

Primer3—new capabilities and interfaces

Andreas Untergasser¹, Ioana Cutcutache², Triinu Koressaar³, Jian Ye⁴,
Brant C. Faircloth⁵, Mairo Remm³ and Steven G. Rozen^{2,*}

¹Zentrum für Molekulare Biologie der Universität Heidelberg, DKFZ-ZMBH Alliance, Heidelberg 69120, Germany, ²Neurobiology and Behavioral Disorders and Centre for Computational Biology, Duke-NUS Graduate Medical School, Singapore 169857, ³Department of Bioinformatics, Institute of Molecular and Cell Biology, University of Tartu, Tartu 51010, Estonia, ⁴National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA and ⁵Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

Received April 27, 2012; Revised May 21, 2012; Accepted May 26, 2012

ABSTRACT

Polymerase chain reaction (PCR) is a basic molecular biology technique with a multiplicity of uses, including deoxyribonucleic acid cloning and sequencing, functional analysis of genes, diagnosis of diseases, genotyping and discovery of genetic variants. Reliable primer design is crucial for successful PCR, and for over a decade, the open-source Primer3 software has been widely used for primer design, often in high-throughput genomics applications. It has also been incorporated into numerous publicly available software packages and web services. During this period, we have greatly expanded Primer3's functionality. In this article, we describe Primer3's current capabilities, emphasizing recent improvements. The most notable enhancements incorporate more accurate thermodynamic models in the primer design process, both to improve melting temperature prediction and to reduce the likelihood that primers will form hairpins or dimers. Additional enhancements include more precise control of primer placement—a change motivated partly by opportunities to use whole-genome sequences to improve primer specificity. We also added features to increase ease of use, including the ability to save and re-use parameter settings and the ability to require that individual primers not be used in more than one primer pair. We have made the core code more modular and provided cleaner programming interfaces to further ease integration with other software. These improvements position Primer3 for

continued use with genome-scale data in the decade ahead.

INTRODUCTION

The design of polymerase chain reaction (PCR) primers, like the laboratory technique of PCR itself, is ubiquitous and diverse. PCR is used for creation of templates for Sanger and next-generation sequencing, detection of the presence or absence of particular deoxyribonucleic acid (DNA) sequences, assessment of the length of simple sequence repeats, creation of constructs for genetic engineering, amplification of complementary DNA to detect splice forms and, using real-time PCR, measuring transcript abundance or DNA copy number. The diversity of PCR applications requires corresponding flexibility in programs for PCR primer design.

The most cited software for primer design in recent years includes PrimerSelect (1), Primer Express (2), Primer Premier (<http://www.premierbiosoft.com/primerdesign/index.html>), the 'OLIGO' software series [most recently OLIGO 7 (3,4)] and Primer3 (5). In addition, Visual oligonucleotide modeling platform (OMP) (6) is noteworthy for its detailed physical modeling of many aspects of PCR primer design, including the effects of modified nucleic acids and PCR buffer additives, as well as a partial ability to simulate PCR reactions.

Primer3, the subject of this article, is a popular group of programs, programming libraries and web interfaces that assist researchers with PCR primer design (Figure 1). The initial version of Primer3, which was released over a decade ago (5), was designed to select primers and hybridization oligonucleotides ('oligos') for the detection of a DNA sequence (i.e. for the design of assays for 'STSS'—sequence tagged sites) and for design of PCR

*To whom correspondence should be addressed. Tel: +65 9857 3213; Fax: +65 6221 8625; Email: steve.rozen@duke-nus.edu.sg

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

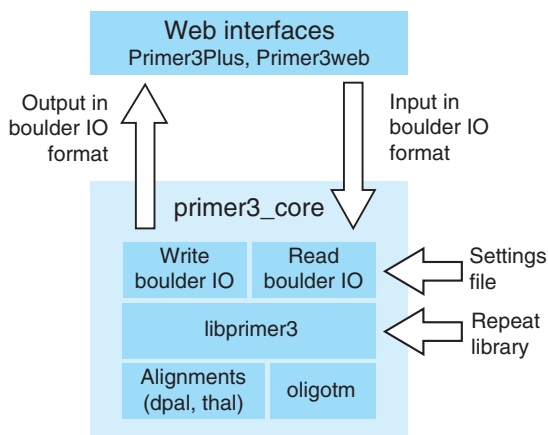


Figure 1. Organization of Primer3. The web interface communicates with primer3_core using the boulder IO format, as described in the text and Figure 4. The primer3_core main program uses a library for reading boulder-IO arguments from the input stream or from a settings file and then calls choose_primers() in the libprimer3 library. Libprimer3 can also read a repeat library, i.e. a library of highly prevalent repeats that primer pairs should avoid amplifying. For example, for the human genome these repeats include Alus, LINEs, endogenous retroviruses and simple sequence repeats (microsatellites, see RepBase, <ftp://ftp.ncbi.nih.gov/repository/repbase/>). Libprimer3 calls several other C/C++ libraries that are part of Primer3. The library for thermodynamic alignments for secondary structures, etc. is called thal. The library for the old-style alignments is dpal. The library for calculating oligo melting temperatures is oligotm. The primer3_core main function returns its results using a library to write boulder-IO format.

assays for microsatellite genotyping. Given a template DNA sequence, the initial version of Primer3 evaluated acceptable primers and primer pairs and then returned the pairs with the lowest penalties. Primer3 consisted of a command-line program that could be integrated into bioinformatic pipelines, whereas a basic web interface, Primer3web (<http://primer3.wi.mit.edu/>), allowed the use of Primer3 through the internet.

Since its initial release, Primer3 has become extremely popular. It was downloaded >20 000 times during 2011, and >7000 manuscripts reference the original publication (Google Scholar). Primer3 likely owes its popularity to several factors that include the availability of a relatively easy-to-use web service, robust engineering, open access to the program source code, suitability for use in high-throughput pipelines for genome-scale research and the ease with which Primer3 can be incorporated into or interoperate with other software. Supplementary Table S1 lists >25 third-party software tools and web services that incorporate Primer3. Primer3's popularity also likely stems from its ability to meet the needs of three groups of users: (i) bioinformaticians, who embed Primer3 within their web services, bioinformatics scripts or pipelines, (ii) expert users who select large numbers of primers and need tight control over the selection process and (iii) occasional users who simply want a convenient way to design a few primer pairs. Bioinformaticians usually rely on the Primer3 command-line interface and run Primer3 on a local machine. The other users, whether 'expert' or 'occasional', usually rely on Primer3's web interfaces.

