In Silico Tools for qPCR Assay Design and Data Analysis

ARTICLE *in* METHODS IN MOLECULAR BIOLOGY (CLIFTON, N.J.) · JANUARY 2011

Impact Factor: 1.29 · DOI: 10.1007/978-1-61779-176-5_18 · Source: PubMed

3 AUTHORS, INCLUDING:

Stephen Andrew Bustin
Anglia Ruskin University

166 PUBLICATIONS **16,719** CITATIONS

SEE PROFILE

Chapter 18

In Silico Tools for qPCR Assay Design and Data Analysis

Stephen Bustin, Anders Bergkvist, and Tania Nolan

Abstract

qPCR instruments are supplied with basic software packages that enable the measurement of fluorescent changes, calculations of quantification cycle (C_q) values, the generation of standard curves and subsequent relative target nucleic acid quantity determination. However, detailed assessments of the technical parameters underlying C_q values and their translation into biological meaningful results require validation of these basic calculations through further analyses such as qPCR efficiency correction, normalization to multiple reference genes, averaging and statistical tests. Some instruments incorporate some of these features, while others offer additional tools to complement the basic running software, in many cases providing those that are described below. In this chapter, there is a detailed description of some of these programs and recommended strategies for the design of robust qPCR assays. Some of the packages available for validation of the resulting C_q data and detailed statistical analysis are described.

Key words: Assay design, real-time PCR, RT-qPCR, PCR efficiency, normalization.

1. Introduction

The broad division of real-time quantitative PCR (qPCR) assays into those that are used within life science research and those used for diagnostic purposes entails two fundamentally different approaches to qPCR assay design (*see* **Note 1**). Research projects tend to be relatively low throughput, but with a requirement for great flexibility with respect to experimental and assay designs. Furthermore, individual researchers have highly variable requirements and coupled with their independence and creativity, this means that they use a wide variety of design and analysis tools, as well as optimization and validation protocols for qPCR assay design. In general, therefore, assays are designed using either

B. Yu, M. Hinchcliffe (eds.), *In Silico Tools for Gene Discovery*, Methods in Molecular Biology 760,
DOI 10.1007/978-1-61779-176-5_18,

no computer assistance or relatively low-powered assay design software, with data analysis software remaining optional and most statistical analysis being performed using tools available in spreadsheets such as Microsoft Excel or Apple's Numbers. Those involved in clinical diagnostics, on the other hand, are increasingly making use of the growing public availability of pathogen genome sequences to implement qPCR assays that rely on sequence-based pathogen identification. Assays must be unique to the pathogen with respect to all other non-target genomes; hence the design of pathogen diagnostic assays is a high throughput activity that involves the computationally expensive comparison of multiple target genomes with all known non-target sequences. Furthermore, accurate, reliable and robust data analysis is essential, resulting in data analysis software with clear process tracking function being an absolute necessity.

This diversity of practice, coupled with the flexibility of qPCR itself, means that numerous in silico tools have been developed to guide the design of qPCR assays and analyse any resulting quantitative data. Many tools are freely available online, while others are bundled with qPCR instruments or are available from various software houses. The most comprehensive information for accessing and evaluating these programs is available at http://www.gene-quantification.de/main-bioinf.shtml, along with additional links to detailed reviews and other publications dealing with qPCR data analysis. With so many different tools available, it is impossible to overview them all and do them justice. Hence we have organized this chapter to present an example of assay design to demonstrate the steps involved at every stage of the process, utilizing the tools we are familiar with and consider to be the best available. This does not mean, of course, that other tools are not equally capable of performing the task.

The MIQE guidelines (1, 2) provide clear guidance on the steps that are important for assay design, and we shall follow these guidelines to design this example assay. The selected target is the human vitamin D receptor (VDR) mRNA. VDR mediates the effects of 1,25-dihydroxyvitamin D3 [1,25-(OH)2D3] and is specified by the *VDR* gene; transcription is directed by distinct promoters that can generate unique transcripts with major N-terminal differences (3).

2. Materials

A computer, access to the Internet and ideally assay design software such as Beacon Designer and data analysis software such as GenEx or qbasePLUS are required for qPCR design and data

analysis. However, as described, there are free online alternatives to these programs. It is also useful to have access to a spreadsheet program such as Microsoft Excel or Apple's Numbers.

