Metadata of the chapter that will be visualized online

Chapter Title	FastPCR Software for PCR, In Silico PCR, and Oligonucleotide Assembly and Analysis					
Copyright Year	2014					
Copyright Holder	Springer Science+Business Media, New York					
Corresponding Author	Family Name	Kalendar				
	Particle					
	Given Name	Ruslan				
	Suffix					
	Division	MTT/BI Plant Genomics Laboratory				
	Organization	Institute of Biotechnology, University of Helsinki				
	Address	65, 00014, Helsinki, Finland				
	Organization	PrimerDigital Ltd				
	Address	00790, Helsinki, Finland				
Author	Family Name	Lee				
	Particle					
	Given Name	David				
	Suffix					
Author	Family Name	Schulman				
	Particle					
	Given Name	Alan H.				
	Suffix					
Abstract	This chapter introduces environment for PCR oligonucleotide proper efficiency. The software primers for most PCR standard, multiplex, lo unique, and overlap ex cloning, as well as h program to design olig by the ligase chain re search includes compri- primer pairs. It calcula degenerate oligonucleo provides analyses for a properties, dimer and complexity, and provid program includes variou with CG or AT skew, and of linguistic seque	s the software FastPCR as an integrated tools primer and probe design. It also predicts ties based on experimental studies of PCR provides comprehensive facilities for designing applications and their combinations, including ng-distance, inverse, real-time, group-specific, xtension PCR for multi-fragment assembly in pisulphite modification assays. It includes a gonucleotide sets for long sequence assembly eaction. The in silico PCR primer or probe rehensive analyses of individual primers and ates the melting temperature for standard and tides including LNA and other modifications, set of primers with prediction of oligonucleotide d G/C-quadruplex detection, and linguistic les a dilution and resuspension calculator. The us bioinformatics tools for analysis of sequences of CG content and purine–pyrimidine skew, ence complexity. It also permits generation of				

	random DNA sequence and analysis of restriction enzymes of all
	types. It finds or creates restriction enzyme recognition sites for
	coding sequences and supports the clustering of sequences. It generates
	consensus sequences and analyzes sequence conservation. It performs
	efficient and complete detection of various repeat types and displays
	them. FastPCR allows for sequence file batch processing, which
	is essential for automation. The FastPCR software is available for
	download at http://primerdigital.com/fastpcr.html and online version at
	http://primerdigital.com/tools/pcr.html.
Keywords	PCR primer design - Primer linguistic complexity - Sequence
(separated by "-")	assembly - Software probe design - Ligase chain reaction - DNA primers
	P

Chapter 18

2

3

4

5

26

FastPCR Software for PCR, In Silico PCR, and Oligonucleotide Assembly and Analysis

Ruslan Kalendar, David Lee, and Alan H. Schulman

Abstract

[AU1]

This chapter introduces the software FastPCR as an integrated tools environment for PCR primer and 6 probe design. It also predicts oligonucleotide properties based on experimental studies of PCR efficiency. 7 The software provides comprehensive facilities for designing primers for most PCR applications and their 8 combinations, including standard, multiplex, long-distance, inverse, real-time, group-specific, unique, and 9 overlap extension PCR for multi-fragment assembly in cloning, as well as bisulphite modification assays. It 10 includes a program to design oligonucleotide sets for long sequence assembly by the ligase chain reaction. 11 The in silico PCR primer or probe search includes comprehensive analyses of individual primers and primer 12 pairs. It calculates the melting temperature for standard and degenerate oligonucleotides including LNA 13 and other modifications, provides analyses for a set of primers with prediction of oligonucleotide proper-14 ties, dimer and G/C-quadruplex detection, and linguistic complexity, and provides a dilution and resus-15 pension calculator. The program includes various bioinformatics tools for analysis of sequences with CG or 16 AT skew, of CG content and purine-pyrimidine skew, and of linguistic sequence complexity. It also permits 17 generation of random DNA sequence and analysis of restriction enzymes of all types. It finds or creates 18 restriction enzyme recognition sites for coding sequences and supports the clustering of sequences. It 19 generates consensus sequences and analyzes sequence conservation. It performs efficient and complete 20 detection of various repeat types and displays them. FastPCR allows for sequence file batch processing, 21 which is essential for automation. The FastPCR software is available for download at http://primerdigital.com/ 22 fastpcr.html and online version at http://primerdigital.com/tools/pcr.html. 23

Key wordsPCR primer design, Primer linguistic complexity, Sequence assembly, Software probe24design, Ligase chain reaction, DNA primers25

Abbreviation

OE-PCR	Overlap extension PCR	27
PCR	Polymerase chain reaction	28
RT-PCR	Real-time PCR	29
SSR	Simple sequence repeat	30

Svein Valla and Rahmi Lale (eds.), DNA Cloning and Assembly Methods, Methods in Molecular Biology, vol. 1116, DOI 10.1007/978-1-62703-764-8_18, © Springer Science+Business Media New York 2014

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

67

68

69

70

71

72

73

Ruslan Kalendar et al.

31 **1** Introduction

The polymerase chain reaction (PCR) is fundamental to molecular biology and is the most important practical molecular technique for the DNA research laboratory. However, the utility of the method is dependent on identifying unique primer sequences and designing PCR-efficient primers. Primer design is a critical step in all types of PCR methods to ensure specific and efficient amplification of a target sequence [1-7]. Even though there are currently many online and commercial bioinformatics tools, primer design for PCR is still not as convenient and practical as it might be for routine use. The adaptation of PCR for different applications has made it necessary to develop new criteria for PCR primer and probe design to cover uses such as RT-PCR, real-time PCR, groupspecific and unique PCR, combinations of multiple primers in multiplex PCR, overlap extension PCR fort multi-fragments assembling cloning, and bisulphite modification assays. There is a need as well as for a program integrating design oligonucleotide sets for long sequence assembly by the ligase chain reaction (LCR), discovery of simple sequence repeats (SSRs) and their amplification as diagnostic markers, and for designing TaqMan, molecular beacon, and microarray oligonucleotides [6, 8, 9].

In developing FastPCR and Java web tools (Table 1), our aim was to create practical and easy-to-use software for routine manipulation and analysis of sequences for most PCR applications. The parameters adopted are based on our experimental data for efficient PCR and are translated into algorithms in order to design combinations of primer pairs for optimal amplification. This software, FastPCR, has been successfully used throughout the scientific community in a wide range of PCR and probe applications, and repeat searches and analysis. The first papers describing this software were published in 2001 and since 1999 this software has been applied in various projects, have been cited together over 500 times in scientific journals, patents, PhD theses, and over 1,000 downloads of the installation file per month. The program code is regularly updated.

66 **2 Softwar**

Software, General Information

The FastPCR software (http://primerdigital.com/fastpcr.html) is written in Microsoft Visual Studio 6.0 and compiled to an executive file that, after installation, can be used with any version of Microsoft Windows. For Linux and Mac it requires "Wine" (http://www.winehq.org/) as a compatibility layer for running Windows programs. It is a completely free alternative implementation of the Windows API also for use with native Windows DLLs.

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

t1.1 Table 1

t1.2 Summary of the FastPCR software for PCR, in silico PCR, and oligonucleotide assembly and analysis

t1.3	Features
t1.4	PCR tool provides comprehensive facilities for
t1.5 t1.6 t1.7 t1.8 t1.9	Design of primers for most PCR applications and their combinations, including standard, multiplex, long-distance, inverse, real-time, unique (specific primers for each from genetically related DNA sequences) or group-specific (universal primers for genetically related DNA sequences), linear-after-the-exponential (LATE)-PCR, bisulphite modification assays, polymerase extension PCR multi-fragment assembly cloning
t1.10 t1.11	Design of long oligonucleotides for microarray analyses and dual-labeled oligonucleotides for probes such as molecular beacons
t1.12 t1.13	Polymerase chain assembly (PCA) or oligo assembly—for automating the design of oligonucleotide sets for long sequence assembly by ligase chain reaction (LCR) and PCR
t1.14 t1.15 t1.16	In silico (virtual) PCR or multiple primer or probe searches, or in silico PCR against whole genome(s) or a list of predictions by chromosome of probable PCR products, and search for potential mismatching locations of the specified primers or probes
t1.17 t1.18	Testing of individual primers, melting temperature calculation for standard and degenerate oligonucleotides including LNA and other modifications
t1.19 t1.20	Evaluation of PCR efficiency, linguistic complexity, dimer and G/C-quadruplex detection, dilution and resuspension calculator
t1.21 t1.22	Analysis of features of multiple primers simultaneously, including T_m , CG content, linguistic complexity, dimer formation; optimal T_a
t1.23 t1.24	Identification of simple sequence repeat (SSR) loci by analyzing the low-complexity regions of input sequences
t1.25 t1.26	Restriction digest analyses for Type I, II, and III restriction enzymes and homing endonucleases, finding or creating restriction enzyme recognition sites for coding sequences
t1.27	Searches for similar sequences (or primers)
t1.28 t1.29	Translation of nucleotide (DNA/RNA) sequences to the corresponding peptide sequence in all six frames for standard and degenerate DNA and modifications (inosine, uridine)
t1.30 t1.31 t1.32	Determination of CG: $(G-C)/(G+C)$, AT: $(A-T)/(A+T)$, SW: $(S-W)/(S+W)$, MK: $(M-K)/(M+K)$, purine–pyrimidine (R–Y)/(R+Y) skews, CG% content and the melting temperature, primer quality and linguistic sequence complexity profiles

The online FastPCR (jPCR) software (http://primerdigital.com/ 74 tools/) is written in Java with NetBeans IDE (Oracle) and requires 75 the Java Runtime Environment (JRE) on a computer. It can be 76 used with any operating system (64-bit OS preferred for large 77 chromosome files). 78 this figure will be printed in b/w

Ruslan Kalendar et al.

79 **3 The Interface**

80	3.1 Inputs to	The software contains the menus, the toolbars, and the ribbon and
81	FastPCR	three text editors. The ribbon is designed to help the user quickly
82		find the commands that are needed to complete a task. Commands
83		are organized in logical groups, which are collected together under
84		tabs (Fig. 1). Each tab relates to a type of activity, such as "PCR
85		Primer Design," "in silico PCR," or "Primer Test."
86		Getting started with a basic project in FastPCR software is as
87		easy as opening a new or existing file than using copy-paste or
88		starting to type. There are three independent text editors on differ-
89		ent tabs within the interface: "General sequence(s)," "Additional
90		sequence(s) or pre-designed primers (probes) list," and "Result
91		report." The first two text editors are necessary for loading
92		sequences for analysis, the text editor "General sequence(s)" is
93		designed for working with the project sequences, and the
94		"Additional sequence(s) or pre-designed primers (probes) list"
95		text editor is used for special and additional sequences such as pre-
96		designed primers, multiple query sequences, or numbers for input.

[AU2]