In light of the multiple uses of PCR, it is not feasible for a single program to meet every need. To overcome this limitation, we envisioned Primer3 as a 'software component' and have encouraged the integration of Primer3 into other software or web services (5). Indeed, Primer3 is unique in the degree to which it has been incorporated into programs and web services tailored to specific primer design needs: multiple software packages and web services have built on Primer3's capabilities by organizing and enhancing Primer3's inputs and outputs to meet specific objectives. We mention several examples in this article and list others in Supplementary Table S1. PRIMEGENS is both a stand-alone program and a web service for high-throughput primer design that offers several integrated functions, which include genome-wide specificity checking and automatic use of gene annotations (7). It also offers dedicated functionality for designing primers to amplify bisulfite-converted DNA (to assess methylation) (8) and for designing specific primers even if highly similar sequences are also present in the template DNA (9). 'The PCR suite' web service (10) integrates Primer3 functionality with GenBank annotations of gene and single-nucleotide polymorphism (SNP) locations and also supports tiling of PCR products across a region. Primer-BLAST is a web service that supports the selection of specific primers by considering opportunities for mispriming across an entire genome or transcriptome (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) (11). The BatchPrimer3 web service (12) allows high-throughput primer design from a list of target sequences and also extends Primer3's selection methods with a focus on SNP typing.

Since we first released Primer3, the fields of bioinformatics, genomics and molecular biology have expanded rapidly, and the amounts and types of sequence data that biologists collect and process are vastly greater and more diverse. In addition, the field has developed more accurate algorithms for predicting the hybridization behavior of nucleotide sequences. These changes have provided opportunities to improve Primer3's functionality:

- The thermodynamic models available for predicting primer melting temperatures, binding and formation of primer dimers and secondary structure have been updated and refined (13–20). Primer3 now offers these up-to-date models, described in detail later.
- Whole-genome sequences are now available for many species. Thus, it is possible in principle to electronically estimate the likely specificity of candidate primer pairs, and one common approach is to search for and then avoid segments of the intended amplification target that are too similar to other sequences (6,7,21). Although we have not added a 'search-and-avoid' capability to Primer3 itself, we have enhanced Primer3, so that it can better support software that uses this approach. This support consists of new ways to precisely constrain primer location, for example, to constrain primers to locations with minimal similarity to other sequences. In addition, for reverse-transcriptase

(RT) PCR, Primer3 can now require primers that span exon-exon junctions.

Finally, the spread of molecular techniques to all corners of biology brings an even wider variety of users to Primer3—from experienced bioinformaticians to high-school students designing and generating their first PCRs. Thus, there has been interest in augmenting Primer3 with interfaces targeted for specific design tasks and in further increasing the ease with which Primer3 can be integrated with other software. This has motivated several enhancements. These include the ability to require that primers not be used in multiple primer pairs, the ability to save and re-use argument settings in both the command line (`primer3_core`) and web interfaces, rationalization of the names of the arguments and a programming library interface for calling Primer3 functionality directly from other programs.

We have greatly expanded Primer3's functionality since the original release. In this article, we detail the major features that improve the capabilities and accuracy of Primer3 for designing primers, simplify users' interaction with Primer3 and simplify the integration of Primer3 with other software.

MATERIALS AND METHODS

Overview of Primer3

The most visible components of Primer3 are the web interfaces, Primer3Plus (<http://primer3plus.com>) (22) and Primer3web (<http://primer3.wi.mit.edu>) and the command-line primer-design program, `primer3_core` (Figure 1). `primer3_core` is mostly useful to bioinformaticians with programming skills and is written with the goal of being computationally efficient, thoroughly testable and easily integrated with other software. Primer3Plus (Figures 2 and 3) and Primer3web offer interfaces more suitable to 'expert' or 'occasional' users. However, `primer3_core` is the engine underlying all primer design operations, and most of the changes described in this article affect both the web interfaces and `primer3_core`, which we collectively refer to as Primer3.

Primer3 can carry out several kinds of design tasks (discussed later) and also check existing primer pairs for correctness. We focus mainly on discussing Primer3 in the context of designing primer pairs for amplifying a DNA template using PCR. To accomplish this task, Primer3 evaluates the primers and primer pairs according to various constraints and sorts acceptable pairs by a penalty function. It uses 'branch and bound' techniques (23) to reduce the search space while still generating the optimal primer pairs according to the penalty function and constraints.

Because the `primer3_core` program accepts >150 parameters and long template sequences—far too much information to be supplied on the command-line—it reads most of its inputs from flat text files. Specifically, `primer3_core` uses the 'boulder IO' (boulder input-output) format, in which each parameter name (termed

a 'tag') and its corresponding value are joined by an '=' sign, with one tag and its value per line (Figure 4). A single input file can specify multiple primer design tasks: the inputs for individual design tasks are separated by lines consisting only of an '=' sign. `primer3_core` also generates its output in boulder IO format (Figure 4B). This format is designed for integration with other software, including scripts and web services, on the assumption that end users will interact with the web interfaces.

Access to Primer3

The most recent `primer3_core` release is available at <http://sourceforge.net/projects/primer3/files/primer3/>.

The Primer3Plus web service is available at <http://primer3plus.com/cgi-bin/dev/primer3plus.cgi>. The most recent Primer3web code is available at <http://sourceforge.net/projects/primer3/files/primer3-web/>, and the Primer3web service is available at <http://primer3.wi.mit.edu/>.

Software organization of `primer3_core`

`primer3_core` consists of a main function and several C++ libraries (Figure 1), including libraries that calculate the melting temperatures of oligos, libraries that calculate the propensity of oligos to form hairpins or dimers or to hybridize or prime from unintended sites in the genome, libraries for reading and writing boulder-IO and the library `libprimer3`, which is a stand-alone C/C++ library that provides Primer3's central primer-design capabilities. When called to select primers, `libprimer3`'s central function, `choose_primers()`, performs an exhaustive search for the 'best' legal primer pairs given the specified template sequence and other input arguments. 'Legal' primers or primer pairs are those that satisfy user-specified constraints, such as minimum and maximum melting temperatures for primers, minimum and maximum product size and so forth. Additional constraints govern possible locations of primers and primer pairs, as we discuss in more detail later. The notion of 'best' primers or primer pairs is operationally defined as minimizing a penalty function, as described in the documentation (http://primer3.sourceforge.net/primer3_manual.htm).

RESULTS

Improved thermodynamic models

We have updated Primer3's thermodynamic models for melting temperature calculations and have added thermodynamic models for estimating the propensity of primers to hybridize with other primers or to hybridize at unintended sites in the template; we have also added thermodynamic models for calculating the stability of potential hairpin structures within primers. All thermodynamic models now use up-to-date nearest-neighbor parameters and salt correction formulae.

Nearest-neighbor thermodynamic models offer the most accurate approach for predicting the energetic stability of DNA structures (19). In brief, these models estimate duplex stability by considering not only the effects of the

Primer3Plus

pick primers from a DNA sequence

[More...](#)
[Help](#)

[Source Code](#)
[About](#)

Load server settings: Default Select primer pairs to detect the given template sequence. Optionally targets and included/excluded regions can be specified.