3. Methods

3.1. qPCR Target Information

A search of any sequence repository for a specific nucleic acid target often reveals a number of variants, making it challenging to deduce which particular pathogen, DNA or RNA has been targeted. Providing an accession number is the most basic, yet frequently omitted piece of information. Common starting points for qPCR target searches are the National Center for Biotechnology Information (NCBI) nucleotide search site (http://www.ncbi.nlm.nih.gov/sites/nuccore) and gene search (http://www.ncbi.nlm.nih.gov/gene) websites, where a search for "human VDR" provides a link to VDR (*Homo sapiens*): vitamin D (1,25-dihydroxyvitamin D3) receptor, with the unique GeneID 7421 and official symbol VDR, both of which are essential MIQE requirements. Following this link opens up a page providing detailed, summarized information on this gene, including a link to Ensembl's genome browser (*see* below). A section labelled "Genomic regions, transcripts, and products" provides a link to the reference sequence with the all-important accession numbers, which confirms that there is more than one transcript. Variant 1 lacks an alternate exon in the 5′-UTR when compared with variant 2; both variants specify the same protein. Following the links to GenBank reveals two reference mRNA sequences, NM_0000376.2 (variant 1) and NM_001017535.1 (variant 2). The requirements of the experiment are crucial considerations at this stage. In some cases, there is a requirement to detect all sequences, regardless of splicing, and so the conserved region is selected as the target. Alternatively, as in this case, the requirement is to distinguish between sequences and so individual assays are designed to target each of the variants. In this case we will design two assays, one targeting each of the two variants.

Both mRNA sequences are imported into the CLC Sequence Viewer (http://www.clcbio.com), a free software package developed for Windows, Mac and Linux platforms that allows basic nucleic acid and protein analysis. An alternative would be to use ClustalW, one of the free tools provided by the European Bioinformatics Institute (EBI) (http://www.ebi.ac.uk/Tools/clustalw2/index.html). The sequences are aligned to reveal the differences at the 5′-end of the mRNA, with the rest of the sequences being identical (**Fig. 18.1**). An analysis of the alignment shows that the first 76 nucleotides from the 5′-end are

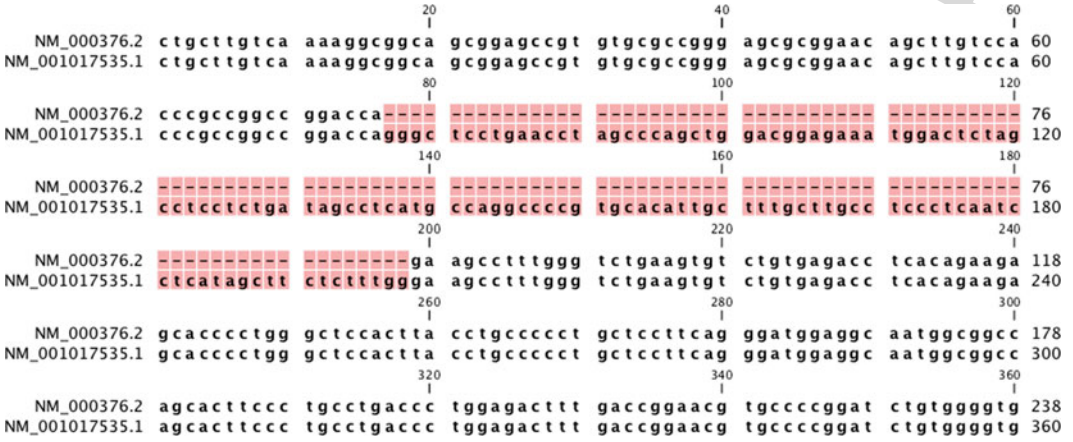


Fig. 18.1. Alignment of the 5′-ends of the two reference sequences specifying the VDR. Two blocks of identical sequences surround the 122-bp sequence (*highlighted*) unique to variant 2. A GC-rich sequence (nucleotides 70–72) precedes the splice junction at position 75.

This figure will be printed in b/w

identical, with variant 2 having a 122-bp insert. This is the critical region of differentiation that provides useful anchors for the variant-specific assay design. An upstream primer centred on nucleotide 76 should allow the generation of variant 1-specific amplicons, whereas an upstream primer located within the variant 2-specific sequence should result in a variant 2-specific amplicons. In either case, the primers will be intron spanning, minimizing any problems associated with genomic DNA (gDNA) contamination.