						-						
FastPCR 6.4												_ 0
Ele Edit Search Con	werting PCR Da	tabase Algnment	Run Help									
	XIDIGM		A 8 8	• 신 크 🎬 🛛	६ स 🕨							
M=(A/C) R=(A/G) W=(A/T	T) S=(G/C) Y=(C/T)	K=(G/T) V=(A/G/C) H	1=(A/C/T) D=(A/G/T) B=(C/G/T) I	N=(A/G/C/T), U=T	and I	TO Group L		000 0	Testa I	01000	DNARNA
PCR Primer Design	IN SILICO PCR F	rimer Test Prime	ers List Anal	ysis Restriction	n Clustering	Searching	LTR Search	MITE Search	SSR Search	Tools	Oligos A	ssembly
Parame	eters for PCR pr	oduct analysis:	_		PCR prime	er design optio	ons:					
Synchronizing	Tm(°C) and dG(cal/mol) for prime	r pair (±):	5	I The se	condary (non-	specific) bindi	ng test	Inverted	PCR		
Limit	it for compatible of	ombination of pair	r primers:	500	✓ Linguis	tic complexity	control		Circular	DNA		
Г	Polymerase e	tension cloning (OE-PCR)		C Overlap	ping primers						
5	7 Multiplex PCR				□ C>>T	bisulphite con	version		Unique F	PCR		
Minimal differ	rence between n	ultiplex PCR prod	ucts (bp):	10					Group-s	pecific F	PCR	
Maximal difference b	between Ta of mi	ultiplex PCR produ-	cts (±°C):	5	☐ Molecu	ilar beacon de	esign					
Limit for multiplex P	CR compatible o	ombination of pair	primers:	500	☐ Microa	rray design						
Comunerat 1 : 2040	1				10000	-						
tatccacccagtgcaca tcaacttcagtagtaag tgctcatgggcctggac acgcgcgagtagttcct aaggcgccactcactta	agcgtttaggtati gctcttatcatgc ctcgctgtctaati ttaccttaaacatc acgaccctatggg	ccatcacggctaat cggttgcagtgcgc cggttcatatatat ccgtcggcctttag agatagcctctgtt	ctagtttga cttgataaa gtagagaac gcaaagcgg cttggctaa	cccaattcgacgg tcgtgcttgcaaa aggtgactcgttt gttgacacgtttt tttatgtaacag	gtaacatogtgoo acacataacogt tggggtcatgtoo ttgtaccagtagt acgctatotogot	cttcctgagtg tgagagtaaaa gagcogttaaa ttgaagtccct tagtgatgcgt	tccaccattgt ttgaaat tcaaagcacct cagcgccactc accgtaccgt	tctagctccgg gggccttacct acgtcg] aacacaaccat	aaagcccgtga taactagacaa tgagaatgggg	atogtto agoctag	cag] cg] cgaagcaa	ctggatecteg
[tatccacccagtgcaca ttcaacttcagtagtag [tgctcatgggctggac [acgogcggtagtcctta gagggccactcactta gaag[gttagccaacg [attcgagtctgtcgg ttatggaaaaatgaac	agcgtttaggtatt getettateatgeg etogetgetetaatt ttacettaaacate acgaecetatggg ggggtetgaetea gaecaagaacaaece ettagaegtegagg	tcatcacggctaat icgatagcagtgcgc cogttcatatatat coctcggcctttag iaccgaaactattc gtcaaggcatcgct cgggttgtgtctcg	cctagtttga cctgataaa gtagagaac gcaaagcgg cctggctaa cctgggtttg cact] ftggcagttc	cccaattcgacgg tcgtggttgcaa aggtgactcgttt gttgacacgtttt tttatgtaacaga agtgacttttggt ccgcatctcaaat	gtaacatcgtgco acacataaccgtf tggggtcatgtco ttgtaccagtagt acgctattcgct ttccgagtggtco tacttggcagtgt	cttcctgagtg tgagagtaaaa gagccgttaaa ttgaagtcctd tagtgatgcgt cttgccctgac tgtgaatttgt	tccaccattgt ttgaaat tcaaagcacct cagcgccactc accgt&ccgtt ggatttaagct gctgacggcco	tctagctccgg gggcettacet acgtcg] aacacaaccat aactg] ggggggggcaac	aaagcccgtga taactagacaa tgagaatgggg gttgatatgct	ategtte ageetag geettta	cag] cg] cgaagcaa atccggcg	ctggatcctcg cgtgggcacct
[Latocacceagtgcaca ttoacttegtagtagt [tgctcatgggcctggac iaaggcgcagtagttoot jaag[gttaagcocaagg [atteggcctcacta iaatotgggtctgtggg [ta [acaggtcagtggag]	agogttaggtati getettateatge tegetgetaata aggaecetatggg gggetetgaetea gaeaagaacaae ettagaegtegag getgaaatatee	tcatcacggctaat tcgatagcagtgcgc cggttcatatat agatagcctttag gtcaaggcatcgct cgggttgtttcg actgggaagccg	cctagtttga cctgataaa gtagagaac ggcaaagcgg ccttggctaa cctgogttg cact] ttggcagttc agtatacta	cocaattogaogg togtgottgoaaa aggtgactogttt gttgacaogtttt tttatgtaacagg agtgacttttggt cogcatctcaaat toatttgccacto	gtaacatogtgcc acacataacogt tggggtcatgtcg ttgtaccagtagt acgctattcgg ttccgagtggtcc tacttggcagtg cggttacatgcg	cttectgagtg tggggtaaaa ggeegttaaa ttggagteett tagtgatgegt cttgeectgae tgtgaatttgt cacatetggtt	tccaccattgt ttgaaat tcaaagcacct accgtaccgt	tctagctccgg gggccttacct acgtcg] accacaaccat aactg] gggggggcaac	aaagcoogtga taactagacaa tgagaatgggg gttgatatgot	acctag gcctta gccttta	cag] cg] cgaagcaa atccggcg	ctggatcctcg cgtgggcacct
[tatccacccagtgcaca ttoacttcagtagcag [tactccaggcctggac [actocggcagtagttcct iaggcgcactcacts jaag [gtaagcccagt [attctggagtcggcg [attctggagtggag [atcactsaggtaggt [accactsaggtaggt [gg [ttacactaagg	aggtttaggtat gotttatatogg ttogtgtttaat ttacttaacat gacacattagg gggtttgactaa gtgacatggag gggtttactaa ttagacgtogg gggactggaattacc coggetcat	tcatcacggctaat cq2tagcagtgcg cggttcatatat ccgtcggctttag lagatagcctcggt cggttgtgtctcg gccaggcagcga actgggtggagcga cacactcggatag gctatggccagg	cctagtttga ccttgataaa gtagagaac gcaaagogg ccttggctaa cctgogtttg tggcagttc aggtatacta cccacc] taggaccaac actactcac	cccaattogaogg togtgottgoaaa aggtgaotogttt tttatgtacagtttt tttatgtacag agtgaottttggt cogcatotcaaat tcatttgocacto acttotgoagate tgogaottotcat	gtaacatogtgco acacataaccgt tggggtcatgtcy tggggtcatgtcy ttccgggtggtggtgg ttccgggtggtgg cggttacatgcg agaaatgacgca tttccgggaggt	Cttootgagtgi tgagagtaaaa gagcogttaaai ttgaagtootk tagtgatgogt cttgocotgack tgggaatttgtk cacacotggtk cagggcaccaai	tccaccattgt ttgaaat tccaaggcccttc accgtaccgt	totagotogg gggoottacot acgtog] aacacaaccat aactg] gggggggggaaac ccccgacgagg tgagcatc]	aaagcoogtga taactagacaa tgagaatgggg gttgatatgot atcoggacgtt	atcgttc agcctag gccttta gcggggg ccctacg	cag] cg] cgaagcaa atccggcg agaacaat	ctggatcctcg cgtgggcacct aactgaggggt
tatocaccoagtgcaca ttoacttcagtgcaca ttgetatgggctggac aaggggcactagtect taggggcactactta aaggggcactaatgac ttatgaaaaatgaa ttatgaaaaatgaa gagttactataggcog gggttactatagcog ggggtgactaatg ggggaggagtgattaaca tagggggggggg	agotttaggtatt gtotttateatgo ttogtgtotatt gtottateatgo ggogottgattagg ggogotgaagacaa ttagacgtoga gottagattatoo toggttaastatoo toggttactoge ggottggctaga ttattotaggat ttototaggctaga	tcatcacgctaat cogatagcagtgcoc cogstcatatat cogatagcagtcac accogastatto coggstcgtatca accoggatgtatca accoggatgtatca accoggatgtagtag ggtcatctaggst ggtcatctggatg agtaggacgtagcaa ggtgcatctaggat	<pre>cctagtttga cctggtagagaac ggagagaac ggagagagac ggcagggtaggtag actiggctagt actiggcagtts actactac cactactac ccgaaaget ccggtacgtg ccggtacgtt</pre>	cocattogacg togtgottgoaa aggigactogtt ttatgoacag ttatgoacagtt ttatgoacag agtgactttago acttotgoagata tgoactotoaa gattactagatg cogattactog cogttactagatg cogttactagatg cogtactagatg	gtaacatogtgoo acacataacogti tggggtcatgtoo ttggggtcatgtoo ttccgggtggtggtoo tacttggcagtggtoo agaaatgacgcao tttccggogatggt ttactgttgco] gtagcaatgatgg gtagcaatgatgg gtagcaatgatgg	tticotgagtg tgagagtaaaa paccettaaa ttgagagtaaa ttgagtcoct ttgagtcoctgacy ttgggaattggt cacactotggty cagggaccaaa tttgctacagti aaaattaactti cacgctagacg	tccaccattgt ttgaaat tccaagcgccctt ccagggccctt accgtacgtt ggstttaagct ggsgst] tgctttccccg tccagtggttt tctacctaggc c	totagotocgg gggoottacot acgtog] aactog] ggggggggacaac coccegacgago tgagcato] agtoacagttt	aaagooogtga taactagaatgggg gttgatatgot atooggaogtt tttattttooo	atcgtto gcctag gcctta gcggggg ccctacg	cag] cg] cgaagcaa atccggcg agaacaat agcttcgt	ctggatcctcg cgtgggcacct aactgaggggt ccttcatcggt
[Latocaccagtgcaca treacttcagtgagaca treacttcagtgagac acgogcagtagttcot aaggogcastaattcot aaggogcastaattcot aaggogcastaattgagg [stastcogstogtcgagg treatgsaaaastgaac treatgsagggagggaggag accactaaggtgaggg accactaaggtgaggg [stasastaaggoggacaggca gag [stasastaatacagg gaggggaggagattgastcogg cogstggaggagattgastc	agotttaggtatt getettatetgo ttogttgattast ttacttaasatt gggettgattast gggettgattas ttagagtetgag gettgastatoo ttagagtetgag gettggettactgo ggettggettas ttettetggetgas tattetetagea gettogetgat ggettgattas tettetetggetgas atattetetagea gettggetgas ggettggetgas ggettggetgas ggettggetgas ggettggetgas ggettggetgas ggettggetgas ggettggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggetggetgas ggas ggetgas ggas ggas ggetgas ggas ggas ggas ggas ggas ggas ggas	tratcacggetaat gegatageagtgege eggttestatata icogteggettage gatagetecteggt iatcogaasetatte ggtagatgeteteg gattaggetagete gattaggetagete gatestggecagt ggtestattggatgega tggaatgaggetag cttaeggegtag agegggagaeag	cctagtttga cctagtttga gcagagac gcagagac gcagagac gcagaggg tctggcttg tcggcagttcg tcgcagttcg tcgcacgt ccgaaagc tcgatacgg gttggccag ggtggccag ggtggccag ggtggccag	cocattogacg togtgottgota aggigatogtt tttatgcacag aggigattit ttatgcacag aggigattit cogtattocaa toattogcag togtttatoc cgatttatoc cgatttatoc cgatttatoc cgatttatoc cgacacg cgacocc tgocgacg cgacocc gacocc tgocgaccocc gacgag cocc	gtaacatogtgoo acacataacogti tggggtcatgtoo ttggggtcatgtoo ttccgggtggtggtoo tacttggcagtgg cggttacatgoo gtaacatgacgaa tttccgggcacaco j cgcccgaagtgg cggttacacgaa gggggtcottaa	ttictigagtg tgagagtaaa ggacgttaaat tgaagtcoct atgaagtcoct ttigectigac ttigectigac ttigectigac ttigectigac ttigectagt aaattaactt cagggcaccaa ttigetaagt aaattaactt aaattaactt aaacaagcagg gtgettggcta	tcoacattgt ttgaat tcaagcactt cagogcactc agogcactc gggst1 ggttacggcc ggggt1 tggtttcacgg tcaagtggtt1 tcattcoccg c c l attgttgcca agggaatca	totagotocgg gggoottacot acgtog] aacacaaccat actg] ggggggggcaac coccgacgago tgagoatc] agtoacagttt cgotacacgta tatcattgta	aaagooogtga taactagacaa tgagaatgggg gttgatatgot atcoggacgtt tttattttoco ttogacgaago goggggttog	acctag gcctta gccttta gcggggg ccctacg ccctacg ctaatat	cag] cgaagcaa atccggcg agaacaat agcttcgt tgcgcaga ccattc]	ctggatcctcg cgtgggcacct aactgaggggt ccttcatcggt gcgtgctcggc
Latocaccagtgcaca tcaacttcagtagtagt tagtcatggcgctggac aagggcgcattagttcot aaggggcattagttcot aaggggcattagtcata (actocgatcgtcgtggg (actocgataggtgggg (accactaaggtagggt caggggggggggtagggt (acagggggggggggggggg (acagggggggggggggggg (acaggggggggggggggggggggggggggggggggggg	agootttaggtatt gotttaggtatt gotttatoatgo ttogottgattaat gggoottgattaat gggoottgattaa gggoottgatta ttagacgooggoog ttagacgooggoog totootggotag attoctoggooggoog gottogocagoot gottogocagoot gottoacaggtog ggttcacogaat gttcacaagtca goga (gottoct)	tratcacggettat cogtagecattatat cogtaggettatatat cogtaggettatatat cogtaggettget cogtaggettgtetagg cogtaggetgtgtetagg cacactcoggatac actaggetgtetagget ggtcastettgttagget ggtcastettgttagget sagagggggatacag T=519.0 G=501.0	cctagtttga cctagtttga gtagaagac ggaagagac ggaagagac cctgggtttg tggcagttc actactag ccccacc] sggaccac ccgaaagc ttggccagtc ccgaaagc ttggtccag ggttgccag ggttgccag cagtaccac	cccattcgacg tcgtgcttgcaa agggactcgtt ttatgcacagttt ttatgcacag agtgactttggt ccgcattccaa acttctgcagate cgattttactcg gatctactacat cgattttactg cgattttactg cgacgacg cgacgacg ccgacgacg ccgacgacgag ccgacgag c	gtaacatogtgoo acacataacogt togggtaataacogt togtacataacogt tuccesgtagt acgotatotogot tuccesgtggt oggttacatgoog agaaatgacgacatg ttaccatgatgg gcoggogacacoco ogcocogaagtgg cogttoatacaga gggggtcottag 0 RY=0.48 CG=	cttoctgagtg ggaggtaaaa ggoogtaaa ggoogtaaa ggoogtaa ttgagtoogt cttgootgao cttgootgao cagggaattagt cagggaattagt cagggaccaa ttggtaattagt aaaattaacti caggcaccaa ggotggootat aaaataaccagg ggotggootat aaaataaccagg	tocaccattgt trgaaat trgaagcactc acogtacogta ggatttasgot gggtttasgot tgotttoccog toagtggtt totacctaggo c atgttatgcca aggaaatcgc	totagotocgg gggoottacot acgtog] ascaceaccat anctg] ggggggggeacc tgagoatc] agtoacagttt cgotacaogta tatoattigto	asageccytys taactagacaa tyagaatgygg gttgatatget atceggaegt tttatttteee tregaegaage geggggteteg with a '>' symb	stogtto sgoctag sgogggg sgogggg scoctacg staatat stogettt satgcta	cag] cgaagcaa atccggcg agaacaat agcttcgt tgcgcaga ccattc] F Read	ctggatcctcg cgtgggcacct aactgaggggt ccttcatcggt gcgtgctcggc

Fig. 1 The FastPCR sequence editor and user interface

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

Program Output 3.2 FastPCR automatically generates results at a third text editor, 97 "Result report," in tabulated format for transferring to spreadsheet 98 software from the clipboard using copy-paste. Output results are 99 easy to save as Excel worksheet (.xls) or as Rich Text Format (.rtf) 100 text files, compatible with MS Excel or Open Office. The separated 101 output of the primer design is a list of primers, a set of primer pair 102 sequences with their theoretical PCR products, and, for multiplex 103 PCR, the result of the calculation of multiple-PCR primers for 104 given target sequences. In addition, the output shows optimal 105 annealing temperature for each primer pair, the size of PCR 106 product, and complete information for each designed primer and 107 for each multiplex PCR product set. 108

3.3 Sequence Entry Sequence data files are prepared using a text editor (Notepad, 109 WordPad, MS Word), and saved in ASCII as text/plain format (.txt) 110 or in .rtf. The program either takes a single sequence or accepts 111 multiple separate DNA sequences in FASTA, tabulated format (two 112 columns from MS Excel sheet or MS Word table), EMBL, MEGA, 113 GenBank, MSF, DIALIGN, simple alignment, or BLAST Queue 114 web alignment formats. The template length is not limited. The 115 FastPCR clipboard allows the user to copy and paste text or tables 116 from MS Word documents or MS Excel worksheets or other programs 117 and to paste them into another Office document. It is important that 118 all target sequences are prepared in the same format. Users can type 119 or import from file(s) into "General Sequence(s)" or "Additional 120 sequence(s) or pre-designed primers (probes) list" editors. 121

FastPCR allows files to be opened in several ways: the original 122 file can be opened as read-only for editing with text editors; files 123 can be opened to memory without using text editors, which allows 124 larger file(s), up to 200 Mb, to be analyzed; files within a folder can 125 be selected and the files opened during task execution without the 126 use of text editor program. Additionally, the program can open 127 files within a selected folder in order to join all these files in a text 128 editor. For example, this feature can be applied to convert all files 129 from a selected folder into a single file of FASTA sequences. 130 Alternatively this feature allows splitting FASTA sequences to indi-131 vidual files in a particular selected folder. 132

When a sequence file is opened, FastPCR displays the information about the opened sequence and its format. The information status bar shows the number of sequences, the total sequence length (in nucleotides), the nucleotide composition, and the purine, pyrimidine, CG percentage, and the melting temperature. Files can be saved in various formats including .rtf, .xls, or txt from the text editor in use.

3.4 FASTA Format Description FastPCR normally expects to read sequence files in FASTA format 140 [10]. FASTA format has the highest priority and is simple, because 141 the raw sequence is merely preceded by a definition line. The 142 definition line begins with a ">" sign and optionally followed 143

151

152

153

154

155

156

157

158

159

Ruslan Kalendar et al.

immediately by a name for the sequence with no length restriction.
Many sequences can be listed in the file, the format indicating a
new sequence at each ">" symbol found. It is important to press
Enter at the end of each line after ">" to help FastPCR recognize
the end and beginning of the sequence and the sequence name. It
is important that the first line of each sequence starts with ">."
Degenerate DNA sequences are accepted as IUPAC code [11],
an extended vocabulary of 11 letters, which allows the description

an extended vocabulary of 11 letters, which allows the description of ambiguous DNA code. Each letter represents a combination of one or several nucleotides: M = (A/C), R = (A/G), W = (A/T), S = (G/C), Y = (C/T), K = (G/T), V = (A/G/C), H = (A/C/T), D = (A/G/T), B = (C/G/T), N = (A/G/C/T), U = T, and I (Inosine). The program accepts amino acid codes: A (Ala), C (Cys), D (Asp), E (Glu), F (Phe), G (Cly), H (His), I (Ile), K (Lys), L (Leu), M (Met), N (Asn), P (Pro), Q (Gln), R (Arg), S (Ser), T (Thr), U (Sec), V (Val), W (Trp), and Y (Tyr).