Task: generic

[Main](#)

[General Settings](#)

[Advanced Settings](#)

[Internal Oligo](#)

[Penalty Weights](#)

[Advanced Sequence](#)

Sequence Id:

Paste template sequence below Or upload sequence file:

```

agttcaacagctatgatccaaatcatataacctgtaaggctaagattaatgtcactgtaattgttcatgtgggaatagc
cagatttcaaaagattatggcttgtttgttacotatccactcaggtctgatgatactcttgcgaaaattgcgaccaaagc
tggcttggatgaaggggtgatcaaaaattcaatcaagatgccaattcagcataggaagtggaatgttccattccag
gaagaggtatgtatctcattttctgccaactgtggttggcacagatggtttgaactctgtcacatccgttgaact
ttgataagctgaaattccgcagttttagattactggtaaatccattataaaatgttaatgtgatttgggtgatttta
tcaaaagtacttgtataagtagtcgagttagataaaaaaattatgacctcttctcgtggaatggactctgataa
ttcataaagtctagccagtgattgtaacaaccaggctttgaacttggacttccaatcaactgacctccaccagacctc
attgaccacttgagtcgaaccctttaaatttcagttagagtataatttaaagctaaagtactctattatttttcaaaagt
atacatggtataaaatttgaagtttatgtatgtattgtttactttgcagatcaaaatggacatttttcccttggatt
ctaggttaagtaacattgattatctcaattttcattttgaaatgattatagaagaagtaaatattgcttcatataaattg
gttatatttttcaacttccattttctttttattttccattcttgcagaacaggtttgtcttttg

```

Mark selected region:

Excluded Regions: < >

Targets: []

Included Region: { }

Primer overlap positions: -

Pair OK Region List:

<input checked="" type="checkbox"/> Pick left primer	<input type="checkbox"/> Pick hybridization probe	<input checked="" type="checkbox"/> Pick right primer
or use left primer below.	(internal oligo) or use oligo below.	or use right primer below (5'→3' on opposite strand).
<input type="text"/>	<input type="text"/>	<input type="text"/>

Figure 2. Input web page for Primer3Plus.

number of hydrogen bonds linking base pairs but also the effects of the stacking of neighboring base pairs along the length of the duplex. For example, because of stacking, the stability of 5'-CT-3' hybridized to 3'-GA-5' is different from that of 5'-CA-3' hybridized to 3'-GT-5', even though the base pairings—C:G and T:A—are the same. To account for the effects of base-pair stacking, nearest-neighbor models capture the thermodynamic stability of overlapping neighboring base pairs rather than considering only single base pairs at a time. For example, nearest-neighbor models calculate the melting temperature of the oligo ACTGCG using the thermodynamic parameters of the overlapping neighboring bases that make up the sequence: AC, CT, TG, GC and CG. These models (and Primer3) also consider specific aspects of the PCR buffer composition—salt and deoxyribonucleotide triphosphate (dNTP) concentrations—that influence the stability of primer binding and potential secondary structures.

Using these thermodynamic models, Primer3 calculates the melting temperatures of oligos when hybridized to

their intended targets and also predicts the strength of three types of secondary structures with possible mismatches: two types of bimolecular hybridization ('ANY' and 'END' interactions) and one type of unimolecular folding. Bimolecular ANY interactions involve the binding of an oligo (a primer or a hybridization oligo), anywhere within its sequence, to another DNA molecule (Figure 5A and B). Bimolecular END interactions involve the binding of the 3'-end of a primer to another single-stranded DNA molecule (Figure 5C and E). Some, but not all, END interactions are also ANY interactions, and vice versa, and the most stable ANY interaction is always at least as stable as the most stable END interaction. END interactions are most relevant to priming polymerases activity. Primer3 filters out primers with stable END interactions to avoid those prone to generating short 'primer-dimers' (Figure 5C). Primer3 also uses a thermodynamic model to assess the propensity of a primer or hybridization oligo to form unimolecular hairpin structures (Figure 5D). Because these can inhibit

Primer3Plus

pick primers from a DNA sequence

[More...](#)
[Help](#)

[Source Code](#)
[About](#)

[< Back](#)

Pair 1: Primer

Left Primer 1: tggtcatggggaatgccaga

Start: 63 Length: 22 bp Tm: 59.4 C GC: 45.5 % Any: 0.0 End: 0.0 TB: 0.0 HP: 36.4 3' Stab: 3.9 Penalty: 2.576

Right Primer 1: gggttcgactcaagtggtca

Start: 582 Length: 20 bp Tm: 59.6 C GC: 55.0 % Any: 0.0 End: 0.0 TB: 0.0 HP: 0.0 3' Stab: 4.0 Penalty: 0.392

Pair: Product Size: 520 bp Any: 0.0 End: 0.0 Penalty: 2.967

[Send to Primer3Manager](#) [Reset Form](#)

```

1      agttcaacag  ctatgatcca  aatcatatac  ctgttaaggc  taagattaat
51     gtaactgtaa  ttgttcatg  tgggaatagc  cagatttcaa  aagattatgg
101    cttgtttgtt  acctatccac  tcaggcttga  tgatactctt  gcgaaaattg
151    cgaccaaagc  tegtcttgat  gaagggttga  tacaaaattt  caatcaagat
201    gccaatttca  gcataggaag  tgggatagtg  ttcattccag  gaagagggat
251    gtattttctc  attttctgcc  aactgtggtt  ggcacagatg  gtttgaactt
301    ctgtcacatc  cgttgtaact  ttgataagtc  tgaattccg  cagttttag
351    attactggtt  aattccatta  taaatgttta  atgtgatttg  gtgattctta
401    tcaaaagtac  ttgtataagt  atgcgagtta  gataaaaaaa  attatgacca
451    tcttgttctc  gtgaaatgg  actctgataa  tcataaaagt  ctagccagtg
501    attgtaacaa  ccaggctttg  aacttggtac  ttccaatcaa  cttgaccttc
551    accagacctc  attgaccact  tgagtccaac  cctttaattt  cagtttagag
601    atatttaaat  gctaagttac  tctattatgt  ttcaaagtat  atacatggta
651    taaattttga  agttttatgt  agttattggt  tactttgcag  atcaaatgg
701    acattttttt  cctttgtatt  ctaggtaagt  aacattgatt  atctcaattt
751    tcatttttga  atgatttata  gaagaagtaa  atattgcttc  atataatttg
801    gttatatttt  tctaactttc  attttctttt  tatttttcca  ttcttcgaga
851    acaggtttgt  cttttg
                    
```

Select all Primers

Pair 2: Primer_1

Left Primer 2: tggttgccacagatggt

Start: 274 Length: 20 bp Tm: 60.1 C GC: 50.0 % Any: 0.0 End: 0.0 TB: 0.0 HP: 38.6 3' Stab: 3.7 Penalty: 0.106

Right Primer 2: acctgttcgaagaatgga

Start: 856 Length: 20 bp Tm: 57.0 C GC: 45.0 % Any: 14.4 End: 0.0 TB: 14.4 HP: 35.4 3' Stab: 3.4 Penalty: 2.957

Pair: Product Size: 583 bp Any: 0.0 End: 0.0 Penalty: 3.064

[Send to Primer3Manager](#) [Reset Form](#)