Ensembl (http://www.ensembl.org/index.html) and the UCSC Genome Bioinformatics Site (http://genome.ucsc.edu/) provide alternative means of acquiring and handling sequence information. Which toolkit to use is down to personal preference, but in our experience, the relative simplicity of NCBI's sites makes them the tools of choice for simple sequence acquisition and browsing of information that is critical for high-quality qPCR assay designs.

3.2. Primer/Probe Design

In theory it should be easy to utilize tens of thousands of published qPCR assays to obtain suitable primer and probe sequences. The Quantitative PCR Primer Database (QPPD) (http://web.ncifcrf.gov/rtp/gel/primerdb/) provides an assembly of assay details from published articles, and in compliance with the MIQE guidelines, it provides information on primer location, amplicon size, assay type, positions of single-nucleotide polymorphisms (SNPs) as well as literature references. However, these assays have not been independently optimized or validated and so may not generate reliable quantitative data. RTPrimerDB (http://medgen.ugent.be/rtprimerdb/) contains more than 8,000 qPCR assays for over 5,500 genes. When an assay is submitted, there is

a requirement for appropriate validation and optimization data, with the option for further users to leave feedback (4–6). The use of such properly validated assays has numerous advantages, obviating the requirement for time-consuming primer design and empirical optimization and critically introducing more uniformity and standardization among different laboratories. RTPrimerDB is particularly flexible and can be queried using the official gene name or symbol, Entrez or Ensembl Gene identifier, SNP identifier or oligonucleotide sequence. Queries can be restricted to a particular application, e.g. mRNA quantification, gDNA copy number quantification, SNP detection, organism or detection chemistry.

However, these databases cannot accommodate every gene or variant; new variants may have been identified, or suitable assays may never have been designed for the organism of interest. Consequently, there are a large number of conditions where a proficiency in good assay design is indispensable. Happily, there are numerous options available if an assay needs to be designed from scratch; less propitiously, the many choices available can be confusing. Several software tools provide comprehensive facilities for designing primers and probes for standard and bisulphite modification qPCR assays, as well as multiplex designs that permit the design of multiple primers simultaneously. Many allow detailed control over numerous features, including assay location, primer and amplicon sizes, probe modifications, annealing temperatures, ionic conditions, GC content and dimer formation. Some can design assays if the sequences of only one primer or probe are provided. We have found two programs to be consistently useful: (i) Premier Biosoft Beacon Designer, which can be used to design single template as well as multiplex assays and supports all common chemistries such as SYBR Green I, standard hydrolysis as well as locked nucleic acid (LNA)-modified probes, hybridization probes, Molecular Beacons and Scorpions. Beacon Designer can predict CpG islands, or CpG regions can be delineated manually, thus permitting the design of MethyLight hydrolysis probes and methyl-sensitive PCR primers, together with suitable control primers and probes for unmethylated and untreated DNA sequences. Beacon Designer is particularly useful for designing high-resolution melting analysis assays, as it uses proprietary algorithms for designing the best primers and the shortest possible amplicons (*see* **Note 2**). (ii) Premier Biosoft AlleleID supports hydrolysis minor grove binding (MGB) probe design and can design assays across exon junctions. Its particular power applies to the design of qPCR-based diagnostic assays for pathogen detection. For cross-species assays, AlleleID identifies the conserved regions to design a universal probe. When it is not possible to identify a significant conserved region for a set of sequences, AlleleID helps design a "mismatch tolerant" probe. Also included

is a "minimal set" option that helps design a minimum number of probe sets that uniquely identify a sequence reducing the overall assay cost. This functionality is also useful to help study gene expression when genomic sequences of the organism under study are not available.

The power of both Beacon Designer and AlleleID is in the customization both programs permit the user. Details are outside the scope of this chapter, but both contain the facility to check primer specificity by avoiding significant cross-homologies identified by automatically interpreting BLAST search results, maximize annealing efficiency by avoiding template structures identified by in silico folding prediction of the template (using mfold, *see* below), can design up to five multiplex hydrolysis probe assays by checking for cross-homologies with all probes and primers to prevent competition, allow the evaluation of predesigned primers and probes and can design wild-type and mutant probes for SNP identification. Previously designed assays can be evaluated and AlleleID can be used to align multiple sequences and select target regions for pathogen detection assay design.