3.5 Alignment There are many different programs that perform different types of 160 Format Description alignment formats. Standardizing on a set of formats enables 161 programs to be written that can read results from many different 162 programs. In all alignment formats, gaps that have been introduced 163 into the sequences to make them align are indicated by the "-" 164 character. The exception to this rule is GCG/MSF format which 165 uses "." as the gap character inside the sequences. The file may 166 begin with as many lines of comment or description as required. 167 The first mandatory line must contain the text "MSF," "Alignment 168 as simple alignment format," "DIALIGN," or "MEGA" to be 169 recognized as alignments from these programs. Following the first 170 line are lines that start with the sequence name, which is separated 171 from the aligned sequence residues by white space. 172

173 4 The PCR Primers or Probe Design Analysis Options

Primer design is one of the key steps for successful PCR. For PCR 4.1 PCR Primer 174 applications, primers are usually 18-35 bases in length and should **Design Generalities** 175 be designed such that they have complete sequence identity to the 176 desired target fragment to be amplified. The parameters, either 177 controllable by the user or selected automatically, are primer length 178 (12-500 nt), melting temperature for short primers calculated by 179 nearest neighbor thermodynamic parameters, theoretical primer 180 PCR efficiency (quality at %) value, primer CG content, 3' end 181 terminal enforcement, preferable 3' terminal nucleotide sequence 182 composition in degenerated formulae, and added sequence tags at 183 5' termini. The other main parameters used for primer selection 184 are the general nucleotide structure of the primer such as linguistic 185 complexity (nucleotide arrangement and composition), specificity, 186

t2.1	Tab	le 2

t2.2 Default primer design selection criteria

t2.3	Criteria	Default	Ideal
t2.4	Length (nt)	20–24	>21
t2.5	$T_{\rm m}$ range (°C) ^a	52–68	60–68
t2.6	T_m^{a} 12 bases at 3' end	30–50	41-47
t2.7	CG (%)	45-65	50
t2.8	3^\prime end composition $(5^\prime\text{-}NNN\text{-}3^\prime)^b$	SWW, SWS, SSW, WSS	SSA, SWS, WSS
t2.9	Sequence linguistic complexity (LC, %) ^c	>75	>90
t2.10	Sequence quality (PQ, %)	>70	>90

t2.11 ^aNearest neighbor thermodynamic parameters according to SantaLucia [13]

t2.12 ^bAmbiguity codes

t2.13 ^cSequence linguistic complexity measurement calculated using the alphabet-capacity *L*-gram method

the melting temperature of the whole primer, the melting 187 temperature at the 3' and 5' termini, self-complementarity, and 188 secondary (nonspecific) binding. 189

The software can dynamically optimize the best primer length 190 for the entered parameters. All PCR primer (probe) design param-191 eters are flexible and changeable according to the specifics of the 192 analyzed SEQUENCES and tasks. Primer pairs are analyzed for 193 cross-hybridization and specificity of both primers and, optionally, 194 selected with similar melting temperatures. Primers with balanced 195 melting temperatures (within 1-6 °C of each other) are desirable 196 but not mandatory. The default primer design selection criteria are 197 shown in Table 2. It is possible to use predesigned primers or 198 probes or, alternatively, predesigned primers can act as references 199 for the design of new primers. The program accepts a list of prede-200 signed oligonucleotide sequences and checks the compatibility of 201 each primer with a newly designed primer or probe. 202

The melting temperature (T_m) is defined as the temperature at 203 which half the DNA strands are in the double-helical state and half 204 are in the "random-coil" state. The $T_{\rm m}$ for short oligonucleotides 205 with normal or degenerate (mixed) nucleotide combinations is 206 calculated in the default setting using nearest neighbor 207 thermodynamic parameters [12, 13]. The CG content of an 208 oligonucleotide is the most important factor that influences the $T_{\rm m}$ 209 value. The melting temperature for mixed bases is calculated by 210 averaging nearest neighbor thermodynamic parameters-enthalpy 211 and entropy values-at each mixed site; extinction coefficient is 212 similarly predicted by averaging nearest neighbor values at mixed 213 sites [2, 3]. Mismatched pairs can be taken into account since the 214 parameters provide for DNA/DNA duplexes and dangling ends, 215

[AU3]

4.2 Melting Temperature Calculation

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

Ruslan Kalendar et al.

which are unmatched terminal nucleotides [14–16]. The melting temperature for primer (probe) self- or cross-dimers and for in silico PCR experiments with oligonucleotides having mismatches to the target is calculated using values for the thermodynamic parameters for a nucleic acid duplex.

The FastPCR allows the choice of other nearest neighbor thermodynamic parameters or simple non-thermodynamic T_m calculation formulae. For non-thermodynamic T_m calculation, we suggest using simple formulae; the Wallace–Ikatura rule [17] is often used as a rule of thumb when primer T_m is to be estimated at the bench. However, the formula was originally applied to the hybridization of probes in 1 M NaCl and is an estimate of the melting temperature for oligonucleotides shorter than 10 bases:

$$T_{\rm m}(^{\circ}\mathrm{C}) = 2(\mathrm{L} + \mathrm{G} + \mathrm{C}),$$

for oligonucleotides longer than ten bases:

$$T_{\rm m}(^{\circ}{\rm C}) = 64.9 + \frac{41([G+{\rm C}]-16.4)}{L}$$

or the formula [18]:

$$T_{\rm m}(^{\circ}{\rm C}) = 77.1 + 11.7 \log_{10} \left[{\rm K}^{+}\right] + \frac{41[{\rm G}+{\rm C}] - 528}{L}$$

where *L* is length of primer, [G+C] is the number of Gs and Cs, and $[K^+]$ is salt molar concentration (default value is 50 mM). The two equations above assume that the stabilizing effects of cations are the same on all base pairs. The melting temperature of the PCR product is calculated using the formula [15]:

$$T_{\rm m}(^{\circ}{\rm C}) = 81.5 + 16.6 \log_{10} \left[{\rm K}^{+} \right] + \frac{41 \left[{\rm G} + {\rm C} \right] - 675}{L}$$

4.3 Linguistic The sequence complexity calculation method can be used to search 240 for conserved regions between the compared sequences in order to Complexity of 241 detect low-complexity regions including SSRs, imperfect direct or Sequence and 242 inverted repeats, polypurine and polypyrimidine triple-stranded Nucleotide-Skew 243 DNA structures, and four-stranded structures (such as G/C-244 Analysis quadruplexes) [19]. Linguistic complexity measurements are 245 performed using the alphabet-capacity L-gram method [20, 21] 246 along the whole sequence length and calculated as the sum of the 247 observed range (x_i) , from 1- to L-size words in the sequence, 248 divided by the sum of the expected (E) value for this sequence 249 length. Linguistic complexity (LC) values for sequence length (s)250 are converted to percentages, in which 100 % means maximal 251 "vocabulary richness" of a sequence: 252

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

$$LC(\%) = \frac{100 \times \sum_{i=1}^{L} x_i}{E},$$
253

254

$$E = \sum_{i=1}^{L} \left\{ \frac{s - i + 1, s < 4^{i} - 1 + i}{4^{i}, s \ge 4^{i} - 1 + i}, \\ L = \left[\log_{4} \left(\frac{s}{3} \right) + 1 \right].$$
255
256

For example, the sequence 5'-ACACACACACACACACAC, 257 16 nt (L=3), contains two nucleotides (A, C), but expected E=4 258 variants; two variants of dinucleotides (AC, CA), but expected 259 E=(16-1) variants; two variants of trinucleotides (ACA, CAC), 260 and expected E=(16-2) variants. The complexity value is 261 LC=100(2+2+2)/(4+16-1+16-2)=18.2%. 262

The LC tries to describe the "uniqueness" ("vocabulary richness") of a sequence and the likelihood of PCR success of each primer; this value varies from 100 for the best to 5 (e.g., poly(N)) 265 for the worst primers. LC values of 80 and higher allow for the rapid choice of the best primer or probe sequences. 267

Our experimental data showed that the primer nucleotide 268 composition and melting temperature of the 12 bases at the 3' end 269 of the primers are important factors for PCR efficiency. The melting 270 temperature of the 12 bases at the 3' terminus is calculated 271 preferably by nearest neighbor thermodynamic parameters [13]. 272 The composition of the sequence at the 3' terminus is important; 273 primers with two terminal C/G bases are recommended for 274 increased PCR efficiency [22]. Nucleotide residues C and G form 275 a strong pairing structure in the duplex DNA strands. Stability at 276 the 3' end in primer template complexes will improve the 277 polymerization efficiency. 278

We specify an abstract parameter called primer quality (PQ) 279 that can help to estimate the efficiency of primers for PCR. PQ is 280 calculated by the consecutive summation of the points according 281 to the parameters of total sequence and purine-pyrimidine 282 sequence complexity and of the melting temperatures of the whole 283 primer and of the terminal 12 bases at both the 3' and 5' ends. 284 Self-complementarity, which gives rise to possible primer-dimer 285 and hairpin structures, reduces the final value. The PQ tries to 286 describe the likelihood of PCR success of each primer; this value 287 varies from 100 for the best to 0 for the worst primers. To meet 288 multiplexing demands, it is possible in the program to select the 289 best primer with an optimal temperature range, allowing the design 290 of qualified primers or probes for any target sequence with any CG 291

4.4 Primer Quality (Virtual PCR Efficiency) Determination where

Ruslan Kalendar et al.

292and repeat content. PQ values of 80 and higher allow for the rapid293choice of the best PCR primer pair combinations. No adverse294effects, due to the modification of the reaction buffer, chosen ther-295mostable polymerases, or variations in annealing temperature, have296been observed on the reproducibility of PCR amplification using297primers with high PQ.

4.5 Hairpin (Loop) and Dimer Formation Primer-dimers involving one or two sequences may occur in a PCR. The FastPCR tool eliminates intra- and inter-oligonucleotide reactions before generating a primer list and primer pair candidates. It is very important for PCR efficiency that the production of stable and inhibitory dimers is prevented, especially by avoiding complementarity in the 3' ends of primers from whence the polymerase will extend. Stable primer-dimer formation is very effective at inhibiting PCR because the dimers formed are amplified efficiently and compete with the intended target.

Primer-dimer prediction is based on analysis of non-gapped local alignments and the stability of both the 3' end and the central regions of the primers (Fig. 2). Primers will be rejected when they have the potential to form stable dimers, depending on the nucleotide composition and with at least five bases at the 3' end or seven bases in the central region. Tools to calculate T_m for primer-dimers with mismatches for pure, mixed, or modified (inosine, uridine, or locked nucleic acid (LNA)) bases, using averaged nearest neighbor thermodynamic parameters, are provided for DNA/DNA duplexes [12–14, 23, 24].

In addition to Watson–Crick base-pairing, there is a variety of other hydrogen bonding configurations possible [19, 25–27], including G/C-quadruplexes and wobble base pairs, which the FastPCR software detects. The program provides for the detection of alternative hydrogen bonding during primer-dimer and in silico PCR primer binding site detection. The mismatch stability is examined in order of decreasing stability: $G-C>A-T>G\cdot G>G\cdot T \ge G\cdot A>T\cdot T \ge A \cdot A > T \cdot C \ge A \cdot C \ge C \cdot C$. Guanine is the most universal base, because it forms the strongest base pair and the strongest mismatches. However, cytosine is the most discriminating base, because it forms the strongest pair and the three weakest mismatches [23, 28]. Therefore, the software also looks for stable guanine mismatches: $G \cdot G, G \cdot T$ and $G \cdot A$.

G-rich (and C-rich) nucleic acid sequences can fold into fourstranded DNA structures that contain stacks of G-quartets [19]. These quadruplexes can be formed by the intermolecular association of two or four DNA molecules, dimerization of sequences that contain two G bases, or by the intermolecular folding of a single strand containing four blocks of guanines. These are easy to eliminate from primer design because of their low linguistic complexity; LC = 32 % for (TTAGGG)₄. The software predicts the presence of putative G- and C-quadruplexes in primer sequences.

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

PCR Primer Design in silico PCR Primer Test	Primers List Analysis	Restriction	Clustering	Searching	LTR Search	MITE Search	SSR Search	Tools	Oligos Assembly
Type of primers combination									
All against all primers	C Individual primer a	nalysis		Show prime	rs Tm for all N	IN thermodynai	nic parameters	s and for	mulas
C Forward against reverse (F/R)	C All pairs against all	l pairs							
C All primers against selected:									
						1			
Sequences: 22:515 Additional sequence(s) or p	pre-designed primers (p	probes) list	Primers anal	lysis result	Primers dimer	s analysis			
7286 5'-catagcatggataataaacgattatc									
<-ctattagcaaataataggtacgatac-5									
5-catagcatggataataaacgattatc-> Tm=19.4°C									
**** Test for Primer:									
2.2449999911999919999419999100									
Potential G/C-quadruplex:									
<-cctggggtaaggggtggggtggggaa-5 5-aagggggttggggtgggaatggggtcc->									
**** Test for Primer: (vb)n 5'-vbvbvbvbvbvbvbvbvbvbvbvbvbvb									
<-bvbvbvbvbvbvbvbvbvbv-5									
5-vbvbvbvbvbvbvbvbvbvbvbvbvb-> Tm=82.2°C									
<-bvbvbvbvbvbvbvbvbvbvbv-5 									
<-bvbvbvbvbvbvbvbvbvbvbv-5 									
Tm=78.1*C									

Fig. 2 One of the FastPCR duplex formation results

Intermolecular G-quadruplex-forming sequences are detected 339 according to the formula $\dots G_{m1}X_nG_{m2}\dots$, where *m* is the number 340 of G residues in each G-tract $(m_1, m_2 \ge 3)$; the gap X_n $(n \le 2^*$ mini-341 mal $(m_1:m_2)$ can be any combination of residues, including G 342 [16]. The gap sequences (X_n) may have varying lengths, and a rela-343 tively stable quadruplex structure may still be formed with a loop 344 more than seven bases long, but in general, increasing the length 345 of the gap leads to a decrease in structure stability. It is also possible 346 for one of the gaps to be zero length when there are long poly-G 347 tracts of >6 bases. 348

4.6 Calculation of Optimal Annealing Temperature The optimal annealing temperature (T_a) is the temperature, 349 generally stated as a range, where efficiency of PCR amplification is 350 maximal but nonspecific products minimal. The most important 351 values for estimating the T_a are the primer quality, the T_m of the 352 primers, and the length of PCR fragment. Primers with high T_m 353 (>60 °C) can be used in PCRs with a wide T_a range compared to 354

4.7 The Secondary

Nonspecific Binding

Test; Alternative

360 361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

Ruslan Kalendar et al.

primers with low $T_{\rm m}$ (<50 °C). The optimal annealing temperature
for PCR is calculated directly as the value for the primer with the
lowest $T_{\rm m}$ ($T_{\rm m}^{\rm min}$). However, PCR can work in temperatures up to
10° higher than the $T_{\rm m}$ of the primer so as to favor primer target
duplex formation:

 $T_{\rm a}(^{\circ}C) = T_{\rm m}^{\rm min} + \ln L$, where L is the length of the PCR fragment.

The specificity of the oligonucleotides is one of the most important factors for good PCR; optimal primers should hybridize only to the target sequence, particularly when complex genomic DNA is used as the template. Amplification problems can arise due to Amplification Products primers annealing to repetitious sequences (retrotransposons, DNA transposons, or tandem repeats). Alternative product amplification can also occur when primers are complementary to inverted repeats and produce multiple bands. This is unlikely when primers have been designed using specific DNA sequences (unique PCR). However, the generation of inverted repeat sequences is exploited in two common generic DNA fingerprinting methods, RAPD and AP-PCR [29, 30]. Because only one primer is used in these PCRs, the ends of the products must be reverse complements and thus can form stem-loops.

> The techniques of inter-repeats amplified polymorphism: interretrotransposon amplified polymorphism (IRAP), retrotransposonmicrosatellite amplified polymorphism (REMAP), inter-MITE amplification [31, 32], and Alu-repeat polymorphism [33, 34] have exploited highly abundant, dispersed repeats as markers. However, primers complementary to repetitious DNA may produce many nonspecific bands in single-primer amplification and compromise the performance of unique PCRs. A homology search with the primer as the query sequence, for example, using BLASTn against all sequences in GenBank or EMBL-Bank, will determine whether the primer is likely to interact with dispersed repeats. Alternatively, one can create a small, local, specialized library of repeat sequences based on those in Repbase [35] or TREP [36] and use this for searches.

> The mismatches at the 3' end of the primers affect target amplification much more than mismatches at the 5' end. A twobase mismatch at the 3' end of the primer prevents amplification. A single-base mismatch at the 3' end as well as several mismatches at the 5' end of the primer permits amplification, albeit with reduced efficiency. However, the presence of multiple primer binding sites does not necessarily lead to alternative amplification products because, for successful amplification, the priming sites for both primers must be both located close to each other, in correct orientation, and sufficiently match the primer sequences.

> By default, FastPCR performs a test for nonspecific binding by repeats masking and low-complexity regions detection and masking for each given sequence.