■ ■ ■

Pair 5: Primer_4

Left Primer 5: tcatggggaatgccagattca

Start: 66 Length: 24 bp Tm: 60.1 C GC: 41.7 % Any: 0.0 End: 0.0 TB: 0.0 HP: 36.4 3' Stab: 2.7 Penalty: 4.080

Right Primer 5: gltgcaatgaggtctggtga

Start: 569 Length: 21 bp Tm: 59.7 C GC: 52.4 % Any: 0.0 End: 0.0 TB: 0.0 HP: 0.0 3' Stab: 4.0 Penalty: 1.348

Pair: Product Size: 504 bp Any: 0.0 End: 0.0 Penalty: 5.428

[Send to Primer3Manager](#) [Reset Form](#)

Statistics:	
Left Primer:	considered 4528, GC content failed 137, low tm 2752, high tm 248, high hairpin stability 2, long poly-x seq 24, ok 1365
Right Primer:	considered 4370, GC content failed 1003, low tm 2469, high tm 125, long poly-x seq 47, ok 726
Primer Pair:	considered 14682, unacceptable product size 14677, primer in pair overlaps a primer in a better pair 1833, ok 5

Figure 3. Example output web page from Primer3Plus. The first major block, labeled 'Pair 1', shows the template sequence with locations of the primers highlighted in blue and yellow and key information about the primers and primer pair. Subsequent blocks ('Pair 2',..., 'Pair 5') show information for alternative primer pairs.

A

```

PRIMER_TASK=generic
PRIMER_NUM_RETURN=1
PRIMER_PAIR_MAX_TEMPLATE_MISPRIMING_TH=80.00
PRIMER_PRODUCT_SIZE_RANGE=75-200
PRIMER_EXPLAIN_FLAG=1
PRIMER_MISPRIMING_LIBRARY=humrep_and_simple.txt
PRIMER_OPT_SIZE=20
PRIMER_MIN_SIZE=19
PRIMER_MAX_SIZE=21
SEQUENCE_EXCLUDED_REGION=37,21
SEQUENCE_ID=example1
SEQUENCE_TEMPLATE=GTAGTCAGTAGACATGACACTGACGATGCAGACNACACACACACAAGCACAC...TGACG
SEQUENCE_TARGET=37,21
=
PRIMER_PRODUCT_SIZE_RANGE=75-100
SEQUENCE_ID=example2
SEQUENCE_TEMPLATE=GANAATGCAGTAATGAGTT...TCAAAGTCTATTCCAAAACGCAAAGCACGCTCCCGATC
=

```

File containing sequences of common human repeats such as Alus and LINES

End of input for 1st primer design task

This changes the allowable product size range from the previous value, while the values of other arguments starting PRIMER_... remain unchanged

B

... echoed input, not shown ...

```

PRIMER_LEFT_EXPLAIN=...
PRIMER_RIGHT_EXPLAIN=considered 87, low tm 33, high tm 12, ok 42
PRIMER_PAIR_EXPLAIN=considered 1, ok 1
PRIMER_LEFT_NUM_RETURNED=1
PRIMER_RIGHT_NUM_RETURNED=1
PRIMER_INTERNAL_NUM_RETURNED=0
PRIMER_PAIR_NUM_RETURNED=1
PRIMER_PAIR_0_PENALTY=1.306831
PRIMER_LEFT_0_PENALTY=1.273327
PRIMER_RIGHT_0_PENALTY=0.033504
PRIMER_LEFT_0_SEQUENCE=AGACATGACACTGACGATGCA
PRIMER_RIGHT_0_SEQUENCE=ATCGATTTGGGTCGGGATCG

```

... more information on the primer pair designed for the first template, not shown ...

```

PRIMER_RIGHT_0_TEMPLATE_MISPRIMING_TH=6.2638
PRIMER_PAIR_0_COMPL_ANY_TH=0.00
PRIMER_PAIR_0_COMPL_END_TH=0.00
PRIMER_PAIR_0_PRODUCT_SIZE=84
PRIMER_PAIR_0_TEMPLATE_MISPRIMING_TH=6.26
=
PRIMER_PRODUCT_SIZE_RANGE=75-100
SEQUENCE_ID=example2
SEQUENCE_TEMPLATE=...
PRIMER_LEFT_EXPLAIN=...
PRIMER_RIGHT_EXPLAIN=considered 216, low tm 167, high tm 13, ok 36
PRIMER_PAIR_EXPLAIN=considered 6, unacceptable product size 5, ok 1
PRIMER_LEFT_NUM_RETURNED=1
PRIMER_RIGHT_NUM_RETURNED=1

```

End of output for first primer design task

... more information on the primer pair designed for second template, not shown ...

=

Figure 4. Examples of boulder IO (A) input to primer3_core and (B) output from primer3_core. Panel A shows two primer design tasks (SEQUENCE_ID's example 1 and example 2) separated by a line consisting only of an = character. PRIMER_NUM_RETURN is 1 for both tasks, because the values of tags beginning with PRIMER_... persist between design tasks. However, the value of tags beginning with SEQUENCE_... such as SEQUENCE_EXCLUDED_REGION = 37,21 in example 1, affect only the current design task. Panel B shows abbreviated output corresponding to panel A.