If there is no need for the additional specificity inherent in a probe-based assay, the use of double-stranded DNA-binding dyes such as SYBR Green is recommended. This has several advantages: (i) costs should be lower as there is no need to order a probe; (ii) there are fewer constraints on the location of the primers and consequently (iii) assay design is easier and may result in a more efficient and sensitive assay. An SYBR Green assay for the VDR would be designed so that amplification of the two variants results in different amplicon lengths which could be easily distinguished through the analysis of the resulting melt curves. Quantification relative to variant-specific dilution curves would then permit accurate quantification of the two variants relative to each other. However, SYBR Green is not an option for diagnostic assays, hence there will be numerous occasions when probe-based assays represent the only recourse and this can mean having to design suboptimal assays that are likely to require extensive empirical optimization without any guarantee of success. Nevertheless, while suboptimal assays may not be accurately quantitative, they will always be qualitative, which often is sufficient in a clinical context.

The difficulty with designing a primer set specific for a probe-based assay detecting VDR variant 1 is that the sequence around the unique splice site is very GC rich (**Fig. 18.1**). Since the forward primer must span the splice junction, the Beacon Designer program feature permitting the use of a predesigned primer is used, allowing Beacon Designer to specify a downstream primer and a probe. The only additional stipulation is a limit on amplicon size of 70–100 bp. The program output is shown in **Fig. 18.2**. The program suggests five alternative probes, and a sequence

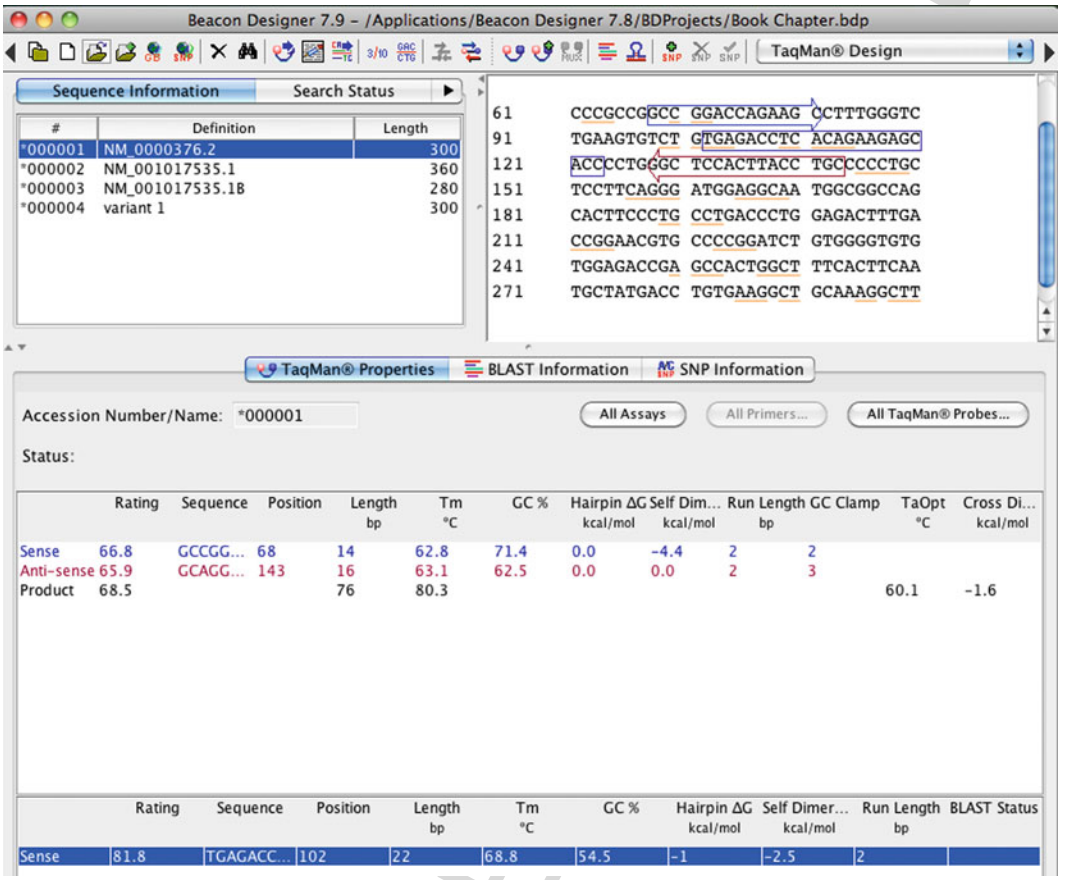


Fig. 18.2. Beacon Designer output following a design for an assay that is variant 1 specific and used a predetermined upstream primer spanning the splice junction.