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

Additionally the software allows this test to be performed 403 against a reference sequence or sequences (e.g., a BAC or YAC) or 404 one's own database. Primers that bind to more than one location 405 on given sequences will be rejected. Even though the test for non-406 specific primer binding is performed as a default for all primers, the 407 user may cancel the operation. Identification of secondary binding 408 sites including mismatched hybridization is normally performed by 409 considering the similarity of the primer to targets along the entire 410 primer sequence. An implicit assumption is that stable hybridiza-411 tion of a primer with the template is a prerequisite for priming by 412 DNA polymerase. FastPCR pays particular attention to the 3' por-413 tion of the primers and calculates the similarity of 3' end to the 414 target (the length is chosen by user) to determine the stability of 415 any potential interactions. 416

The secondary nonspecific primer binding test is based on 417 repeat masking using a quick local alignment screen (which allows 418 one mismatch within a hash index of 12-mers) between the reference and input sequences. 420

Once the input files are selected or sequences copied and pasted to the **General Sequence(s)** text editor, the FastPCR provides various execution features. Figure 3 provides an example for primer design from the user's perspective. 425

5.1 Execution of the Selected Task
Selected Task
The user selects the ribbon having the task needed. The program 426 will only perform the selected task. The name of the selected 427 executive task is shown on the status bar by "Press F5." The task is 428 executed by using key F5 or by clicking the arrow on the toolbar 429 using the mouse. Once the executive task is completed, the result 430 is shown in the Result report text editor (e.g., see Fig. 3).

The "PCR Primer Design" Tab contains various execution 5.2 PCR Primer 432 options for commonly selected types of PCR and for the most **Design Options** 433 important PCR parameters (Fig. 1). The option panel of "PCR 434 Primers or Probe Design Options" is shown in Fig. 4. Once the 435 user selects any attribute, the option attribute value field shows 436 the default attributes value, which can then be modified. "PCR 437 Primers or Probe Design Options" affects all sequences. PCR 438 primer design options can be customized for each sequence using 439 special commands at the header of the sequence (http:// 440 primerdigital.com/soft/pcr_help.html). Typically, it is not 441 necessary to use these commands to manage typical PCR primer 442 design and these are applied to advanced tasks. Default global 443 parameters for primer design will be assigned by typing the help 444 command "/?" in text editor: 445

421

Author's Proof

Ruslan Kalendar et al.

Sequences: 3 : 8	344 Additional sequence	(s) or pre-desig	ned prime	ers (probe	s) list P	CR prim	ers desig	gn result					
PrimerID	Sequence (5'-3')	Length (bp)	Tm (°C) dG(kca	al/mol)	Tm_3'e	end(°C)	CG(%)	Linguis	stic_Co	mplexit	:y(≹)	Primer_Quality(%)
example1						-				-			_
1F1_1_18-39	tcgtattcaggcgtacctto	tg 22	56.5	-27.3	41.2	50.0	88	88					
1F2_1_42-62	gcctcgggctgcgttactto	g 21	64.0	-30.0	40.3	66.7	85	85					
1F3_1_53-73	cgttacttcgttgcggatag	ig 21	56.6	-26.3	41.4	52.4	87	87					
1F4_1_65-85	gcggataggactatctcggt	g 21	57.3	-26.3	37.2	57.1	85	83					
1F5_1_78-97	tctcggtgttttctgactg	20 55.4	-25.0	36.3	50.0	76	73						
1F6_1_99-118	tggcttccgcgagtcattg	20 61.1	-27.9	43.4	60.0	92	92						
1R1_1_348-368	atgtgggtttgccttacaag	c 21	56.7	-26.1	40.2	47.6	92	90					
1R2_1_264-286	tgttggttacttctcggtaa	ggt 23	55.3	-27.4	38.8	43.5	80	80					
1R3_1_161-180	atgatgccctcatgtccgta	20 55.5	-25.1	37.7	50.0	82	82						
1R4_1_136-157	tgttttatcgacaccttcgt	cc 22	55.2	-26.6	42.8	45.5	85	80					
Forward Prime	rID Sequence (5'-3	') Tm (*	C) Prime	r Qualit	.v(\$)	Revers	e Prime	TD	Semien	ce (51-3		Tm (*C)	Primer Quality(%)
PCR Fr	agment Size (bp)	Topt (*C)	o, rranc	-×	-1(+)	I.C. FERS			bedaen	0010 0	· /	***(0)	
1F1 1 18-39	togtattcaggcgtacctto	ta 56.5	88	1R1 1	348-368	atgtgg	attaco	ttacaag	c	56.7	90	351	62.0
1F1 1 18-39	togtattcaggogtacctto	tg 56.5	88	1R2 1	264-286	tatta	ttactto	tcootaa	aat	55.3	80	269	61.0
1F1 1 18-39	togtattcaggogtacctto	tg 56.5	88	1R3 1	161-180	atgato	ccctcat	gtccgta	55.5	82	163	61.0	
1F1 1 18-39	togtattcaggogtacctto	tg 56.5	88	1R4 1	136-157	tattt	atcgaca	ccttcat	cc	55.2	80	140	60.0
1F3 1 53-73	cgttacttcgttgcggatag	a 56.6	87	1R1 1	348-368	atgtgg	attaco	ttacaag	c	56.7	90	316	62.0
1F3 1 53-73	cgttacttcgttgcggatag	a 56.6	87	1R2 1	264-286	tatta	ttactto	tcggtaa	aat	55.3	80	234	61.0
1F3 1 53-73	cgttacttcgttgcggatad	a 56.6	87	1R3 1	161-180	atgato	ccctcat	gtccgta	55.5	82	128	60.0	
1F3 1 53-73	cgttacttcgttgcggatag	g 56.6	87	1R4 1	136-157	tgttt	atcgaca	ccttcgt	cc	55.2	80	105	60.0
1F4 1 65-85	gcggataggactatctcggt	g 57.3	83	1R1 1	348-368	atgtgg	gtttgcc	ttacaag	с	56.7	90	304	62.0
1F4_1_65-85	gcggataggactatctcggt	g 57.3	83	1R2_1	264-286	tgttg	ttactto	toggtaa	ggt	55.3	80	222	61.0
example2 2F1_1_9-28 2F2_1_58-77 2F3_1_84-105 2F4_1_96-115 2R1_1_96-115 2R3_1_67-15 2R3_1_67-28 2R4_1_56-75 2R5_1_16-35 2R6_1_5-25 Forward_Prime	gagcacataggaacgctco taatogcggcacacggagg gtgctacaacacggtag acggtagttctacgtgca gaggtagttctacgg gactacggtgttgttag agcacatatttacctcogt ctcgtgtgcggagttcta agcgttcctatggtgctcog rID Sequence (5'-3	20 57.3 20 60.8 ttc 22 20 55.3 20 55.3 ac 22 tg	-25.8 -28.1 56.8 -25.2 -25.2 56.8 55.2 -27.9 -26.6 60.7	44.5 47.6 -27.0 40.3 37.6 -27.0 -26.4 46.4 42.0 -28.9 r_Qualit	55.0 60.0 40.6 50.0 38.3 41.8 65.0 65.0 46.1 ±y(%)	92 84 50.0 82 82 50.0 45.5 87 84 57.1 Revers	92 80 88 80 82 88 90 87 84 87 87	88 86 80 87 erID	Sequen	ce (5'-3	(')	Tm(°C)	Primer_Quality(%)
PCR_Fr	agment_Size(bp)	Topt(°C)									<i></i>		
2F1_1_9-28	gagcacataggaacgctcca	57.3 92	2R1_1 2P2_1	96-115	ttggac	gtgaga	tattaccgt	55.3	82 56 9	107	60.0	61 0	
ig. 3 Prime	r design result v	vindow		0	Ċ		5						
				Y V									

Length range (12-500 bases)	Special settings for primers or probes	Tm of oligo calculation parameters
Min: 22 Max 26	C Default primers	K+ (and Na+, NH4+, Tris+) concentration (mM): 50
Tm range (20.72%)	Best primers	Mg2+ concentration (mM): 0.0
Tm range (20-72°C)	C Long-distance PCR	DMSO concentration (%): 0.0
Min: 60 Max: 69	C Quantitative PCR	Oligo concentration (µM): 0.25
Tm at 3'-end (10-50°C)	C Microarray design	Ends nucleotide composition
Min: 40 Max: 50	C Low CG% sequence	6' 00 cow owo woo 1
CG content range (20-75%)	C Degenerate sequence	Minimal size is 2 letter (maximum - primer length) 5'-nn-3' is
Min: 40 Max 70	C Any primers	any, 5'-sww-3' - all variants for 5'-(C/G)(A/T)(A/T)-3'
Non-specific primer(probe) binding te	st parameters	
Load dataset file with reference DNA	sequence(s):	Primer qualty limit (0-100%): 70
		Primer linguistic complexity limit (0-100%): 75
C Synchronizing with dataset by seq	uence names	
C Excluding synchronizing with data	sat by sequence names	Primer's Tm optimisation

Fig. 4 The "PCR Primers or Probes Design Options" window

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

"{-ln20-23 -tm55-68 -3tm37-50 -cg41-70 -q70 446 -lc75 -npr400 -c5[NN] -c3[SWW SSW SWS WSS]}," 447 where "-ln20-23" determines the range of primer length (20-448 23 bases), "-tm55-68" determines the range of primer $T_{\rm m}$ (55-449 68 °C), "-3tm37-50" determines the range of primer $T_{\rm m}$ at 3' 450 end (37-50 °C), "-cg41-70" determines the range of primer 451 CG% contents (between 41 and 70 %), "-npr400" shows the 452 maximum number of primers (400) designed to each target, 453 "-c5[NN]" denotes a primer having no specific sequence pattern 454 for 5' ends, and "-c3[SWW SSW SWS WSS]" specifies primers 455 that conform to particular patterns of ambiguity, such as that 456 shown here for example. 457

5.3 Examples for Users can specify, individually for each sequence, multiple locations 458 **Primer Selection** for both forward and reverse primer placement with the commands: 459 "-FpdN1-N2" for forward primers and "-RpdN1-N2" for reverse Region 460 primers, where from N1 to N2 are bases from the selected regions; 461 "-pdN1-N2" (see more at: http://primerdigital.com/soft/pcr_ 462 help.html). Alternatively, users can specify multiple locations for 463 both forward and reverse primers using ["and"] inside each 464 sequence: the software allows multiple and independent locations 465 for both forward and reverse primers inside each of the sequences. 466 Whilst PCR primer design will be performed independently for 467 different targets, multiplex PCR primer design can be performed 468 simultaneously with multiple amplicons within a single sequence as 469 well as for different sequences, i.e., all possible combinations of 470 ["and"] inside one or more sequences. By default, the software 471 designs primers within the entire sequence length. Optionally, 472 users can specify, individually for each sequence, multiple locations 473 for both forward and reverse primers with the commands: 474 1. The same location for both forward and reverse primers will be 475 designed in the central [nnnnnnnn] part ("[]" is used 476 only once): 477[nnnnnn].... 478 2. Different locations for forward and reverse primers; forward 479 primers will be chosen inside the "[1nnnnn]" location and 480 reverse primers inside "[2nnnnn]" location (twice "[]"): 481[lnnnnn]....[2nnnnn].... 482

- 3. Primers must flank the central "]nnnnnn["; forward primers 483 will be chosen from 1 to "A]" bases and reverse primers will be 484 chosen from base "[C" to the end of sequence:A] 485 nnnnnn[C.....
- 4. Forward and reverse primers have an overlapping part 487 "[nnnnnn]"; forward primers will be chosen from "[A to n]" 488 bases and reverse primers will be chosen from "[n base to C]": 489[A....[nnnnn].....C].... 490



Ruslan Kalendar et al.

491		The software allows the selection of any number of independent	
492		PCR primer (or probe) designing tasks for each sequence using mul-	
493		tiple combinations of "[]" and -FpdN1-N2, -RpdN1-N2, or	
494		-point-n2 commands. Multiplex FOR can be carried out simulta-	[AU4]
495		different sequences with multiple tasks or a combination of both	
490		All possible combinations of "[1]" (forward) with "[1]"	
497		(reverse) within the sequence(s):	
499			
500		2.] [
501		3. [] []	
502		4. [[]]	
503		5.([] [])n or/and ([[]])n.	
504	5.4 User-Defined	The PCR product size can be specified in a similar way, with the	
505	PCR Product Size	command: "(N1-N2)"; these values can be specified in the form	
506		of minimum and maximum values for the product size. For	
507		example, the "(400-500)" line defines that the product size	
508		ranges from 400 to 500 base pairs. In case a user wants to specify	
509		a fixed product size, the command should be a single number, for	
510		example, "(500)." FastPCR is flexible and allows PCR product	
511		sizes from 50 to 10,000 base pairs in length.	
512	5.5 PCR Set-up	1. Where primers have already been designed, FastPCR can be	
513	Examples with	used to predict the optimal annealing temperature and PCR	
514	Individual Commands	product length for one or more predesigned primers (with the	
515		"-npd" command, which prohibits the primer design). For	
516		example, the command:	
517		"-fpr[ggagagtagcttacctcgct cggtaaggttct-tcatgc]	
518		-npd"	
519		will analyze the selected sequence between the two primers (5'	
520		ggagagtagcttacctcgct and $5'$ cggtaaggttcttcatgc).	
521		2. Design primers with a difference in T_m of about 10°, e.g., for	
522		LATE-PCR:	
523		"-Ftm50-55 -Rtm64-68 -pTMs10."	
= 0 /			
524		3. Design primers with a specific restriction enzyme site at the $3'$	[AU5]
524 525		3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e	[AU5]
524 525 526		3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e NameEnzyme," where "NameEnzyme" is the name of the	[AU5]
524 525 526 527		3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e NameEnzyme," where "NameEnzyme" is the name of the restriction enzyme: "-z3eXceI." The alternative command ("-c3NN") is also used for a special primer location. For symp	[AU5]
524 525 526 527 528		3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e NameEnzyme," where "NameEnzyme" is the name of the restriction enzyme: "-z3eXceI." The alternative command ("-c3NN") is also used for a special primer location. For exam- ple. "-c3YCATCP" is the same as "-z3eYceI"; both com	[AU5]
524 525 526 527 528 529 530		3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e NameEnzyme," where "NameEnzyme" is the name of the restriction enzyme: "-z3eXceI." The alternative command ("-c3NN") is also used for a special primer location. For exam- ple, "-c3YCATGR" is the same as "-z3eXceI": both com- mands will design primers with the recognition site for Xce I	[AU5]
524 525 526 527 528 529 530 531		3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e NameEnzyme," where "NameEnzyme" is the name of the restriction enzyme: "-z3eXceI." The alternative command ("-c3NN") is also used for a special primer location. For exam- ple, "-c3YCATGR" is the same as "-z3eXceI": both com- mands will design primers with the recognition site for <i>Xce I</i> endonucleases 5'-YCATGR-3' at the 3' end of the primers	[AU5]
524 525 526 527 528 529 530 531		 3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e NameEnzyme," where "NameEnzyme" is the name of the restriction enzyme: "-z3eXceI." The alternative command ("-c3NN") is also used for a special primer location. For example, "-c3YCATGR" is the same as "-z3eXceI": both commands will design primers with the recognition site for <i>Xce I</i> endonucleases 5'-YCATGR-3' at the 3' end of the primers. 4. Additional bases can be added to primer ends using the commands 	[AU5]
524 525 526 527 528 529 530 531 532 532		 3. Design primers with a specific restriction enzyme site at the 3' end ("-z3eNameEnzyme," "-Fz3eNameEnzyme," "-Rz3e NameEnzyme," where "NameEnzyme" is the name of the restriction enzyme: "-z3eXceI." The alternative command ("-c3NN") is also used for a special primer location. For example, "-c3YCATGR" is the same as "-z3eXceI": both commands will design primers with the recognition site for <i>Xce I</i> endonucleases 5'-YCATGR-3' at the 3' end of the primers. 4. Additional bases can be added to primer ends using the commands "-5eNN" or "-3eNN" where 5 or 3 denotes the end at 	[AU5]

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

which to add the extra bases and "NN" is a given sequence of 534 one to more bases. For example, "-F5eCGACG -R5eTTTTT," 535 means adding the sequence "CGACG" to forward primers and 536 "TTTTTT" to reverse primers, both at the 5′ ends. 537

5.6 Bisulphite-
Modified DNAThe "C \gg T bisulphite conversion" option allows the design of
specific PCR primers for in silico bisulphite conversion for both
strands. Only cytosine not followed by guanidine (CpG 540
methylation) will be replaced by thymine:

5'aaCGaagtCC 3' 5'aaCGaagtTT 3' ||||||||| -> |||||| | 3'ttGCttCagg 5' 3'ttGCttTagg 5' 542

5.7 Uniqueness Optionally, the user can synchronize the primer test for secondary, 543 of Primers nonspecific binding with a dataset of sequence names. The 544 program recognizes that a given sequence in the screening library 545 dataset (from loading the dataset file) is the same as the sequence 546 for which it is designing primers; this allows sequence selection to 547 be made even if the selection matches the screening sequence 548 perfectly. This allows the same dataset to be used for both primer 549 design and screening without having to make a new screening 550 database for each sequence. In other words, a dataset that contains 551 sequences A, B, C, and D can be used both for choosing primers 552 and for checking primer specificity. Alternatively, the user can 553 input preexisting primers into a second "Additional sequence(s) 554 or pre-designed primers (probes) list" text editor. These primers 555 or probes will be checked for compatibility (inter-primer-dimer 556 formation) with newly designed primers. The number of 557 preexisting primers is not limited to one or two; it can be as many 558 as the user needs. 559

5.8 PCR Primer Design The PCR primer design algorithm generates a set of primers having 560 a high likelihood of success in virtually any amplification protocol. 561 All PCR primers designed by FastPCR can be used for PCR or 562 sequencing experiments. The program is able to generate either 563 long oligonucleotides or PCR primers for the amplification of 564 gene-specific DNA fragments of user-defined length. FastPCR 565 provides a flexible approach to designing primers for many 566 applications and for both linear and circular sequences. It will 567 check if either primers or probes have secondary binding sites in 568 the input sequences that may give rise to additional PCR products. 569 The selection of the optimal target region for the design of long 570 oligonucleotides is performed in the same way as for PCR primers. 571 The basic parameters in primer design are also used to measure the 572 oligonucleotide quality and to evaluate the thermodynamic stability 573 of the 3' and 5' terminal bases. 574

Author's Proof

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

Ruslan Kalendar et al.