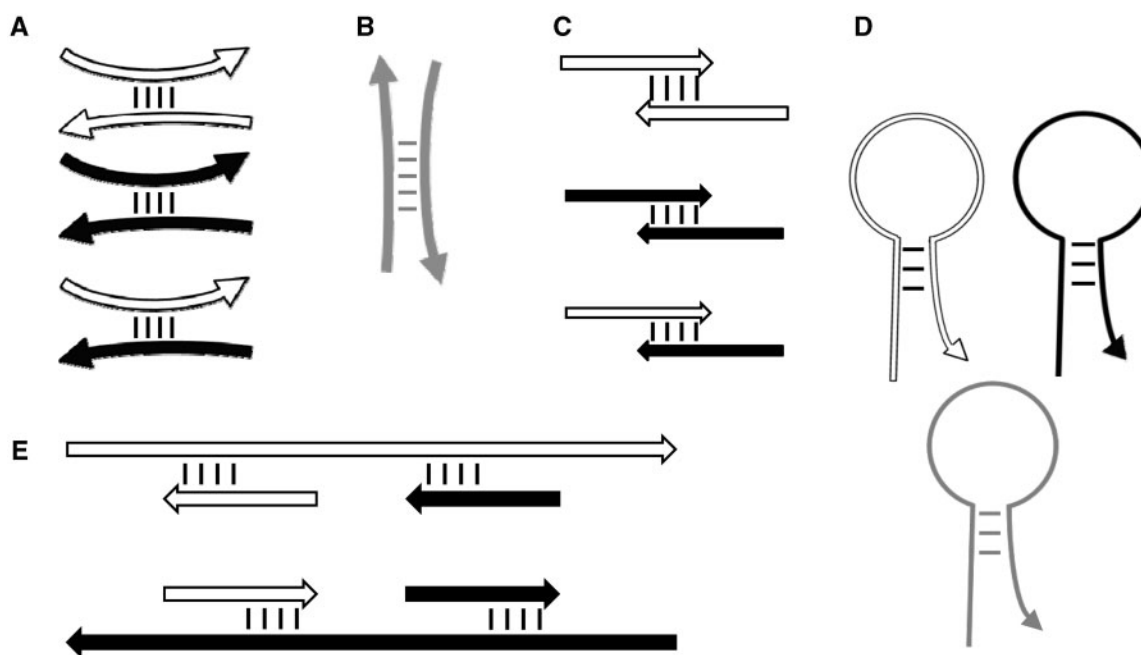


Figure 5. Primer3's thermodynamic models for predicting the stability of secondary structures. (A–C) Interactions between primers. (D) Hairpin structures. (E) Undesirable binding of primers to template sequence. Short white and black arrows represent forward and reverse primers, respectively. Gray arrows represent hybridization oligos. Long white and black arrows represent forward and reverse template. Curved oligos (including primers) indicate ANY interactions (A, B and D), and straight oligos indicate END (i.e. 3'-anchored) interactions (C and E). Primer3_core and the Primer3 web interfaces now use thermodynamic models by default. The Primer3 manual (http://primer3.sourceforge.net/primer3_manual.htm) and Supplementary Figure S1 document the arguments governing temperature calculations for secondary structures.

primer hybridization to template molecules and cause PCR failure (20), Primer3 filters out primers likely to form hairpins.

Primer3 expresses the stability of ANY and END bimolecular interactions and of hairpin structures as melting temperatures—the temperatures at which there is a 50% possibility that a molecule is in a single-stranded form and a 50% possibility that a molecule has formed a duplex. By default, Primer3 requires these temperatures to be substantially lower than the minimum primer melting temperature.

The methods for calculating melting temperatures for ANY and END bimolecular interactions and of hairpin structures are extensions of those used to calculate melting temperatures for oligos when hybridized without mismatches to their intended target; in addition to considering the stability of neighboring base pairs, the secondary structure models also take into consideration the thermodynamic consequences of mismatched base pairs, gaps (physically, bulges) and loops. The relevant parameters are taken from the following: (13–20).

The Primer3 package also provides two stand-alone command-line programs: *oligotm* calculates oligo melting temperatures and *ntthal* (nucleotide thermodynamic alignment) predicts the propensity of oligos to form dimers or hairpins.

Controlling primer location

The availability of whole-genome sequences from many species (24) offers the possibility of using an organism's genome sequence to improve primer-pair specificity. One

approach to increasing PCR specificity is to search the intended target sequence against the rest of the genome and then restrict primers to regions within the target sequence that are minimally similar to other sequences in the genome. This approach can also be applied to primers for RT-PCR, by searching the intended target sequence against the entire transcriptome. This 'search-and-avoid' approach is used by GENOMEMASKER (21) and is one of the techniques used by Primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>)(11).

Furthermore, if almost the entire intended template is similar to another region of the genome, we might be satisfied if either the forward or reverse primer is in a region having minimal similarity to other regions of the genome. To flexibly support these situations, Primer3 now offers the argument `SEQUENCE_PRIMER_PAIR_OK_REGION_LIST`. This argument takes a list of region pairs and demands that the selected forward and reverse primers be in regions specified by at least one of the pairs in the list. For example, the list in Figure 6 specifies that there are three, equally acceptable alternatives for the locations of the forward and reverse primers:

- Forward primer in the 50 bp region starting at position 100 and reverse primer in the 50 bp region starting at position 300.
- Forward primer in the 50 bp region starting at position 400 and reverse primer anywhere.
- Forward primer anywhere and reverse primer in the 30 bp region starting at position 460.

An additional new argument in Primer3 supports a common requirement for RT-PCR, in which, to lessen

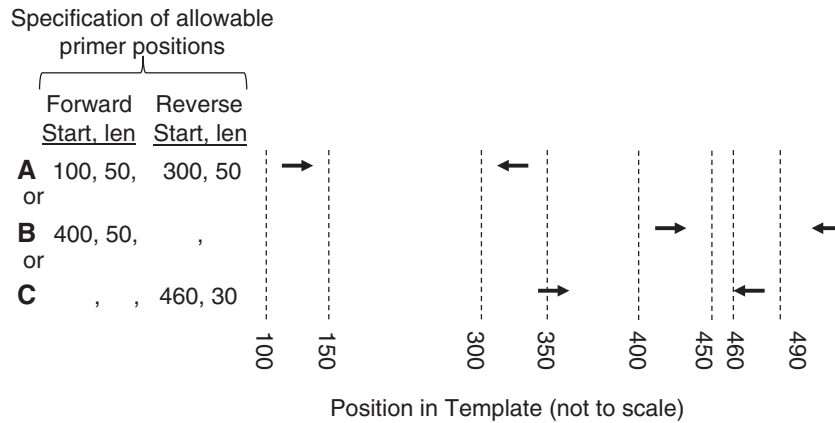


Figure 6. Constraining the positions of forward and reverse primers using `SEQUENCE_PRIMER_PAIR_OK_REGION_LIST`. Arrows represent primers. In this example, primer pairs A, B and C are all allowed. In the specification of allowable primer positions, blanks for ‘Start’ and ‘len’ indicate no constraint (example B, reverse primers, example C, forward primers).

the likelihood of amplifying genomic DNA or unspliced transcripts, it is often desirable to have one of the primers fall across the junction of two exons. To specify this, Primer3 offers the argument `SEQUENCE_OVERLAP_JUNCTION_LIST`. The value of this argument is a list of positions (e.g. of exon–exon junctions), and either the forward or the reverse primer must overlap one of these positions. By default, the minimum overlaps default to four bases at the 3′-end of the primer and seven bases at the 5′-end of the primer.

Primer3 retains the other, previously available arguments for constraining the locations of primers and hybridization oligos. These include `SEQUENCE_INCLUDED_REGION`, `SEQUENCE_EXCLUDED_REGION`, `SEQUENCE_TARGET` and others. Please see the documentation (http://primer3.sourceforge.net/primer3_manual.htm) for details regarding the use of these arguments.

Usability improvements

Unique primers

The current version of Primer3 remedies a common source of frustration for users of earlier versions: Primer3 often re-used the same forward or reverse primer in multiple primer pairs (Figure 7). In other words, the same primer was used several times in combination with a different mate. Thus, although primer pairs were unique, the primers within the pairs were not. Primer3 now allows the selection of primer pairs where each individual primer is used only once. In addition, one can guarantee that the multiple primer pairs returned by Primer3 are sufficiently different from each other by using the tag `PRIMER_MIN_THREE_PRIME_DISTANCE`. This tag ensures that the 3′-end of a forward (or reverse) primer has a minimum distance from the 3′-end of the forward (respectively, reverse) primer in any other primer pair (Figure 7).

