctcgaccacccacccactttcggaaccatggccgcagtggcagcagcctcgg
cagaactgctcatcatcggctggtacatcttccgcgtgctgctgcaggtgagtatgtaccccgggctttgggtgggagagg
gcaccccgaaacgcgcccccgggacccccttaatccatatcgcgattgccgcccctacccaacccatcctatcccaattgccg
caatccctgtgtaccaaagtgccgcggtggtactatttgtccccgcccgttcccgtgcccacatcctcgcgcagaccctg
tccccagtgcgcccgcgcccagccgggaacaaagactcaggggtgcggcggtcgcccgctgcggacgatcgcgaccaacg
cggtaagaaatacccgctatattcaggctcgatcccaggagggaggcggggcacttgtgtgctgtgaaactttacagttc
gcacaacagaaaattttcaccttaaaaaattgcgcaatgcaggaaattttcttttaaatacactatgtacacacatataa
cacttgggtgggtgaaaaagaagttagaaaatggcactcctcaggaaaaataacaaaaagaatctcaataaggggaaag
ggaggaaatagagaaaacgcgcattgcaggaaaaataagccagacaaaggcacttactaggagaaaaatggattagataa
aagtgcattggaaaagaaacaaagaaacgaaaaaaaaaaaaaaaaaaaaaccgaattggaaacacactagagaaaaaatgc
cccttgcaggacaaaaacggaaggggaaagaacatttcaaaaaaaaaaaaaaaaagaaagaaagaaagaaaaagaaaaaa
gaaagaaagaaagaaagaaagggggcaaatcattttatggaaaaagagagtgcgcgatatcgcgcgcaaaagtgcttctct
taaagccaaactggaagaaagcgctttgtgcaaaaaaaaactatctaccccaagagctccctttgcaggaatcctgctaa
aggaatcctttgccaaaggaatcgcatattttccttcaaggtgttcctggaatgctgcatttactgggtaggattcgcttt
tcgaaatcctccagggacacagcccattgcgagaagtgaggtatacttaagttgtgggtccaatcagcttgcagccatgc
agctctcagcacagttggaaaagctccagctgccctgactggtggacaagctgtgcctcggccgccactccaggctatgc
tggacaggcgggcggggcaggggcgggccgggcgggcaaagcagcagcagcaaggcagcagtagcagcagtagcacaca
ggatcatgggtacttctctaagggtgggtcctgggtctctcaccgcaggtgttcaggtactccctgcagaagctggcgca
cacggtgtcccggaccgggcggcaggtgctgggggagcgcaggcagcgagcccccaactgaggccccagctcccagccct
gggcggccgtgtcatcaggtgctcctgtgcttctcgaccagcatgggagccagtgccgcgcaggaatgggggggtccctg
tgttccctcgtcAGAGGAGCACTTGCCAAGGT
```

Fig. 7 The part of "Amplicons details" section of the MFEprimer-2.0 result page

Parameters

Database: Mouse.rna, Mouse.genomic	K value: 9
Degenerate sequence: No	PPC cutoff: 10.0 (%)
Size start: 50 (bp)	Size stop: 2000 (bp)
Tm start: 10.0 (°C)	Tm stop: 80.0 (°C)
Concentration of monovalent cations: 50.0 (mM)	Concentration of divalent cations: 1.5 (mM)
Concentration of annealing oligos: 50.0 (nM)	Concentration of dNTPs: 0.25 (mM)
Time used: 0 m 1 s	Date: 2013-04-03 14:04:01

Citation

Wubin Qu, Yang Zhou, Yanchun Zhang, Yiming Lu, Xiaolei Wang, Dongsheng Zhao, Yi Yang and Chenggang Zhag. (2012) MFEprimer-2.0: A fast thermodynamics-based program for checking PCR primer specificity. Nucleic Acids Res. (Web Server Issue): W205-8. DOI: 10.1093/nar/gks552.

Wubin Qu, Zhiyong Shen, Dongsheng Zhao, Yi Yang and Chenggang Zhang. (2009) MFEprimer: multiple factor evaluation of the specificity of PCR primers. Bioinformatics, 25(2), 276-278.

Fig. 8 The "Parameters" and "Citation" sections of the MFEprimer-2.0 result page

4.4 The General Rules to Analyze the MFEprimer Results

In general, there are several steps required to analyze the MFEprimer results before drawing the conclusion:

1. Finding the target amplicon. If there is no amplicon returned or the program is unable to find the target amplicon, then the primer pair is failed to amplify the target DNA and should be redesigned.

2. Analyzing the red amplicons. In most cases, the target amplicon is located in this red group and the real PCR conditions should be optimized for amplifying the target DNA. Other amplicons in red are the ones which have thermodynamics features close to the target amplicon and could be probably amplified in the PCR. Therefore, if there are other amplicons in red group, the users may run a multiple align [15, 16] or cluster [7] analysis with the amplicons to disclose the relationship between the target amplicon and these amplicons:

    (a) If they are homologous sequences, such as transcript variants from one gene, then these amplicons should be regarded as specific amplicons.