PCR Primer Design in silico PCR Primer Test Primers List Analysis Restriction Clustering Searching LTR Search MITE Search	ch SSR Search Tools Oligos Assembly
Parameters for PCR product analysis: PCR primer design options:	
Synchronizing Tm(*C) and dG(kcal/mol) for primer pair (±): 5 The secondary (non-specific) binding test	Inverted PCR
Limit for compatible combination of pair primers: 500	Circular DNA
Polymerase extension cloning (OE-PCR) Overlapping primers	
Villiplex PCR	Unique PCR
Minimal difference between multiplex PCR products (bp) 10	Group-specific PCR
Maximal difference between Ta of multiplex PCR products (+*C)	
Limit for multiplex PCR compatible combination of pair primers: 500	
Construction manageer for comparison containing on primeros, 500	atible pair primers: 500
Sequences: 1: 2048 Additional sequence(s) or pre-designed primers (probes) list PCR primers design result Multiplexes PCR compa	aubie pair primers. 500
Forward_PrimerID Sequence(5'-3') Im('C) Primer_Quality(%) Reverse_PrimerID Sequence(5'-3') PCR Fragment Size(bp) Topt('C)	Im("C) Primer_Quality(%)
1F1_1_55-79 ccaattcgacggtaacatcgtgcct 61.3 94 1R1_1_325-350 ggcttgtctagttaaggtaaggtcca 60.3 80	296 66.0
IFI_2_376-400 cttaaacatccgtcggcctttagg 60.5 87 IRI_2_676-700 aatccgtcagggcaacgaccatcg 64.7 91	325 66.0
IF1_4_1019-1044 gCCCaccCcgCccaccCcggaaccc 05.0 07 IR1_4_1204-1200 CCCgaccgcagcaaccacggggga	61 1 63 347 67 0
151 6 1684-1707 charageastacgettcaccec 60.9 87 182 6 1962-1964 gatetcagattcaccecettage	61.8 87 281 67.0
1F1 7 107-129 ccggaaagccggtatcgttcca 63.9 87 68.0	
1R2 8 1512-1534 gccttgtcaaccgatgaaggacg 61.3 80 65.0	
1R3_9_1289-1312 cccgcttcttgatgctcaaaacca 60.5 80 64.0	
1F1 1 55-79 ccaattcgacggtaacatcgtgcct 61.3 94 1R1 1 325-350 ggcttgtctagttaaggtaaggccca 60.3 80	296 66.0
1F1 2 376-400 cttaaacatcccdtcggcctttagg 60.5 87 1R1 2 676-700 aatccdtcagggcaaggaccactcg 64.7 91	325 66.0
1F1 4 1019-1044 gtttacctgcccacactccggatacc 63.6 87 1R1 4 1264-1288 cttgaactgtagcaaacatgcgcga	60.3 80 270 66.0
1F1_5_1310-1333 gggatcaggttctctggctagacg 61.3 93 1R2_5_1634-1656 gcacgtcatcttggacctgttcg	61.1 93 347 67.0
1F1_6_1684-1707 cagagagaatacggcttcaccgca 60.9 87 1R2_6_1942-1964 gagctctggtatccccgcttagg	61.8 87 281 67.0
1F1_7_107-129 ccggaaagcccgtgatcgttcca 63.9 87 68.0	
1R2_8_1512-1534 gccttgtcaaccgatgaaggacg 61.3 80 65.0	
1R4_9_1264-1288 cttgaactgtagcaaacatgcgcga 60.3 80 64.0	
1F1_1_55-79 ccaattcgacggtaacatcgtgcct 61.3 94 1R1_1_325-350 ggcttgtctagttaaggtaaggtcca 60.3 80	296 66.0
1F1_2_376-400 cttaaacatcccgtcggcctttagg 60.5 87 1R1_2_676-700 aatccgtcagggcaaggaccactcg 64.7 91	325 66.0
1F1_4_1019-1044 gtttacctgcccacactccggatacc 63.6 87 1R1_4_1264-1288 cttgaactgtagcaaacatgcgcga	60.3 80 270 66.0
1F1_5_1310-1333 gggatcaggttctctggctagacg 61.3 93 1R2_5_1634-1656 gcacgtcatcttggacctgttcg	61.1 93 347 67.0
1F1_6_1684-1707 cagagagaatacggcttcaccgca 60.9 87 1R2_6_1942-1964 gagctctggtatccccgcttagg	61.8 87 281 67.0
1F1_7_107-129 ccggaaagccortgatcgttcca 63.9 87 68.0	
102_0_1012_0_1014 gcccugtcaaccgatgaaggacg 0.1.3 80 05.0	×
1: 2048 nt A=502.0 T=519.0 G=501.0 C=526.0 R=502.0 Y=1045.0 R/Y=0.48 CG=50.1% Tm=85 Open FASTA Sequence	es with a '>' symbol 🛛 🔽 Reading sequence(s)

Fig. 5 An example of multiplex PCR result

Both the proposal of primer pairs and the selection of the best pairs are possible. The user can vary the product size or design primer pairs for the whole sequence without specifying parameters by using default or preset parameters. Preset parameters are specified for various situations: for example, for sequences with low CG content, for long-distance PCR, for degenerate sequences, or for manual input. A list of the best primer candidates and all compatible primer pairs that are optimal for PCR is generated. Users can specify, individually for each sequence, multiple locations for both forward and reverse primers within each sequence, whilst PCR design will be performed independently for different targets. Primers for multiplex PCRs can be designed from a single or from multiple targets (Fig. 5).

The program generates primer pairs (and probes) from the input sequences and shows the optimal annealing temperature for each primer pair and the sizes of PCR products together with information for each designed primer. Suggested primers and primer pairs are produced in tabulated format for MS Excel or Open Office (Table 3). The spreadsheets show the following properties: the automatically generated primer name, primer sequence, sequence location, direction, length, melting temperature, CG content (%), molecular weight, molar extinction coefficient, linguistic complexity (%), and *PQ*. For compatible primer pairs, the annealing temperature and PCR product size are also provided.

ta.a Primer ID	Sequence (5'-3')	Length (nt)	T _m (°C)	dG, kcal/mol	T _m 3′end (°C)	% 9 0	DT	PQ
t3.41F1_1_6-27	gggcggatcacttgaggtcagg	22	63.0	-29.7	40.0	63.6	88	87
t3.51F2_1_107-132	ggcaggagaatcacttcaacctggga	26	62.5	-33.6	40.2	53.8	89	89
t3.61F3_1_133-154	ggcggaggttgcagtgaaccga	22	64.4	-31.3	43.9	63.6	90	90
t3.71F4_1_153-175	gagaccgcgccactgcactccag	23	67.2	-33.7	42.8	69.69	76	76
t3.81F5_1_166-187	tgcactccagcctgggcgacag	22	66.3	-32.1	49.5	68.2	85	85
t3.91F6_1_176-197	cctgggcgacagcctgagactc	22	64.8	-30.8	41.4	68.2	80	80
t3.1dF1_2_859-880	cttacggaggccgagatgggca	22	63.4	-30.6	46.9	63.6	88	87
t3.11F2_2_913-934	tggccaacatggtgaaaccctg	22	60.9	-28.9	41.6	54.5	82	80
t3.1 1 F3_2_956-977	aattagetgggcatggtggcac	22	60.9	-29.1	47.8	54.5	80	73
t3.13 F4_2_101 3-1036	tggagctgaaccactgcactccag	24	63.2	-32.3	42.8	58.3	79	79
t3.14 R1_1_412-434	acccagaagagcctgagtgggca	23	63.9	-31.7	46.7	6.09	83	83
t3.1 f R2_1_309-330	ccttgctcagctctggccatcc	22	63.6	-30.0	44.8	63.6	80	80
t3.1dR3_1_299-320	ctctggccatcccagttcaagc	22	61.5	-28.9	42.2	59.1	88	80
t3.11R1_2_1206-1231	agatggggtttcaccatgttggccca	26	64.1	-34.8	44.6	53.8	80	80
t3.1d R2_2_1166-1187	acctcaggtgatccacctgcct	22	61.7	-29.5	44.2	59.1	82	80
t3.1dR3_2_1150-1174	cacctgcctcagcttcccaaagtgc	25	65.3	-34.2	40.6	60.0	84	84
t3.2dR4_2_1132-1157	caaagtgcttggattacaggcgtgag	26	61.0	-33.0	42.4	50.0	93	93
t3.2The separated output	of a set of primer pair sequence	es with their th	eoretical P	CR products				
								(continued)

 $\mathfrak{B}.\mathbf{P}$ rogram output of the primer design t3.Table 3

Primer ID	Sequence (5'–3')	Length (nt)	T _m (°C)	dG, kcal/mol	<i>T</i> _m 3'end (°C)	% 9 0	27	PQ	
3.2Forward primer ID 13.23	Sequence (5'-3')	$T_{\rm m}$ (°C)	PQ	Reverse primer ID	Sequence (5'-3')	$T_{\rm m}$ (°C)	ЪQ	PCR Fragment Size (bp)	T ^{opt} (°C)
t3.24 F1_1_8-27	gcggatcacttgaggtcagg	58.9	87	1R3_1_301-320	ctctggccatcccagttcaa	57.6	80	313	63
t3.2 đ F1_1_8-27	gcggatcacttgaggtcagg	58.9	87	1R4_1_291-310	cccagttcaagccatcccct	60.2	76	303	65
t3.2d F2_1_49-68	caacgtggagctaggtatgg	56.0	80	1R1_1_409-429	gaagagcctgagtggggcacaa	60.09	85	381	62
t3.21 F2_1_49-68	caacgtggagctaggtatgg	56.0	80	1R3_1_301-320	ctctggccatcccagttcaa	57.6	80	272	62
t3.2d F2_1_49-68	caacgtggagctaggtatgg	56.0	80	1R4_1_291-310	cccagttcaagccatcccct	60.2	76	262	62
t3.24 F3_1_109-130	caggagaatcacttcaacctgg	56.4	87	1R1_1_409-429	gaagagcctgagtggggcacaa	60.09	85	321	62
t3.3d F3_1_109-130	caggagaatcacttcaacctgg	56.4	87	1R3_1_301-320	ctctggccatcccagttcaa	57.6	80	212	62
t3.31F3_1_109-130	caggagaatcacttcaacctgg	56.4	87	1R4_1_291-310	cccagttcaagccatcccct	60.2	76	202	62
t3.3 3 F4_1_134-154	gcggaggttgcagtgaaccga	62.6	06	1R1_1_409-429	gaagagcctgagtggggcacaa	60.09	85	296	66
t3.33 F4_1_134-154	gcggaggttgcagtgaaccga	62.6	06	1R2_1_311-333	tgtccttgctcagctctggccat	62.3	83	200	68
t3.34 F4_1_134-154	gcggaggttgcagtgaaccga	62.6	06	1R3_1_301-320	ctctggccatcccagttcaa	57.6	80	187	63
t3.3 đ F4_1_134-154	gcggaggttgcagtgaaccga	62.6	06	1R4_1_291-310	cccagttcaagccatcccct	60.2	76	177	65
t3.3d F5_1_153-172	gagaccgcgccactgcactc	64.3	73	1R2_1_311-333	tgtccttgctcagctctggccat	62.3	83	181	68
t3.3 1 F5_1_153-172	gagaccgccactgcactc	64.3	73	1R4_1_291-310	cccagttcaagccatcccct	60.2	76	158	65
					5				

Table 3 (continued) 5.9 Multiplex PCR Primer Design

Multiplex PCR is an approach commonly used to amplify several 599 DNA target regions in a single reaction. The simultaneous 600 amplification of many targets reduces the number of reactions that 601 needs to be performed; multiplex PCR thus increases throughput 602 efficiency. The design of multiplex PCR assays can be difficult 603 because it involves extensive computational analyses of primer pairs 604 for interactions. The multiplex PCR algorithm is based on the fast 605 non-recursion method, with the software performing checks on 606 product size and primers' thermodynamics parameters (enthalpy-607 dH and Gibb's free energy-dG) compatibility and cross-dimer 608 formation for all primers. To achieve uniform amplification of the 609 targets, the primers must be designed to bind with equal efficiencies 610 to their targets. FastPCR can quickly design a set of multiplex PCR 611 primers for all input sequences and/or multiplex targets within 612 each sequence. PCR conditions may need to be adjusted, for 613 example, by increasing or decreasing the annealing temperature so 614 that all products are amplified equally efficiently. 615

To achieve uniform amplification, most existing multiplex 616 primer design packages use primer melting temperature. In practi-617 cal terms, the design of almost identical T_{a} s and T_{m} s is very impor-618 tant. The melting temperatures of the PCR products are also 619 important because these are related to annealing temperature val-620 ues. The T_m of a PCR product directly depends on its CG content 621 and its length; short products are more efficiently amplified at low 622 PCR annealing temperatures (100 bp, 50–55 °C) than are long 623 products (>3,000 bp, 65–72 °C). For most multiplex PCRs, there 624 is usually a small variation (up to 5 °C) between the optimal T_{a} s of 625 all primer pairs. The annealing temperature must be optimal in 626 order to maximize the likelihood of amplifying the target genomic 627 sequences whilst minimizing the risk of nonspecific amplification. 628 Further improvements can be achieved by selecting the optimal set 629 of primers that maximize the range of common $T_{\rm m}$ s. Once 630 prompted, FastPCR calculates multiplex PCR primer pairs for 631 given target sequences. The speed of calculation depends on the 632 numbers of target sequences and primer pairs involved. 633

An alternative way to design compatible multiplex PCR primer 634 pairs is to use predesigned primers as references for the design of 635 new primers. The user can select input options for the PCR prod-636 ucts such as the minimum product size differences between the 637 amplicons. Primer design conditions can be set individually for 638 each given sequence or all primers can be designed using the same 639 values; selected settings have higher priority for PCR primer or 640 probe design than the general settings. The results include primers 641 for individual sequences, compatible primers, product sizes, and 642 annealing temperatures. Because clear differentiation of the prod-643 ucts is dependent on using compatible primer pairs in the single 644 reactions, the program recovers all potential variants of primer 645 combinations for analyses of the chosen DNA regions and provides, 646

Ruslan Kalendar et al.