Rationalization and documentation of boulder IO ‘tags’

For easy integration of primer3_core with other software, the boulder IO input tags must be straightforward to

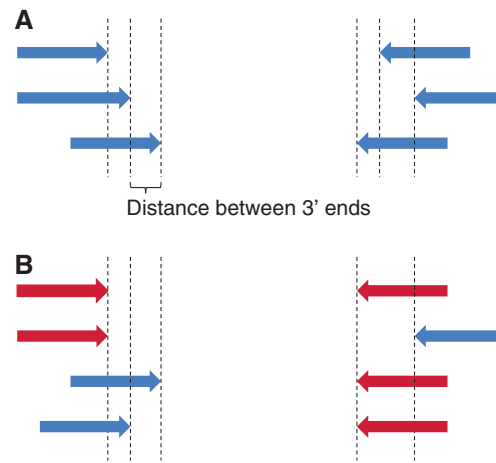


Figure 7. Unique and non-unique primers. (A) The 3′-end of each forward and reverse primer has a unique position; Primer3 can now ensure this and can ensure a minimum distance between the 3′-ends of all forward and/or all reverse primers. (B) Each primer pair is unique, but the red primers are used more than once.

create, and the output tags must be straightforward to interpret by other programs or scripts. To this end, we have systematized tag names throughout the program. In particular, tags providing information specific to the current template sequence start with ‘`SEQUENCE_`’ (Figure 4A) and only have effect for a single primer design task; tags providing global information required for primer selection start with ‘`PRIMER_`’ and are active until they are over-ridden. We have also systematized the structure of the output tags to ease automatic parsing.

We have updated Primer3’s documentation (http://primer3.sourceforge.net/primer3_manual.htm) to explain each input and output tag in detail and to explain the details of the penalty calculations, which are governed by tags. The documentation also provides a table detailing the correspondences between old and new tag names.

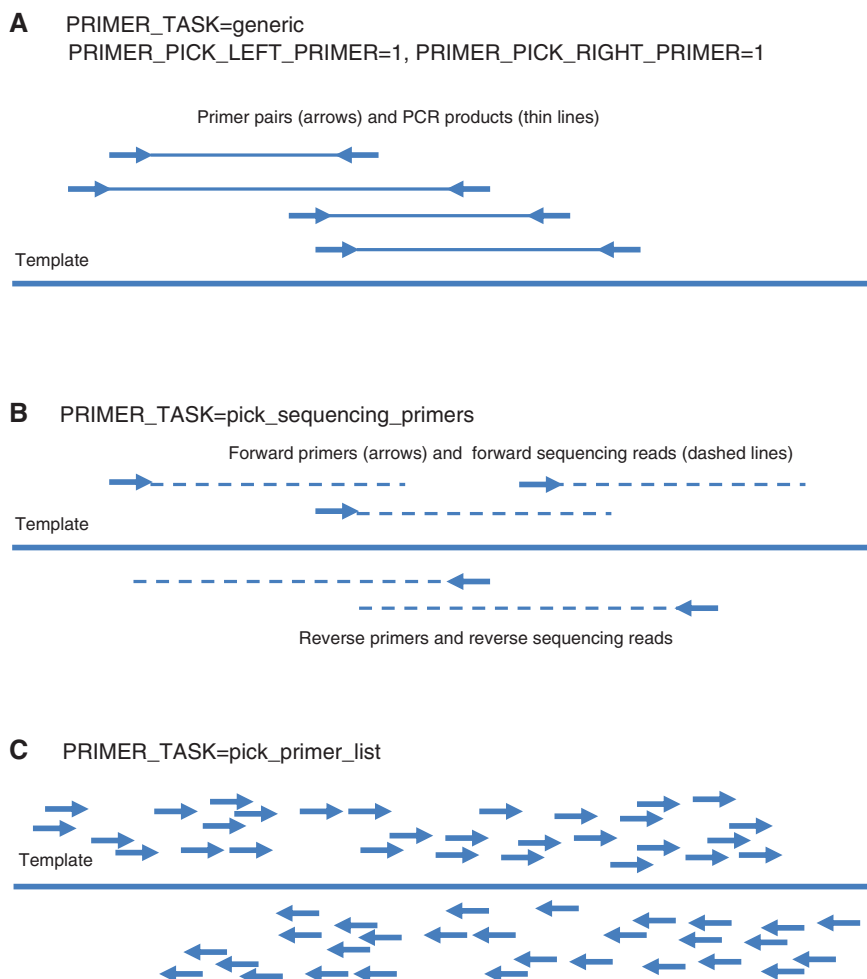


Figure 8. Three types of Primer3 task. Arrows represent PCR primers with 3'-ends at the arrow heads. Thick solid lines represent the template sequence. (A) The task *generic* directs Primer3 to design primer pairs for PCR when both forward (left) and reverse (right) primers are requested. Thin solid lines represent PCR products. (B) The task *pick_sequencing_primers* directs Primer3 to pick multiple unpaired primers for Sanger sequencing reads that will cover the template. Thin dashed lines represent Sanger sequencing reads. (C) The task *pick_primer_list* directs Primer3 to generate lists of legal forward and reverse primers without regard to their positions relative to each other.

Reusable argument settings

Because the behavior of Primer3 is governed by >150 arguments, which often require tuning for specific tasks, we have added the ability to store and re-use argument values in 'settings files'. These are text files containing a header followed by boulder-IO-formatted input that sets the values of 'global' tags (input tags prefixed by 'PRIMER_'). Settings files offer a means to save and re-use argument settings developed for a particular application and a means for expert users to provide less-experienced users with parameter values suitable for their needs. Settings files can also be provided in publications as supplemental data to improve experimental reproducibility. For users of the web interfaces (Primer3Plus and Primer3web), settings files have the critical function of easing repetitive primer selections by obviating the need to manually change multiple parameters at the start of each design session. Users of the web interfaces can also download their current parameter values to a settings file and later, re-use these parameter values by uploading

them back to the web interface or by using them on the command line with `primer3_core`.

Design tasks

To make Primer3 easier to use, we have re-factored the set of 'tasks' that Primer3 can carry out to include the following (Figure 8):

generic: This task includes the standard design of primer pairs for PCR amplifying a DNA template (Figure 8A). When this task is specified, Primer3 can optionally design a hybridization oligo within the amplified region.

check_primers: Check existing primers or primer pairs, even in the absence of a template sequence. Primer3 assesses whether the primer or primer pair satisfies the user specified constraints and also returns additional measures describing aspects of the primers and primer pairs (primer melting temperature, propensity to form secondary structure, product size, etc.)

pick_sequencing_primers: Pick multiple unpaired primers for Sanger sequencing reads that will cover the template (Figure 8B). The spacing between successive forward primers and successive reverse primers is governed by arguments (described in the manual, http://primer3.sourceforge.net/primer3_manual.htm), but as the primers are designed to prime Sanger sequencing reactions, forward and reverse primers are not paired. This is different from selecting primer pairs to 'tile' a large region with PCR products for subsequent sequencing, which Primer3 does not offer [although the PCR Suite (10) layers this capability on top of Primer3.]

pick_primer_list: Generate lists of legal forward and reverse primers without regard to their positions relative to each other (Figure 8C). This is approximately a re-implementation of the capability governed by P3_FILE_FLAG, which directs Primer3 to generate files containing all acceptable primers evaluated. These files were used by many bioinformatics scripts that further evaluated the lists of primers and then selected pairs according to the needs of the particular application. Primer3 can now provide these primer lists on the output stream in boulder-IO format. Thus, applications that use primer3_core to generate these lists no longer need to manage temporary files.