    (b) If they are not homologous sequences, these amplicons might usually be nonspecific amplicons, and the primer pair should be discarded. However, if we have to use the primer pair or the primer pair is quite difficult to redesign, a further analysis is required. In this situation, if the detection method for amplicons can distinguish the target and nonspecific amplicons, the primer pair is also acceptable. For example, if the user uses electrophoresis to detect the amplicons and none of them has similar size with the target amplicon, then the primer pair is acceptable. However, if we use hybridization (e.g., probe) to detect the amplicons, the primer pair would be accepted only when the probe hybrids to the target amplicon specifically.

3. Be careful to the rest of the amplicons. If there are too many predicted amplicons (*see* **Note 8**), then the primer pair should be redesigned as recommended [17]. On the other side, although these amplicons have lower binding energy (lower T_m and higher ΔG value), they also have the possibility to be amplified [18], especially when they have small sizes or the annealing temperature is low. As a result, our suggestion to avoid amplifying these amplicons is using a higher annealing temperature.

4.5   Analyzing the Example MFEprimer Results

Figure 6 shows the brief description of 344 potential amplicons predicted by MFEprimer-2.0 for the primer pair of the mouse Nnat gene. Following the steps described in the last section:

1. We can find the target amplicons: NM_010923 (ID: 2) and NM_180960 (ID: 3).

2. Besides the two target amplicons, there is another amplicon (ID: 1, GenBank No. NC_000068) with size of 1,632 bp in the red group. Multiple sequence alignment result shows that this amplicon contains the other two amplicons as shown in Fig. 9. The search result of NM_010923 on GenBank (http://www.ncbi.nlm.nih.gov/gene/?term=NM_010923.2) indicated that the sequence NC_000068 (ID: 1) is the genomic DNA sequence region of the mouse Nnat gene. Therefore, this amplicon should

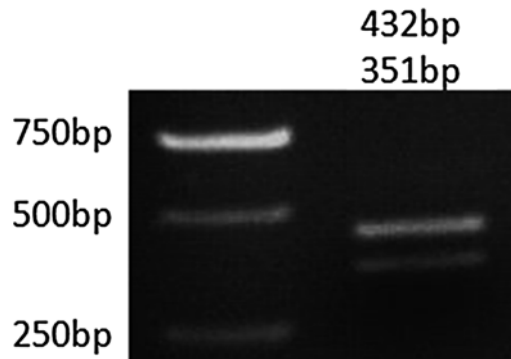**Fig. 9** Multiple sequence alignment of the three amplicons in *red font*

Fig. 10 Electrophoresis analysis of the two transcript variants of the mouse Nnat gene

not be considered as nonspecific amplicon. To be noticed that, if the sample was contaminated by genomic DNA, then the amplicon would be amplified.

3. Besides the amplicons in red font, there are also 343 amplicons predicted by MFEprimer-2.0. Compared to the target amplicon, these amplicons have significantly lower T_m values and higher ΔG values. Therefore, when we use an annealing temperature optimized for the target amplicon, these amplicons should not be amplified. In other words, we can use a higher annealing temperature to suppress the amplification of these nonspecific amplicons.

4. Directed by the MFEprimer-2.0 result, we experimentally validated the primer pair and obtained a high-quality result as shown in Fig. 10.

5   Notes

1. According to Rychlik's work [18], a continuous primer template duplex stable more than –11 kcal/mol will initialize the extension reaction. Therefore, if the stability of 3′-end of a primer duplex is equal or smaller than –11 kcal/mol, the duplex will probably be amplified. MFEprimer uses –9 kcal/mol as a strict cutoff value for safety consideration.

2. We would like to clarify here that the sequence alignment results (e.g., perfect match, mismatch) only indicate the nucleotide composition of the binding sites but not reflect the binding stability of the duplex. For example, the G-C and A-T matches have equal scores in the search results of National Center for Biotechnology Information's Basic Local Alignment Search Tool (NCBI BLAST) [7], but it is well known that the G-C match is more stable than the A-T match. Therefore, it is inappropriate to use NCBI BLAST or other similar tools [19] to check the specificity of the PCR primers.

3. The annealing temperature of a successful PCR is usually greater than 55 °C to suppress nonspecific PCR amplification. Therefore, primers with T_m values less than 55 °C would not be amplified in PCR reaction [2]. MFEprimer-2.0 uses a strict cutoff value of 48 °C to check the 3′-end stability of the primer duplex. To determine the maximum size of the 3′-end subsequence to represent its stability, we calculated all of the k-mer values, with k ranging from 5 to 9. The result shows that $k=9$ (the T_m of most stable 9-mer sequence is 48.4 °C) is suitable for the maximum size of the 3′-end subsequence. Moreover, the command-line version of MFEprimer-2.0 also supports custom k values, for example to set $k=8$ for a more strict 3′-end stability checking.