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

in tabular form, their compatibility including information on primer-dimers, cross-hybridization, product size overlaps, and similar alternative primer pairs based on $T_{\rm m}$. The user may choose those alternative compatible primer pair combinations that provide the desired product sizes. Using the program, researchers can select predesigned primer pairs from a target for their desired types of PCRs by changing the filtering conditions as mentioned above. For example, a conventional multiplex PCR requires differently sized (at least by 10 bp) amplicons for a set of target genes, so the value for the minimum size difference between PCR products can be selected.

In addition to avoid amplifying different amplicons of the same size, multiplex PCR must also minimize the generation of primerdimers and secondary products, which becomes more difficult with increasing numbers of primers in a reaction. To avoid the problem of nonspecific amplification, FastPCR selects primer pairs that give the most likelihood of producing only the amplicons of the target sequences by choosing sequences which avoid repeats or other motifs. The program allows the user to design not only compatible pairs of primers but also compatible single primers for different targets or sequences. The input data can be either a single sequence with a minimum two internal tasks or many sequences with or without internal tasks. Most of the parameters on the interface are self-explanatory. Optionally, the user is asked to provide the sequence and select oligonucleotide designing parameters.

On the PCR Primer Design tab, the user clicks on the Multiplex PCR option. The user then selects the limit for the number of primer pairs (the default is 100), the minimal size difference between multiplex PCR products (the default 10 bp), and the maximal difference between the T_a s of the PCR products (the default is ±5 °C). After specifying inputs and primer design options, the user can execute the primer design task. Once the design of the primer set is completed, the result will appear in two **Result** text editors: PCR primer design result and Multiplex PCR compatible pair primers. Figure 8 shows the access to the PCR primer design output. The result text editor PCR primer design result will display the individual PCR primer design data, including the primer list and the compatible primer pairs for all the sequences and their internal tasks. The second Multiplex PCR compatible pair primers text editor collects final search results that are presented as a list of the sets of the compatible primer pairs for multiplex PCR.

689	5.10	Group-Specific
690	PCR F	Primers
691		
692		
693		

Group-specific amplification, also called family-specific and sequence-specific amplification, is an important tool for comparative studies of related genes, sequences, and genomes that can be applied to studies of evolution, especially for gene families and for cloning new related sequences. Specific targets such as homologous genes or members of a transposable element family can be amplified to 694 uncover DNA polymorphisms associated with these sequences or 695 other genetic investigations. The overall strategy of designing 696 group-specific PCR primers uses a hash index of 12-mers to identify 697 common regions in the target sequences, followed by standard 698 PCR primer design for the current sequence, and then testing of 699 complementarity of these primers to the other sequences. FastPCR 700 performs sequences multiple alignment or does accept alignment 701 sequences input, giving it the flexibility to use a different strategy 702 for primer design. If required, it can design degenerate PCR primers 703 to amplify the polymorphic region of all related sequences. 704

The FastPCR package designs large sets of universal primer 705 pairs for each given sequence, identifies conserved regions, and 706 generates suitable primers for all given targets. The steps of the 707 algorithm are performed automatically and the user can influence 708 the settings for the primer design options. The quality of primer 709 design is dependent on sequence relationships, genetic similarity, 710 and suitability of the consensus sequence for the design of good 711 primers. The software is able to generate group-specific primers for 712 each set of sequences independently, which are suitable for all 713 sequences. Primer alignment parameters for group-specific PCR 714 primers are similar to those used for in silico PCR. The user chooses 715 Group-specific PCR on the PCR Primer Design tab. After speci-716 fying inputs and PCR primer design options, the user can execute 717 the PCR primer design task. The program takes multiple separate 718 DNA sequences either in FASTA or at alignment formats. 719

Once the primer set design is complete, the result will appear 720 in Result text editor, as the PCR primer design result. Figure 6 721 shows the access to PCR primer design output (Table 3). The 722 result text editor PCR primer design result displays the individual 723 group-specific PCR primer design data, including the primer list 724 and compatible primer pairs for all the sequences and their internal 725 tasks where suitable primers are found. In the case where an align-726 ment has been input, the result text editor displays only one 727 group-specific PCR primer design set, including degenerate prim-728 ers, in the primer list as well compatible primer pairs for all the 729 sequences. 730

5.11 Simple Sequence Repeat Locus Search and PCR Primer Design SSRs or microsatellites are short tandem repeats of one or more 731 bases. Microsatellites are ubiquitously distributed throughout 732 eukaryotic genomes, often highly polymorphic in length, and 733 thereby an important class of markers for population genetic 734 studies. Our approach to locating SSRs is to analyze low-complexity 735 regions in DNA by using linguistic sequence complexity. This 736 method allows the detection of perfect and imperfect SSRs with a 737 single, up to 10-base, repeat motif. Each entry sequence is 738 processed for identification of SSRs and the SSR flanking regions 739 are used to design compatible forward and reverse primers for their 740 amplification by PCR. 741

[AU6]

Author's Proof

Ruslan Kalendar et al.

Sequences: 2 :	3945 Additional sequence(s) or p	ore-design	ed prime	ers (prob	es) list	PCR pri	imers des	ign resul	t Multiple	exes PC	R compa	tible pair	r primers	: 500	
PrimerID	Sequence(5'-3') Lengtl	h(bp)	Tm(°C)	dG(kca	1/mol)	Tm 3'	end(°C)	CG(%)	Linguis	tic Co	mplexit	y(\$)	Primer	Quality(%)	
1F1_1_19-39	cgttatgacaaagttcccggt	21	55.6	-25.9	41.3	47.6	92	80		-	-				
1F2_1_42-62	aaatggattctgagtcggtgg	21	55.3	-25.4	44.5	47.6	85	80							
1F3_1_51-73	ctgagtcggtggaattgggcacg	23	63.3	-31.3	43.4	60.9	93	93							
1F4_1_81-103	tcaggctgaaaccttgagcatcg	23	59.8	-29.8	41.9	52.2	93	93							
1F5_1_142-163	ttagagttcgttgaacgtggca	22	56.3	-27.6	42.8	45.5	88	88							
1F6_1_153-172	tgaacgtggcactggattgg 20	58.5	-26.3	42.4	55.0	84	84								
1F7_1_191-213	ccatagaagggttaagaaccgca	23	57.1	-28.0	37.7	47.8	85	85							
181_1_1024-10	44 agcagataagtcactctgc	gt	21	55.4	-26.1	43.3	47.6	92	80						
182 1 1007-10	2/ gcgtgtggaaagtgggtac	at	21	55.1	-27.0	36.0	52.9	02	01						
104 1 070-000	agtgggtacatgtcaatgc	acg so o	-26 5	37.9	-2/.9	91.2	50.0	00	0/						
185 1 967-988	ttaccactagtacctatascaa	22	61.9	-29.9	43.3	59.1	88	87							
1R6 1 956-976	cctgtgacagcatttaacgga	21	55.6	-25.8	36.6	47.6	97	82							
1R7 1 943-962	taacggaacgttttgggcat 20	55.1	-25.1	41.8	45.0	82	80								
1R8 1 922-941	tgcccatctctcaacgtgtt 20	56.3	-25.6	37.6	50.0	84	84								
1R9_1_892-912	tcgtttacgggaggtcttgaa	21	55.4	-26.0	37.9	47.6	90	85							
1R10_1_867-88	17 tcggagttttgtattgcgc	ct	21	56.9	-26.9	42.1	47.6	87	87						
1R11_1_834-85	3 gtaggtacgcccattgggt	c 20	58.8	-25.8	46.4	60.0	87	87							
1R12_1_811-83	acaaattetgtettggegg	ag	21	56.0	-25.9	45.6	47.6	92	92						
L		-				-			-						
Forward_Prime	rID Sequence(5'-3')	Tm(°C)	Primer	_Qualit	Y(\$)	Rever	se_Prime	erID	Sequenc	ce (5'-3	3*)	Tm(°C)	Primer	C_Quality(%))
PCR_Fr	agment_Size(bp) Topt(°C)										~~		<i>co o</i>	
181_1_19-39	cgttatgacaaagttcccggt	55.6	80	181_1_	1024-10	199	agcaga	taagtca	ctctgcg	5	55.4	80	1026	62.0	
121_1_19-39	cgttatgacaaagttcccggt	55.0	80	1010 1	067-00	27	gegege	ggaaagt	gggtaca	5	50.1	01	1009	63.0	
181 1 19-39	cottatgacaaagttccccggt	55.6	80	1012 1	811-83	1	acaaat	tetatet	tagegee	~	56.0	92	813	62.0	
1F2 1 42-62	aaatggattctgagtcggtg	55.3	80	1R1 1	1024-10	44	accada	taagtca	ctctaca	9 T.	55.4	80	1003	62.0	
1F2 1 42-62	aaatggattetgagteggtgg	55.3	80	1R2 1	1007-10	27	acatat	ggaaagt	gggtaca		58.1	81	986	62.0	
1F2 1 42-62	aaatggattctgagtcggtgg	55.3	80	1R3 1	996-101	7	agtggg	tacatgt	caatgca	ca	57.9	87	976	62.0	
1F2 1 42-62	aaatggattctgagtcggtgg	55.3	80	1R10 1	867-88	7	toggag	ttttgta	ttgcgcc	t	56.9	87	846	62.0	
1F2_1_42-62	aaatggattctgagtcggtgg	55.3	80	1R12_1	811-83	1	acaaat	tetgtet	tggcgga	a	56.0	92	790	62.0	
1F4_1_81-103	tcaggctgaaaccttgagcatcg	59.8	93	1R1_1_	1024-10	44	agcaga	taagtca	ctctgcg	t	55.4	80	964	62.0	
1F4_1_81-103	tcaggctgaaaccttgagcatcg	59.8	93	1R2_1_	1007-10	27	gcgtgt	ggaaagt	gggtaca	5	58.1	81	947	65.0	
1F4_1_81-103	tcaggctgaaaccttgagcatcg	59.8	93	1R3_1_	996-101	.7	agtggg	tacatgt	caatgca	cg	57.9	87	937	65.0	
1F4_1_81-103	tcaggctgaaaccttgagcatcg	59.8	93	1R5_1_	967-988	ttaco	gctggtgd	ctgtgac	ag	61.9	87	908	67.0		
1F4_1_81-103	tcaggctgaaaccttgagcatcg	59.8	93	1R6_1_	956-976	cctgt	gacagcat	ttaacgg	ja.	55.6	82	896	62.0		
1F4_1_81-103	tcaggctgaaaccttgagcatcg	59.8	93	1R9_1_	892-912	togtt	tacgggag	gtettga	a	55.4	85	832	62.0		
184_1_81-103	tcaggctgaaaccttgagcatcg	59.8	93	1810_1	_867-88	2	toggag	ttttgta	ttgcgcc	5 50 0	56.9	87	807	64.0	
124 1 01-103	teaggetgaaacettgageateg	59.5	33	1P12 1	_034-85	13	gtaggt	acgccca	Legggte	35.5	56 0	02	05.0	63 0	
185 1 142-149	traggetgaaacettgageateg	56.3	99	101 1	1024-10	44	acadāt	Taagtot	-ugcgga	-	55.0	92	003	62.0	
185 1 142-163	tragagteegtegaacgeggca	56.3	88	182 1	1007-10	27	ageaga	adaguCa	agataca		58.1	81	886	63.0	-
1110_1_142-103		0010		102_1_			20000		AAAcaca	•		**			
2:3945 nt	A=1007.0 T=1009.0	G=969.0	C=960.0	R=1007	.0 Y=196	9.0 R/Y	=0.511 CO	G=48.9%	Tm={ Op	en FAS1	TA Seque	nces wit	h a '>' sy	mbol	Reading sequence(s)

Fig. 6 An example of group-specific PCR result

FastPCR identifies all SSRs within each entry sequence and designs compatible PCR primer pairs for each SSR locus. The default PCR primer design parameters are that the primers must be within 100 bases from either side of the identified SSR. Often the sequences available around SSR loci are not suitable for designing good primers; the user can increase or decrease the distance from either side to find more efficient and compatible primer pairs. The capabilities of FastPCR make it a complete bioinformatics tool for the use of microsatellites as markers, from discovery through to primer design. For example, the user can specify PCR primer design to SSR loci within 200 bp around an SSR, with the command: "-ssr/200." The software finds all SSR sites and then will design PCR primers and compatible primer pairs independently for each SSR locus.

756	5.12 Oligonucleotide
757	Design for In Vitro
758	Long Sequence
759	Synthesis
760	
761	
762	

742

743

744

745

746

747

748

749

750

751

752

753

754

755

Ie The application to make long synthetic DNA molecules relies on the in vitro assembly of a set of short oligonucleotides, either for the LCR [37] or for assembly PCR [38]. These oligonucleotides should be adjacent on the same strand and overlap the complementary oligonucleotides from the second strand. There are two major parameters for designing oligonucleotides for gene synthesis for LCR or assembly PCR. First, the oligonucleotides

PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

should share similar $T_{\rm m}$ values. Second, a given oligonucleotide 763 sequence should be unique to avoid multiple nonspecific binding 764 that may lead to incorrect assembly. The software must dynamically 765 choose the length of the oligonucleotides to ensure both the 766 specificity and the uniform $T_{\rm m}$. The algorithm of FastPCR is able 767 to design oligonucleotides for long sequences containing repeats 768 and to minimize their potential nonspecific hybridization during 3' 769 end extension in PCR. For long sequence assembly, oligonucleotide 770 design starts from the 5' end of a given sequence; the oligonucleotide 771 length is dynamically changed until a unique 3' end has been found 772 and the $T_{\rm m}$ of the oligonucleotide has reached the $T_{\rm m}$ threshold. All 773 oligonucleotides are designed without gaps between them. The 774 other strand is used for the design of the overlapping 775 oligonucleotides using the same algorithm as above but with the 776 $T_{\rm m}$ of the overlapping regions reaching the $T_{\rm m}$ -15 °C threshold. 777 The composition of the sequence at the 3' terminus is important 778 because stability at the 3' end of the double-stranded complexes 779 will improve the specificity of extension by the polymerase. To 780 reduce nonspecific polymerase extension (and ligation), the 781 algorithm chooses only unique sequences for the 3' terminus. 782 Minimally, the last two nucleotides at the 3' terminus must not be 783 complementary to any nonspecific targets. Other complementary 784 regions are less important for assembling multiple fragments by 785 PCR and ligation. 786

The input data can comprise either a single or many sequences. 787 Most of the parameters on the interface are self-explanatory. The 788 user is asked to provide the sequence and select oligonucleotide 789 designing parameters. The user clicks on Oligo options on the 790 Oligos Assembly tab, and chooses the minimal oligonucleotide 791 length and $T_{\rm m}$ threshold, which by default are 40 nt and 60 °C, 792 respectively. The interface allows changing $T_{\rm m}$ calculation parame-793 ters. The search process runs after pressing F5 or from menu bar 794 or toolbox. The research result is presented as a list of oligonucle-795 otides for both strands. On each strand, all oligonucleotides are 796 adjacent with no gap between neighboring primers. An oligonu-797 cleotide will overlap two oligonucleotides from the complemen-798 tary strand. The algorithm pays attention to avoid nonspecific 799 oligonucleotide hybridization to repeated regions. Where it is not 800 possible to design primers outside of repeated sequences, it is like-801 wise difficult to find short specific oligonucleotides. The solution 802 to this problem is to divide the sequence into short segments, 803 design a set of oligonucleotides for each segment independently, 804 and then combine all these segments in the second PCR for final 805 amplification. 806

5.13 Polymerase Extension PCR for Fragment Assembly Sequence-independent cloning, including ligation-independent 807 cloning, requires generation of complementary single-stranded 808 overhangs in both the vector and insertion fragments. Similarly, 809

Author's Proof

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839 840

841

842

843

844

845 846

Ruslan Kalendar et al.

multiple fragments can be joined or concatenated in an ordered manner using overlapping primers in PCR. Annealing of the complementary regions between different targets in the primer overlaps allows the polymerase to synthesize a contiguous fragment containing the target sequences during thermal cycling, a process called "overlap extension PCR" (OE-PCR) (Chapter 8) [39]. The efficiency depends on the T_m and on the length and uniqueness of the overlap. To achieve this, the program designs compatible forward and reverse primers at the ends of each fragment, and then extends the 5' end of primers using sequences from the primers of the fragment that will be adjacent in the final product. The input sequence can be made of either a single or many sequences. The user needs to pay special attention to the preparation of the given sequences for assembly.