The Primer3 manual (http://primer3.sourceforge.net/primer3_manual.htm) describes a few additional, more specialized, tasks.

Programming library interface to Primer3

Primer3 now offers a C/C++ programming library, named libprimer3, that offers the same primer-design functionality as the command line interface (Figure 1). The rationale is that some third-party programs using Primer3 functionality will benefit from the close integration offered by a library function-call interface. For example, the NCBI Primer-BLAST service (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) (11) powers part of their primer design service with libprimer3.

Primer3Plus—an improved web interface

Primer3 has two web interfaces—the original web interface, Primer3web, and the newer Primer3Plus (22), which is designed to be an easier-to-use, task-oriented interface that still harnesses the full power of Primer3. Primer3Plus offers several tabs from which all Primer3 parameters are accessible, including the new parameters controlling the thermodynamic models for secondary structure and those providing precise control over possible primer locations (Figure 2). We have set the parameters in Primer3Plus to reasonable default values, which should allow occasional users to design working primers by simply providing a template DNA sequence in the first tab of the web interface and clicking the 'pick primers' button. Experienced users can adjust arguments presented through the remaining tabs to exert fine control over the primer selection process.

When the user clicks the 'pick primer' button, the web application sends the template sequence and design parameters to primer3_core for primer selection (Figure 1). Primer3_core selects primer pairs and returns the boulder IO-formatted results to Primer3Plus, which parses these results and presents them to the user in graphical format (Figure 3). Users can then optionally store the primers using Primer3Manager, a simple tool for managing primer collections. We have updated Primer3Manager to support the extensible markup language (XML) based Real-time PCR Data Markup Language (RDML) format for primer information storage and exchange (25).

New search algorithm

During the implementation of some of the new features described in this article, it became obvious that the primer search algorithm needed to be re-designed. The original algorithm looked for the best primer pairs by considering legal combinations of reverse and forward primers. It used a branch and bound approach based on the penalties associated with the individual forward and reverse primers. However, the algorithm inefficiently reevaluated possible pairs every time it had to find the next best pair. In the new implementation, to speed up the search, each pair is evaluated at most once, and then its characteristics are stored in memory. This increase in computational efficiency comes at the expense of higher memory usage. To minimize the amount of memory used, the new algorithm stores characteristics of primer pairs in hash maps and, therefore, allocates memory for a pair only when necessary. The algorithm also postpones computing some expensive-to-compute characteristics of primers until it is known that the primer might be part of a legal primer pair. These characteristics include prediction of the primer's propensity to form hairpins or hybridize to other oligos or ectopic sites in the template and the primer's propensity to amplify known high-abundance genomic repeats.

DISCUSSION

Over the last decade, Primer3 has proven to be a robust, versatile and widely used tool for primer selection. Evidence supporting these assertions stems both from the abundant use of Primer3 directly and from the many software programs and web services that rely on Primer3. The changing landscapes of molecular biology, genetics, genomics and bioinformatics have changed Primer3's uses. These changes, plus experience with continued use of the program, indicated several opportunities for improvement of Primer3. We have augmented Primer3's functionality and usability to take advantage of these opportunities. New capabilities include updated thermodynamic models for predicting oligo melting temperatures and thermodynamic models for predicting the propensity of oligos to form hairpin structures, to hybridize to each other or to hybridize to undesirable sites in template sequences. The new capabilities also include the ability to more precisely control the locations of primers. Additional enhancements improve ease of use and of integration with

other software. These include the ability to require that primers not be used in multiple primer pairs, the ability to save and re-use argument settings, rationalization of the names of the boulder-IO arguments to primer3_core and a C/C++ programming library interface for calling Primer3 functionality directly from other programs. We are confident that, with the updates and improvements of the last several years, Primer3 and its web interfaces will remain useful and will serve as starting points for future primer design applications.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figure 1 and Supplementary references [26–49].

ACKNOWLEDGEMENTS

The authors thank Fran Lewitter, John R. McPherson, Choon Kiat Ong, Kathryn Rozen-Gagnon, Helen Skaletsky and Zhejiang Zang for comments on the manuscript. They thank Whitehead Institute for hosting Primer3web and Fran Lewitter and Dennis Ristuccia for installing and testing Primer3web at Whitehead Institute.

FUNDING

The Singapore Ministry of Health (to S.G.R.); the Singapore Agency for Science, Technology, and Research (to S.G.R.); the E.U. European Regional Development Fund through the Estonian Centre of Excellence in Genomics (to M.R. and T.K.); the Estonian Ministry of Education and Research [SF0180026s09 to M.R. and T.K.]; National Science Foundation of the U.S.A. [DEB-1136626 to B.C.F.]; Intramural Research Program of the NIH, National Library of Medicine (to J.Y.). Funding for open access charge: The Singapore Ministry of Health (to S.G.R.); the Singapore Agency for Science, Technology, and Research (to S.G.R.).

Conflict of interest statement. None declared.