4. Although there are several methods available for predicting melting temperature, according to Panjkovich's [20] and Ahsen's [21] works, the most accurate T_m predicting method is the Nearest-Neighbor method [22].

5. We defined a parameter primer pair coverage (PPC) to score the ability of the primer pair (a forward primer coupled with a reverse primer) binding to the DNA template using the following formula (Fig. 11):

    Where Fm and Rm are sequence overlaps of the forward primer and reverse primer with the DNA template, and Fl and Rl are the full lengths of the forward primer and reverse primer, respectively. CVfr is the coefficient of variability of matched length of forward primer (Fm) and reverse primer (Rm). The maximum value of PPC is 100 %, indicating a pair of primers with the same length and both of them binding to the template completely.

6. It is recommended to select the "Genomic DNA" as background DNA to prevent the sample from being contaminated with genomic DNA.

7. Knowing the gene structure is essential for PCR primer design. For the gene has several transcript variants, it is important to know: (1) Which exon did the primers locate? (2) Can this primer pair distinguish these transcript variants? (3) Is the primer located at the junction site between the two adjacent exons?

8. According to Andreson's [17] work, the more the predicted amplicons, the larger the experimental failure rate. In detail, if the number of predicted amplicons is in 2–20, the failure rate is about 20 %; but if the number is in 10–100, the failure rate is about 40 %; and if the number is >10,000, the failure rate is >70 %.

$$\text{PPC} = \frac{Fm}{Fl} \times \frac{Rm}{Rl} \times (1 - CVfr)$$

Fig. 11 The formula of PPC

Acknowledgement

References

1. Qu W, Shen Z, Zhao D et al (2009) MFEprimer: multiple factor evaluation of the specificity of PCR primers. Bioinformatics 25(2):276–278

2. Qu W, Zhou Y, Zhang Y et al (2012) MFEprimer-2.0: a fast thermodynamics-based program for checking PCR primer specificity. Nucleic Acids Res 40(W1):W205–W208

3. Lexa M, Horak J, Brzobohaty B (2001) Virtual PCR. Bioinformatics 17(2):192–193

4. Lexa M, Valle G (2003) PRIMEX: rapid identification of oligonucleotide matches in whole genomes. Bioinformatics 19(18):2486–2488

5. Boutros PC, Okey AB (2004) PUNS: transcriptomic- and genomic-in silico PCR for enhanced primer design. Bioinformatics 20(15):2399–2400

6. Andreson R, Kaplinski L, Remm M (2007) Fast masking of repeated primer binding sites in eukaryotic genomes. Methods Mol Biol 402:201–218

7. Altschul SF, Madden TL, Schaffer AA et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402

8. SantaLucia J Jr (2007) Physical principles and visual-OMP software for optimal PCR design. Methods Mol Biol 402:3–34

9. SantaLucia J, Hicks D (2004) The thermodynamics of DNA structural motifs. Annu Rev Biophys Biomol Struct 33:415–440

10. Onodera K, Melcher U (2004) Selection for 3′ end triplets for polymerase chain reaction primers. Mol Cell Probes 18(6):369–372

11. Miura F, Uematsu C, Sakaki Y et al (2005) A novel strategy to design highly specific PCR primers based on the stability and uniqueness of 3′-end subsequences. Bioinformatics 21(24):4363–4370

12. SantaLucia J (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor hermodynamics. Proc Natl Acad Sci U S A 95(4):1460–1465

13. Lindblad-Toh K, Wade CM, Mikkelsen TS et al (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438(7069):803–819

14. Zhou Y, Qu W, Lu Y et al (2011) VizPrimer: a web server for visualized PCR primer design based on known gene structure. Bioinformatics 27(24):3432–3434

15. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302(1):205–217

16. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797

17. Andreson R, Möls T, Remm M (2008) Predicting failure rate of PCR in large genomes. Nucleic Acids Res 36(11):e66

18. Rychlik W (1995) Priming efficiency in PCR. Biotechniques 18(1):84–86, 88–90

19. Kent WJ (2002) BLAT—the BLAST-like alignment tool. Genome Res 12(4):656–664

20. Panjkovich A, Melo F (2005) Comparison of different melting temperature calculation methods for short DNA sequences. Bioinformatics 21(6):711–722

21. von Ahsen N, Wittwer CT, Schutz E (2001) Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections for Mg2+, deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. Clin Chem 47(11):1956–1961

22. Crothers DM, Zimm BH (1964) Theory of the melting transition of synthetic polynucleotides: evaluation of the stacking free energy. J Mol Biol 9:1–9