Users can specify the locations for both forward and reverse primers design using "[]" to bracket the region. The bracketed sequences will be used by the program for designing the overlapping primers. The program selects the overlapping area so that the primers from overlapping fragments are similar in size with optimal annealing temperatures. The program adds the required bases so that the T_m of the overlap is similar to, or higher than, the T_m of the initial primers. Primers are tested for dimer formation within the appropriate primer pairs. The user chooses Polymerase extension cloning (OE-PCR) on the PCR Primer Design tab and selects the limit for multiple-PCR-compatible combinations of pair primers (default is 100). After specifying sequence inputs and PCR primer design options, the user can execute the search task. Once the design of the primer sets is complete, the result will appear in two text editors: PCR primer design result and PCR fragments assembling compatible pair primers. The text editor **PCR primer design result** window displays the individual PCR primer design data, including the primer list and the compatible primer pairs for all sequences whose primers are found. The PCR fragments assembling compatible pair primers text editor collects the final search result and presents it as a list of sets of compatible primer pairs for individual fragment amplification and assembly. Figure 7 shows a sample result visualization window.

Modelling the hybridization of primers to targeted annealing sites 5.14 In Silico PCR 847 is the only way to predict PCR products [7, 24, 40–44]. The last 848 10-12 bases at the 3' end of primers are important for binding 849 stability; single mismatches can reduce PCR efficiency, the effect 850 increasing with proximity to the 3' terminus. FastPCR allows 851 simultaneous testing of single primers or a set of primers designed 852 for multiplex PCR. It performs a fast, gapless alignment to test the 853 complementarity of the primers to the target sequences. For in 854 silico PCR, a quick alignment to detect primer locations on the 855 reference sequence is performed by analyses of both strands using 856



PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

PCR Primer De	esign in silico PCR Primer Test Primers List Analysis Re	estriction Clustering Searching LTR Sea	rch MITE Search SSR Search Tools	Oligos Assem	bly	
	Parameters for PCR product analysis:	PCR primer design options:				
Synchro	onizing Tm(*C) and dG(kcal/mol) for primer pair (±): 5	 The secondary (non-specific) 	binding test			
	Limit for compatible combination of pair primers: 500	Linguistic complexity control	Circular DNA			
	Polymerase extension cloning (OE-PCR)	C Overlapping primers				
	Multiplex PCR		Unique PCR			
Minim	al difference between multiplex PCR products (bp) 10	- I carrotation	Group-specific P	CR		
Maximal diffe	rence between Ta of multiplex PCR products (+*C) 5	 Molecular beacon design 				
Limit for mult	tiniar PCR compatible combination of pair primars: 600	- Microarray design				
Sequences: 4 :	1269 Additional sequence(s) or pre-designed primers (prob	es) list PCR primers design result PCR fr	aoments assembling compatible pair prim	ers: 500		
[gaagcgctctg gggggatgaatt cctcgcccgcgc [ttgcagtggca	<pre>tccactacaccatgattactt) tacaatattoctacgaagatogtctagogettgatacgtgccgattt gatgagttgaaggatcgccttgacgtgacg</pre>	ttatotgtagoaatgatgaaaattaactttotacot sooogotagatgo	aggcagtcacagtttttatttccctaatata	gettegteett	categgttg	acaaggogaaagoca
[cgaagggttcg	ttaccggcttcaacgatcaacgattcaacccc]					
aagttccaacgt	cgaaggtggaaatttcggcggagatggctactcctattctaggacgcga	aatcccggcattgggtaacatttgcccagacgcgc	cacgggtgcggatatcccgactaggtaacgggc	aaggagaaagc	tggtaggcg	catgatgcaggaatt
[aatttattcat	accccgacgcatgttcggatttgacctggtcgagaaccgccgcatacgt	[gtaatagcagtgtcgat]	9-99 9-5			
>	agageacacactagagatgcctgggaagecct]					
aatagcccgtaa	taaccettageggeetgeggteeceeeggggagaettateatgtgtaat	cccctcccatcagtgcgcgtgatattttatgcaga	gtaaaccatgaaagataaacgacactctaacga	togtataatgt	caccocgo	gtgcgatattgtccc
gac[gaacgtac	cacattgcactttgcgtcacttggccccgatacctaaactatgctgagg	10a]				
[gaagcgctctg ggggggatgaatt cctcgcccgcgc	tccactacaccatgattactt) taacaatattcctacgcaagatcgtctagcgcttgatacgtgccgattt tgatgaagttgaagactcgccttgacgttaactagatgccggcgacacc	statetgtagcaatgatgaaaattaaettteteet seaegetagatge	aggcagtcacagtttttattttccctaatata	gettegteett	categgttg	acaaggcgaaagcca
(ttgcagtggca	ctgagttctgctcccgcaaagacact]					
Convences: 4	1250 Additional equipace(e) or are designed primare (proh	Net DCD primare design result PCR fr	anments assembling compatible pair prim	ers: 500		
Sequences. 4	Additional Sequence(s) of pre-designed primers (pro-	Port primers design result [Sort		1010.000 pcp	F	Size (he)
Topt ("C) Sequence(5-5-) Im(C) Primer_Quality	cy(4) Keverse_Frimerid Sequent	Se(S-S-) Im(C) Frimer_Quali	cy(t) FCR	_rragment	
151 1 1-20	APPROACT AT A CAPACITA A 57 1 80 101 1 285-300	5 AND DAATTANATATATATATAAAAAAAAAAAAAAAAAA	60 2 73 315 63 0			
2F2_1_1-21	gctcccgcaaagacactcgaagggttcgttaccggctt 60.1	87 2R1_1_361-386 casatcaagggate	cgacactgctattacacgtatgcg 58.6	80 409	65.0	
3F2_1_1-21	gcatacgtgtaatagcagtgtcgatcccttgatttggagagcacgc	59.9 80 3R1_1_246-271 cagage	cttctccctcagcatagtttaggtatcggg	59.4 80 328 63	302	65.0
111_1_1-10	-facacecaaaceachechaddhadaadedeeesheecaca	anicosourgegee	coopeyyyayeayaac oon /o	020 001		
1F1_1_1-20	gaagegetetgtecactaca 57.1 80 1R1_1_285-306	6 cgaaccettcgagtgtetttgcgggagcagaac	60.2 73 315 63.0	80 409	65.0	
3F2_1_1-21	gcatacgtgtaatagcagtgtcgatcccttgatttggagagcacgc	59.9 80 3R1_1_246-271 cagage	cttctccctcagcatagtttaggtatcggg	59.4 80	302	65.0
4F1_1_1-20	cgatacctasactatgctgagggagaagcgctctgtccactaca	57.1 80 4R2_1_284-306 agtgtct	tttgcgggagcagaact 60.4 80	327 63.	0	
1F1_1_1-20	gaagegetetgtecactaca 57.1 80 1R1_1_285-306	6 cgaaccettegagtgtetttgegggageagaac	60.2 73 315 63.0			
2F2_1_1-21	gctcccgcaaagacactcgaagggttcgttaccggctt 60.1	87 2R1_1_361-386 casatcaagggat	cgacactgctattacacgtatgcg 58.6	80 409	65.0	
4F2 1 1-21	gcatacgtgtaatagcagtgtcgatcccttgatttggagagcacgc cgatacctaaactatgctgagggagaagcgctctgtccactacac	58.7 80 4R1 1 285-306 agtgtct	tttgcgggagcagaac 60.2 73	329 64.	302	65.0
			~ ~ ~ ~ ~ ~ ~			
2F2 1 1-21	gaagcgctctgtccactaca 57.1 80 1R1_1_285-306 gctcccgcaaagacactcgaagggttcgttaccggctt 60.1	cgaacccttcgagtgtctttgcgggagcagaac 87 2R1 1 361-386 caaatcaaggcagta	eu.2 /3 315 63.0 cgacactgctattacacgtatgcg 58.6	80 409	65.0	
3F2_1_1-21	gcatacgtgtaatagcagtgtcgatcccttgatttggagagcacgc	59.9 80 3R1_1_246-271 cagage	gettetecetcagcatagtttaggtateggg	59.4 80	302	65.0
4F2_1_1-21	cgatacctaaactatgctgagggagaagcgctctgtccactacac	58.7 80 4R2_1_284-306 agtgtct	tttgcgggagcagaact 60.4 80	328 64.	0	
1F1_1_1-20	gaagegetetgteeactaca 57.1 80 1R1_1_285-306	6 cgaaccettegagtgtetttgegggageagaac	60.2 73 315 63.0			
lama a a ma	gctcccgcaaagacactcgaagggttcgttaccggctt 60.1	87 2R1_1_361-386 casatcaagggat	cgacactgctattacacgtatgcg 58.6	80 409	65.0	
282 1 1-21	GCALACULGLAALAGCAGLGLCGALCCCLLGALLLGGAGAGAGCACGC	55.5 du 3K1_1_240-2/1 Cagages	serececetcadestadittaddigtg		302	66 0
2F2_1_1-21 3F2_1_1-21 4F2_1_1-21	cgatacctasactatgctgaggggagaagcgctctgtccactacac	58.7 80 4R3_1_283-306 agtgtct	tttgcgggagcagaactc 61.4 87	327 64.	0	65.0
2F2_1_1-21 3F2_1_1-21 4F2_1_1-21	cgatacctasactatgctgagggagaagcgctctgtccactacac	58.7 80 4R3_1_283-306 agtgtct	tttgcgggagcagaactc 61.4 87	327 64.	0	65.0
2F2_1_1-21 3F2_1_1-21 4F2_1_1-21 1F1_1_1-20 2F2_1_1-21	cgatacctaaactatgctgagggagaagcgctctgtccactacac gaagcgctctgtccactaca 57.1 80 1R1_1_285-300 gctcccgcaaagacactcgaagggttcgttaccggctt 60.1	58.7 80 4R3_1_283-306 agtgtct 6 cgaaccottogagtgtottttgogggagcagaac 87 2R1 1 361-386 caaatcaannoat	tttgcgggagcagaactc 61.4 87 60.2 73 315 63.0 cgacactgctattacacgtatgcg 58.6	327 64. 80 409	65.0	65.0
2F2_1_1-21 3F2_1_1-21 4F2_1_1-21 1F1_1_1-20 2F2_1_1-21 3F2_1_1-21	cgatacotaaactatgctgagggagaagggottctgtocactacac gaagggottctgtocactaca 57.1 80 1R1_1_285-304 gottocogcaaagacactogagggttcgttacoggott 60.1 gottacogtgtaataggagtocgatcocttgatttggagacgoc	S8.7 80 4R3_1_283-306 agtgtct 6 cgaaccettegagtgtetttgegggageagaac 87 2R1_1_361-386 caaatcaagggat 59.9 80 3R1_1_246-271 cagagcg	tttgcggggggggggggggggggggggggggggggggg	327 64. 80 409 59.4 80	0 65.0 302	65.0
2F2_1_1-21 3F2_1_1-21 4F2_1_1-21 1F1_1_1-20 2F2_1_1-21 3F2_1_1-21 4F3_1_1-22	cystacctasactatyctysgysgaagcyctutytocactacac gaagogototytocactaca 57.1 80 IR1_285-304 gotocogoaagaacactogaaggitogitaccogoot 60.1 gotacogytaatagoagyogaaccottgattoggagacogo cystacctasactatyctyggggaagcyctutytocactacac	58.7 80 4R3_1_223-306 agtgtc1 6 cgaacccttcdgagtgtctttgcgggagcaggac 87 2R1_1_361-386 casatcaaggact 59.9 80 3R1_246-271 cagagot 60.7 73 4R1_1_285-306 agtgtc1	tttgcggggagcagaacte 61.4 87 60.2 73 315 63.0 ggacactgctattacagtatgcg 58.6 gcttctccctcagcatagtttaggtatcggg tttgcggggagcagaac 60.2 73	80 409 59.4 80 330 66.	65.0 302	65.0
2F2_1_1-21 3F2_1_1-21 4F2_1_1-21 1F1_1_1-20 2F2_1_1-21 3F2_1_1-21 4F3_1_1-22 1F1_1_1-20	cystacctasactatyctyspysgyagogctotytocactacac gaagogototytocactaca 57.1 80 1R1_285-304 gotocogoaaggacactogaaggatogttacgtacggot 60.1 gotacgytastagoagytogtacottgatoggagacgot cystacgtastatgotygoggagagogototytocactacac gaagogototytocactaca 57.1 80 1R1_285-304	58.7 80 4R3_1_283-306 agtgtc1 6 opascottcgagtgtcttttgoggagcagaac 87 2R1_361-366 castcaaggat 59.9 80 3R1_1_246-271 cagagcg 60.7 73 4R1_1_285-306 agtgtc1 6 cgaaccottcgagtgtctttgggggagcagaac 6	tttgcgggagcagaactc 61.4 87 60.2 73 315 63.0 cgacactgctattacacgtattgcg 58.6 58.6 cytttdccctagacastgtttagggtatcggg 58.6 60.2 73 60.2 73 315 63.0 53.0	327 64. 80 409 59.4 80 330 66.	65.0 302	65.0 65.0

Fig. 7 An example of polymerase extension PCR for fragment assembly result

a hash index of 7- to 12-mers (allowing up to one mismatch) and 857 by calculating the local similarity for the whole primer. The 858 parameters can be altered to allow different degrees of mismatches 859 at the 3' end of the primers. The parameters for quick alignment 860 may be set: the minimum is 0–5 mismatches (default 2 mismatches) 861 at 3' end of primer. The program can also handle degenerate 862 primers or probes, including those with 5' or 3' tail sequences. It 863 includes the detection of non-Watson-Crick base-pairing in in 864 silico PCR, e.g., the stable guanine mismatches $G \cdot G$, $G \cdot T$, and 865 $G \cdot A$. Probable PCR products can be found for both linear and 866 circular templates in both standard and inverse PCR, as well as in 867 multiplex PCR and using bisulphite-treated DNA. This in silico 868 tool is useful for quickly analyzing primers or probes against target 869

Author's Proof

Ruslan Kalendar et al.

sequences, for determining primer location, orientation, and efficiency of binding, and for calculating primer T_m and annealing temperature in PCR.

The user must input a preexisting primer list into a second Additional sequence(s) or pre-designed primers (probes) list text editor. The number of preexisting primers is not limited; it can be as many as the user needs. The target sequences can be entered either as multiple separate DNA sequences or by opening files from the selected folders. For in silico PCR against whole genome(s) or a list of chromosomes, the user must specify the directory containing the input. The program will be consistent: it will look at each file to find the position of the primers. The user can execute the search task with F5 on the in silico PCR tab or can specify search options including stringency and PCR product detection settings. For the stringency options, users can specify the number of mismatches that the primers are allowed at 3' terminus. The default specificity settings allow a maximum two mismatches within the 3' end region of the primers. These mismatches within the 3' end of the primers should not be located close to each other. Once the primer set design is complete, the results will appear in the text editors In silico PCR Result.

In silico **PCR Result** text editor reports the specificity of the primers (locations, including target position, similarity, and T_m), a summary of primer pairs in relation to the PCR template, and detailed information on each primer pair, including its length and T_a . It will show the target-specific primers that have been found. The actual targets will be listed along with detailed alignments between primers and targets (Fig. 8).