REFERENCES

- Plasterer, T.N. (1997) PRIMERSELECT. Primer and probe design. *Methods Mol. Biol.*, **70**, 291–302.
- Applied Biosystems. (2004) Primer Express[®] Software Version 3.0 Getting Started Guide.
- Rychlik, W. (2007) OLIGO 7 primer analysis software. *Methods Mol. Biol.*, **402**, 35–60.
- Rychlik, W. (2010) OLIGO Primer Analysis Software Version 7.
- Rozen, S. and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz, S. and Misener, S. (eds.), *Bioinformatics Methods and Protocols*. Humana Press, Totowa, NJ, pp. 365–386.
- SantaLucia, J. Jr (2007) Physical principles and visual-OMP software for optimal PCR design. *Methods Mol. Biol.*, **402**, 3–34.
- Srivastava, G.P. and Xu, D. (2007) Genome-scale probe and primer design with PRIMEGENS. *Methods Mol. Biol.*, **402**, 159–176.
- Srivastava, G.P., Guo, J., Shi, H. and Xu, D. (2008) PRIMEGENS-v2: genome-wide primer design for analyzing DNA methylation patterns of CpG islands. *Bioinformatics*, **24**, 1837–1842.
- Srivastava, G.P., Hanumappa, M., Kushwaha, G., Nguyen, H.T. and Xu, D. (2011) Homolog-specific PCR primer design for profiling splice variants. *Nucleic Acids Res.*, **39**, e69.
- van Baren, M.J. and Heutink, P. (2004) The PCR suite. *Bioinformatics*, **20**, 591–593.
- Sayers, E.W., Barrett, T., Benson, D.A., Bolton, E., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., Dicuccio, M., Federhen, S. et al. (2012) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **40**, D13–D25.
- You, F.M., Huo, N., Gu, Y.Q., Luo, M.C., Ma, Y., Hane, D., Lazo, G.R., Dvorak, J. and Anderson, O.D. (2008) BatchPrimer3: a high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics*, **9**, 253.
- Allawi, H.T. and SantaLucia, J. Jr (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry*, **36**, 10581–10594.
- Allawi, H.T. and SantaLucia, J. Jr (1998) Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry*, **37**, 9435–9444.
- Allawi, H.T. and SantaLucia, J. Jr (1998) Thermodynamics of internal C.T mismatches in DNA. *Nucleic Acids Res.*, **26**, 2694–2701.
- Allawi, H.T. and SantaLucia, J. Jr (1998) Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry*, **37**, 2170–2179.
- Bommarito, S., Peyret, N. and SantaLucia, J. Jr (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.*, **28**, 1929–1934.
- Peyret, N., Seneviratne, P.A., Allawi, H.T. and SantaLucia, J. Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry*, **38**, 3468–3477.
- SantaLucia, J. Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA.*, **95**, 1460–1465.
- SantaLucia, J. Jr and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 415–440.
- Andreson, R., Reppo, E., Kaplinski, L. and Remm, M. (2006) GENOMEMASKER package for designing unique genomic PCR primers. *BMC Bioinformatics*, **7**, 172.
- Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.*, **35**, W71–W74.
- Land, A.H. and Doig, A.G. (1960) An automatic method of solving discrete programming problems. *Econometrica*, **28**, 497–520.
- Genome 10K Community of Scientists. (2009) Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J. Hered.*, **100**, 659–674.
- Lefever, S., Hellemans, J., Pattyn, F., Przybylski, D.R., Taylor, C., Geurts, R., Untergasser, A. and Vandesompele, J. (2009) RDML: structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic Acids Res.*, **37**, 2065–2069.
- McKay, S.J. and Jones, S.J. (2002) AcePrimer: automation of PCR primer design based on gene structure. *Bioinformatics*, **18**, 1538–1539.
- Wrobel, G., Kokocinski, F. and Lichter, P. (2004) AutoPrime: selecting primers for expressed sequences. *Genome Biol.*, **5**, P11.
- You, F.M., Huo, N., Gu, Y.Q., Lazo, G.R., Dvorak, J. and Anderson, O.D. (2009) ConservedPrimers 2.0: a high-throughput pipeline for comparative genome referenced intron-flanking PCR primer design and its application in wheat SNP discovery. *BMC Bioinformatics*, **10**, 331.
- Wu, X. and Munroe, D.J. (2006) EasyExonPrimer: automated primer design for exon sequences. *Appl. Bioinformatics*, **5**, 119–120.
- Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
- Shen, Z., Qu, W., Wang, W., Lu, Y., Wu, Y., Li, Z., Hang, X., Wang, X., Zhao, D. and Zhang, C. (2010) MPprimer: a program

- for reliable multiplex PCR primer design. *BMC Bioinformatics*, **11**, 143.
32. Li, L.C. and Dahiya, R. (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics*, **18**, 1427–1431.
33. Faircloth, B.C. (2008) msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Resour.*, **8**, 92–94.
34. Koressaar, T., Joers, K. and Remm, M. (2009) Automatic identification of species-specific repetitive DNA sequences and their utilization for detecting microbial organisms. *Bioinformatics*, **25**, 1349–1355.
35. Brown, A.M., Lo, K.S., Guelpa, P., Beaudoin, M., Rioux, J.D., Tardif, J.C., Phillips, M.S. and Lettre, G. (2010) Optimus primer: a PCR enrichment primer design program for next-generation sequencing of human exonic regions. *BMC Res. Notes*, **3**, 185.
36. Gadberry, M.D., Malcomber, S.T., Doust, A.N. and Kellogg, E.A. (2005) Primaclade—a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics*, **21**, 1263–1264.
37. Xu, D., Li, G., Wu, L., Zhou, J. and Xu, Y. (2002) PRIMEGENS: robust and efficient design of gene-specific probes for microarray analysis. *Bioinformatics*, **18**, 1432–1437.
38. Pessoa, A.M., Pereira, S. and Teixeira, J. (2010) PrimerIdent: a web based tool for conserved primer design. *Bioinformation*, **5**, 52–54.
39. Tsai, M.F., Lin, Y.J., Cheng, Y.C., Lee, K.H., Huang, C.C., Chen, Y.T. and Yao, A. (2007) PrimerZ: streamlined primer design for promoters, exons and human SNPs. *Nucleic Acids Res.*, **35**, W63–W65.
40. Meglec, E., Costedoat, C., Dubut, V., Gilles, A., Malausa, T., Pech, N. and Martin, J.F. (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics*, **26**, 403–404.
41. Kim, N. and Lee, C. (2007) QPRIMER: a quick web-based application for designing conserved PCR primers from multigenome alignments. *Bioinformatics*, **23**, 2331–2333.
42. Arvidsson, S., Kwasniewski, M., Riano-Pachon, D.M. and Mueller-Roeber, B. (2008) QuantPrime—a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics*, **9**, 465.
43. Piriyaongsa, J., Ngamphiw, C., Assawamakin, A., Wangkumhang, P., Suwannasri, P., Ruangrit, U., Agavatpanitch, G. and Tongsima, S. (2009) REXPrimer: an integrated primer designing tool increases PCR effectiveness by avoiding 3' SNP-in-primer and mis-priming from structural variation. *BMC Genomics*, **10**(Suppl 3), S4.
44. Zhang, R., Zhu, Z., Zhu, H., Nguyen, T., Yao, F., Xia, K., Liang, D. and Liu, C. (2005) SNP Cutter: a comprehensive tool for SNP PCR-RFLP assay design. *Nucleic Acids Res.*, **33**, W489–W492.
45. Weckx, S., De Rijk, P., Glassee, W., Van Broeckhoven, C. and Del-Favero, J. (2007) SNPbox: web-based high-throughput primer design with an eye for repetitive sequences. *Methods Mol. Biol.*, **402**, 179–200.
46. Weckx, S., De Rijk, P., Van Broeckhoven, C. and Del-Favero, J. (2005) SNPbox: a modular software package for large-scale primer design. *Bioinformatics*, **21**, 385–387.
47. Andreson, R., Puurand, T. and Remm, M. (2006) SNPmasker: automatic masking of SNPs and repeats across eukaryotic genomes. *Nucleic Acids Res.*, **34**, W651–W655.
48. Vijaya Satya, R., Kumar, K., Zavaljevski, N. and Reifman, J. (2010) A high-throughput pipeline for the design of real-time PCR signatures. *BMC Bioinformatics*, **11**, 340.
49. Martins, W.S., Lucas, D.C., Neves, K.F. and Bertioli, D.J. (2009) WebSat—a web software for microsatellite marker development. *Bioinformation*, **3**, 282–283.