898 6 Primer Analyses

Individual and sets of primers are evaluated using FastPCR or the online software. They calculate primer $T_{\rm m}s$ using default or other formulae for features of the primers including normal and degenerate nucleotide combinations, CG content, extinction coefficient, unit conversion (nmol per OD), mass (µg per OD), molecular weight, and linguistic complexity and consider primer PCR efficiency. Users can select either DNA or RNA primers (online: PrimerAnalyser, http://primerdigital.com/tools/ PrimerAnalyser.html) with normal or degenerate oligonucleotides or modifications with various labels (for example, inosine, uridine, or fluorescent dyes). Tools allow the choice of other nearest neighbor thermodynamic parameters or non-thermodynamic $T_{\rm m}$ calculation formulae.

For LNA modifications the four symbols, dA=E, dC=F, dG=J, and dT=L, are used. Both programs perform analyses on-



PCR Primer and Probe Design and Oligonucleotide Assembly and Analysis

PCR Primer Design in silico PCR Primer Test Primer	s List Analysis RAPD PCR Restriction Clustering Searching LTR Search MITE Search SSR Search Tools Oligos Assembly
Maximal PCR product size (bp)	Mismatches allowed in 3-end 0.5 nt
PCR product prediction	Prope search
Circular sequence	Pattern search
Restrict analysis to F/R primer pairs	Show matches sequences
I C >> T bisulphite conversion	
Sequences: 1 : 1139 Pre-designed primers list 2 : 43 In	n silico PCR Result
>1 atatococtocgtocotaatatttgacgoogttgactt actatagttttacatattocoaaagtttttgatag tacotoGCTTGTCCTCAAGCANHANNAtocoaaacat gggaaggtacaattocttcagaagcatgaagaa aaccacataatcatcattocttttgutttocacaatco tacotaggggggggtggggttgagttocaattoc tacottagggggggtggggttgagttocaattga aggegttctgactgococatcottgoggcocttaga aggagggacggaggaaacatcatct	III taaaacatattigacogticgtottattoaaaaaaattaagtaattattaattoottiootatoattigattoattgtaaatatactittagtat acgaacggtcaaacatgttgacgaaggttaaagagggggggg
Sequences: 1: 1139 Pre-designed primers list: 2: 43	n silico PCR Result
In silico Primer(s) search for: 1 1 5'-gettgteeteaagegaaaassa	
Position: 251->272 22bp 88% Tm = 51.2°C	
5-gcttgtcctcaagcgaaaassa-> ::::: tcgcttgtcctcaagcgarrrnnaagtg	
1 5'-gcttgtcctcaagcgaaaassa	
Position: 285->306 22bp 86% Tm = 51.1°C	
5-gcttgtcctcaagcgaaaassa-> :::::: tcgcttgtcctcaagcgawrwnnratcc	
2 5'-cgcagcgttctcataaggtcr	
Position: 1074<-1094 21bp 95% Tm = 55.4°C	
<-retggaatactettgegaege-5 ::!!!!!! egssacettatgagaaegetgegae	
1 5'-gcttgtoctcaagogaaaassa 2 5'-ogoagogttotoataaggtor FCR product size: 844bp Ta=66°C	
1 5'-gottgtootoaagogaaaassa 2 5'-ogoagogttotoataaggtor FCR product size: 810bp Taw66°C	
A=357.5 T=323.0 G=213.5 C	2=245 0 N=13 0 R=362 0 Y=568 0 RY=0 637 CG=40 3% Tm=8 Open FASTA Sequences with a >' symbol

Fig. 8 An example of in silico PCR result

type, allowing users to see the results immediately on screen. They 914 can also calculate the volume of solvent required to attain a specific 915 concentration from the known mass (mg), OD, or moles of 916 oligonucleotide. 917

All primers are analyzed for intra- and inter-primer interactions 918 regarding formation of dimers. Primer(s) can efficiently hybridize 919 using the 5' end or the middle of the oligonucleotides. Even 920 though such interactions are not efficiently extended by DNA 921 polymerase, their formation reduces the effective primer concen-922 tration available for binding to the targets and their presence can 923 strongly inhibit PCR because double-stranded DNA at high con-924 centrations is a strong inhibitor of DNA polymerase (Fig. 9). 925



Ruslan Kalendar et al.

PCR Primer Design in silico PCR Primer Test Primers List Analysis F	Restriction Clustering	Searching	LTR Search	MITE Search	SSR Search	Tools	Oligos Assembly
Type or paste (Ctrl-V) sequence of oligo with universal degenerate DNA co	de (5'-3'):						
catagcatggataataaaciattatc							
Tm of oligo calculation parameters:	Amount of olig	0:					
K+ (or Na+, NH4+, Tris+) total concentration (mM): 50	OD (A260):	1	Target	concentration (j.	IM): 100		
Mg2+ concentration (mM): 0	Mass (µg):						
Oligo concentration (µM): 0.25	nmol:						
General Sequence(s) Additional sequence(s) or pre-designed primers (pr	obes) list Oligo analysis						
<pre>S'-catagcatagcatataaciattatc-3' Length=26 A=11.0 G=3.0 T=7.0 C=4.0 CG=26.9% Linguistic complexity = 82% Primer's PCR efficiency = 25% dG = -25.0 kcal/mol dH = -184.0 kcal/mol dS = -512 Tm = 48.7°C (Allawi's thermodynamics parameters (Bio Tm = 52.6°C (Tm = 77.1 + 11.7Log[K+] + (41(G + C) - 1 Tm = 50.1°C (Tm = 64.9 + 41(G + C - 16.4)/L) Extinction coefficient = 238000 L/(mol·cm) Molecular weight = 7971 g/mol 00260 = 1.000 µg = 33.492 nmol = 4.202</pre>	.5 cal/K mol chemistry,1997, 34 528)/L)	6:10581-	.10594)				
100µM = dissolve in 42.0 µl of MQ-water or TE buffer							
Dimer:							
<-ctattaicaaataataggtacgatac-5 S-catagcatggataataaaciattatc-> Tm=22.5°C							
1		_					

Fig. 9 Example result of the oligonucleotide analysis

₉₂₆ 7 Availabilit	V
------------------------------	---

927	The FastPCR software is available for download at http://
928	primerdigital.com/fastpcr.html; the online version is available at
929	http://primerdigital.com/tools/pcr.html. The program manual,
930	licence agreement, and installation files can be found at http://
931	primerdigital.com/fastpcr/. YouTube tutorial videos have been
932	placed at http://www.youtube.com/user/primerdigital. Web tools
933	are accessible at http://primerdigital.com/tools/.

934 Acknowledgments

935Web tools are available free to academic institutions, provided that936they are used for noncommercial research and education only.937They may not be reproduced or distributed for commercial use.938This work was partially supported by the companies PrimerDigital939Ltd. and Oligomer Ltd. and by the Academy of Finland, Project940134079.

References

- 942 1. Untergasser A et al (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115. doi:10.1093/nar/gks596
- 945
 946
 946
 947
 948
 948
 949
 949
 940
 941
 941
 942
 943
 944
 945
 945
 946
 946
 947
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 949
 949
 949
 940
 940
 941
 941
 942
 942
 943
 944
 944
 945
 945
 946
 946
 947
 947
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 948
 949
 948
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
 949
- 849
 949
 950
 950
 951
 952
 952
 952
 953
 954
 954
 955
 955
 955
 956
 957
 957
 957
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
 958
- 953
 954
 954
 955
 956
 956
 957
 958
 959
 959
 959
 950
 950
 950
 950
 951
 951
 952
 952
 952
 953
 954
 954
 955
 956
 956
 956
 957
 958
 958
 958
 958
 959
 959
 959
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
 950
- 957 5. Owczarzy R et al (2008) IDT SciTools: a suite
 958 for analysis and design of nucleic acid oligo959 mers. Nucleic Acids Res 36:W163–9.
 960 doi:10.1093/nar/gkn198
- 961 6. Bekaert M, Teeling EC (2008) UniPrime: a
 962 workflow-based platform for improved univer963 sal primer design. Nucleic Acids Res 36:e56.
 964 doi:10.1093/nar/gkn191
- 965 7. Ye J et al (2012) Primer-BLAST: a tool to
 966 design target-specific primers for polymerase
 967 chain reaction. BMC Bioinformatics 13:134.
 968 doi:10.1186/1471-2105-13-134
- 8. Giegerich R, Meyer F, Schleiermacher C
 (1996) GeneFisher—software support for the
 detection of postulated genes. Proc Int Conf
 Intell Syst Mol Biol 4:68–77
- 973
 97. Gadberry MD et al (2005) Primaclade—a flexible tool to find conserved PCR primers across multiple species. Bioinformatics 21:1263– 1264. doi:10.1093/bioinformatics/btil34
- 977 10. National Center for Biotechnology
 978 Information, National Library of Medicine,
 979 Building 38A, Bethesda, MD, 20894. http://
 980 blast.ncbi.nlm.nih.gov/blastcgihelp.shtml
- 11. Nomenclature for incompletely specified bases
 in nucleic acid sequences (1984) http://www.
 chem.qmul.ac.uk/iubmb/misc/naseq.html
- 12. Allawi HT, SantaLucia J Jr (1997)
 Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry 36:10581–
 10594. doi:10.1021/bi962590c
- 13. SantaLucia J (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci USA 95:1460–1465
- 14. Le Novere N (2001) MELTING, computing
 the melting temperature of nucleic acid duplex.
 Bioinformatics 17:1226–1227. doi:10.1093/
 bioinformatics/17.12.1226

- 15. Bolton ET, McCarthy BJ (1962) A general method for the isolation of RNA complementary to DNA. Proc Natl Acad Sci USA 48: 1390–1397
 999
- 16. Guedin A et al (2010) How long is too long? 1000 Effects of loop size on G-quadruplex stability. 1001 Nucleic Acids Res 38:7858–68. doi:10.1093/ 1002 nar/gkq639 1003
- 17. Wallace RB et al (1979) Hybridization of synthetic oligodeoxyribonucleotides to ΦX 174 1005
 DNA: the effect of single base pair mismatch. Nucleic Acids Res 6:3543–57. doi:10.1093/ 1007
 nar/6.11.3543 1008
- 18. von Ahsen N, Wittwer CT, Schutz E (2001) 1009
 Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections 1011
 for Mg²⁺, deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. Clin 1014
 Chem 47:1956–1961
 1009
- 19. Kypr J et al (2009) Circular dichroism and 1016 conformational polymorphism of DNA. 1017 Nucleic Acids Res 37:1713–25. doi:10.1093/ 1018 nar/gkp026
- 20. Gabrielian A, Bolshoy A (1999) Sequence complexity and DNA curvature. Comput Chem 23:263–74. doi:10.1016/S0097-8485 (99)00007-8
- 21. Orlov YL, Potapov VN (2004) Complexity: an
internet resource for analysis of DNA sequence
complexity. Nucleic Acids Res 32:W628–33.
doi:10.1093/nar/gkh4661027
- 22. Gilson MK et al (1997) The statisticalthermodynamic basis for computation of 1029 binding affinities: a critical review. Biophys J 1030 72:1047–69. doi:10.1016/S0006-3495(97) 1031 78756-3 1032
- Peyret N et al (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. doi:10.1021/bi9825091
- 24. Watkins NE Jr, SantaLucia J Jr (2005) Nearestneighbor thermodynamics of deoxyinosine pairs in DNA duplexes. Nucleic Acids Res 33:6258–67. doi:10.1093/nar/gki918
 1041
- 25. Sen D, Gilbert W (1992) Guanine quartet 1042 structures. Methods Enzymol 211:191–199 1043
- Il'icheva IA, Florent'ev VL (1992) Fourstranded complexes of oligonucleotidesquadruplexes. Mol Biol (Mosk) 26:512–531 1046
- 27. Shing HP (1994) The non-B-DNA structure of 1047 d(CA/TG)n does not differ from that of Z-DNA. 1048 Proc Natl Acad Sci USA 91:9549–9553 1049

941

1020

1021

1022

1023

1033

1034

1035

1036

1037

Ruslan Kalendar et al.

- 28. SantaLucia J Jr, Hicks D (2004) The thermodynamics of DNA structural motifs. Annu Rev Biophys Biomol Struct 33:415–40. doi:10.1146/ annurev.biophys.32.110601.141800
- 1054 29. Williams JGK et al (1990) DNA polymor1055 phisms amplified by arbitrary primers are use1056 ful as genetic-markers. Nucleic Acids Res
 1057 18:6531–5. doi:10.1093/nar/18.22.6531
- 30. Welsh J, Mcclelland M (1990) Fingerprinting
 genomes using PCR with arbitrary primers.
 Nucleic Acids Res 18:7213–8. doi:10.1093/
 nar/18.24.7213
- 1062 31. Kalendar R, Schulman A (2006) IRAP and
 1063 REMAP for retrotransposon-based genotyp1064 ing and fingerprinting. Nat Protoc 1:2478–84.
 1065 doi:10.1038/nprot.2006.377
- 1066 32. Chang RY, O'Donoughue LS, Bureau TE
 1067 (2001) Inter-MITE polymorphisms (IMP): a
 1068 high throughput transposon-based genome
 1069 mapping and fingerprinting approach. Theor
 1070 Appl Genet 102:773–781
- 1071 33. Nelson DL et al (1989) Alu polymerase chain
 1072 reaction: a method for rapid isolation of
 1073 human-specific sequences from complex
 1074 DNA sources. Proc Natl Acad Sci USA 86:
 1075 6686–6690
- 1076 34. Sinnett D et al (1990) Alumorphs—human
 1077 DNA polymorphisms detected by polymerase
 1078 chain reaction using Alu-specific primers.
 1079 Genomics 7:331–334
- 35. Jurka J et al (2005) Repbase update, a database of
 eukaryotic repetitive elements. Cytogenet Genom
 Res 110:462–7. doi:10.1159/000084979

- 36. TREP, the Triticeae Repeat Sequence Database 1083 (2008) http://wheat.pw.usda.gov/ITMI/ 1084 Repeats/ 1085
- 37. Landegren U et al (1988) A ligase-mediated gene1086detection technique. Science 241:1077–10801087
- 38. Higasa K, Hayashi K (2002) Ordered catenation of sequence-tagged sites and multiplexed SNP genotyping by sequencing. Nucleic Acids Res 30:E11
 1088
 1089
 1090
 1091
- 39. Quan J, Tian J (2009) Circular polymerase 1092 extension cloning of complex gene libraries 1093 and pathways. PLoS One 4:e6441. 1094 doi:10.1371/journal.pone.0006441 1095
- 40. Cao YF et al (2005) Information theory-based
algorithm for *in silico* prediction of PCR prod-
ucts with whole genomic sequences as tem-
plates. BMC Bioinformatics 6:190. doi:10.1186/
1471-2105-6-1901096
1097
1100
- 41. Rubin E, Levy AA (1996) A mathematical 1101 model and a computerized simulation of PCR 1102 using complex templates. Nucleic Acids Res 24:3538–45. doi:10.1093/nar/24.18.3538 1104
- 42. Lexa M, Valle G (2003) PRIMEX: rapid identification of oligonucleotide matches in whole genomes. Bioinformatics 19:2486–2488
 1107
- 43. Nishigaki K et al (2000) Whole genome1108sequence-enabled prediction of sequences per-
formed for random PCR products of *Escherichia*1109*coli*. Nucleic Acids Res 28:1879–18841111
- 44. Rotmistrovsky K, Jang W, Schuler GD (2004) 1112
 A web server for performing electronic PCR. Nucleic Acids Res 32:W108–12. doi:10.1093/ 1114
 nar/gkh450 1115



Author Queries

Chapter No.: 18 0002067381

Queries	Details Required	Author's Response
AU1	Please check if the edit made in the sentence "It calculates the melting resuspension calculator." is appropriate.	
AU2	Please check if the edit made in the sentence "Getting started with a to type." is appropriate.	
AU3	Please check if the edit made in the sentence "Primer pairs are analyzed melting temperatures." is appropriate.	
AU4	Please check if the edit made in the sentence "Multiplex PCR can be of both." is appropriate.	
AU5	Please provide closing parenthesis in sentence "Design primers with a specific restric- tion"	
AU6	Please check if the edit made in the sentence "The program takes multiple align- ment formats." is appropriate.	
	unonecie	