This figure will be printed in b/w

analysis reveals one possible sense primer dimer (unavoidable), no antisense primer dimer, no primer hairpins, two weak cross-dimers, no runs of greater than two and no repeats (**Fig. 18.3**). Primers and probe are categorized as good and best, respectively. In principle, therefore, this should be an acceptable assay for the specific detection of variant 1. Testing an additional upstream primer, obtained by shifting the binding site by +1 nucleotide, gives the sense primer a slightly higher rating, resulting in a better primer pair rating. Care must be taken with the interpretation of the ratings. These are a measure of how well the assay fits to the criteria defined by the user, hence if a huge tolerance of conditions is permitted, the assay will be rated higher than if a narrower tolerance is permitted. For empirical testing, it is best to test more than one primer set per amplicon, hence both sense primers will be tested against the reverse primer. Primers, probe and their characteristics are shown in **Fig. 18.3**. Alternatively, it would have been possible to designate the reverse primer as the variant-specific primer, with the forward primer in the common

Variant 1 Sense Primer A	Length	Tm	GC%	Rating	Position	Pair Quality	Pair Rating	Product Length	Product Tm	Product TaOpt
GCCGGACCAGAAGC	14	62.8	71.4	66.8	68	Good	68.5	76	80.3	60.1
Anti-sense Primer A	**Length**	**Tm**	**GC%**	**Rating**	**Position**					
GCAGGTAAGTGGAGCC	16	63.1	62.5	65.9	143					
Sense Primer B	**Length**	**Tm**	**GC%**	**Rating**	**Position**	**Pair Quality**	**Pair Rating**	**Product Length**	**Product Tm**	**Product TaOpt**
CCGGACCAGAAGCC	14	61.9	71.4	73.5	69	Best	71.4	75	80	59
Anti-sense Primer B	**Length**	**Tm**	**GC%**	**Rating**	**Position**					
GCAGGTAAGTGGAGC	15	59.7	62.5	78.8	143					

Probe Sequence	Length	Tm	GC%	Rating	Position	Quality
TGAGACCTCACAGAAGAGCACC	22	68.8	54.5	81.8	102	Best

Amplicon A

GCCGGACCAGAAGCCTTTGGGTCTGAAGTGTCTGTGAGACCTCACAGAAGAGCACCCCTGGGCTCCACTTACCTGC

Amplicon B

CCGGACCAGAAGCCTTTGGGTCTGAAGTGTCTGTGAGACCTCACAGAAGAGCACCCCTGGGCTCCACTTACCTGC

Fig. 18.3. Amplicon characteristics for variant 1-specific assay (Beacon Designer). Primer length and ΔT_m between primers and probe are not ideal, but governed by the absolute need to place the forward primers across the splice junction. If the resulting assay is not optimal, it is important to note this during the analysis step and, for example, use the data in a qualitative rather than a quantitative way.

5′-region. Designing a splice junction-specific probe with general primers is another possibility.

Variant 2 is a more straightforward design, as there are 122 nucleotides from the additional exon that can be used for a variant-specific assay. Again an additional upstream primer shifted to +1 is included for the empirical optimization and the primers and probe are shown in **Fig. 18.4**.

Variant 2 Sense Primer A	Length	Tm	GC%	Rating	Position	Pair Quality	Pair Rating	Product Length	Product Tm	Product TaOpt
CACATTGCTTTGCTTGCC	18	62.2	50	73.2	153	Best	80	75	76.1	56.8
Anti-sense Primer A	**Length**	**Tm**	**GC%**	**Rating**	**Position**					
CTCACAGACACTTCAGACC	18	61.2	52.6	84.9	227					
Sense Primer B	**Length**	**Tm**	**GC%**	**Rating**	**Position**	**Pair Quality**	**Pair Rating**	**Product Length**	**Product Tm**	**Product TaOpt**
ACATTGCTTTGCTTGCCT	18	62.4	44.4	72.5	154	Best	78.7	90	77.4	58
Anti-sense Primer B	**Length**	**Tm**	**GC%**	**Rating**	**Position**					
TGCTCTTCTGTGAGGTCTC	19	62.6	52.6	83.8	243					

Probe Sequence	Length	Tm	GC%	Rating	Position	Quality
TTCTCTTTGGGAAGCCTTT	19	70.2	42.1	62.8	189	Good

Amplicon A

CACATTGCTTTGCTTGCCTCCCTCAATCCTCATAGCTTCTCTTTGGGAAGCCTTTGGGTCTGAAGTGTCTGTGAG

Amplicon B

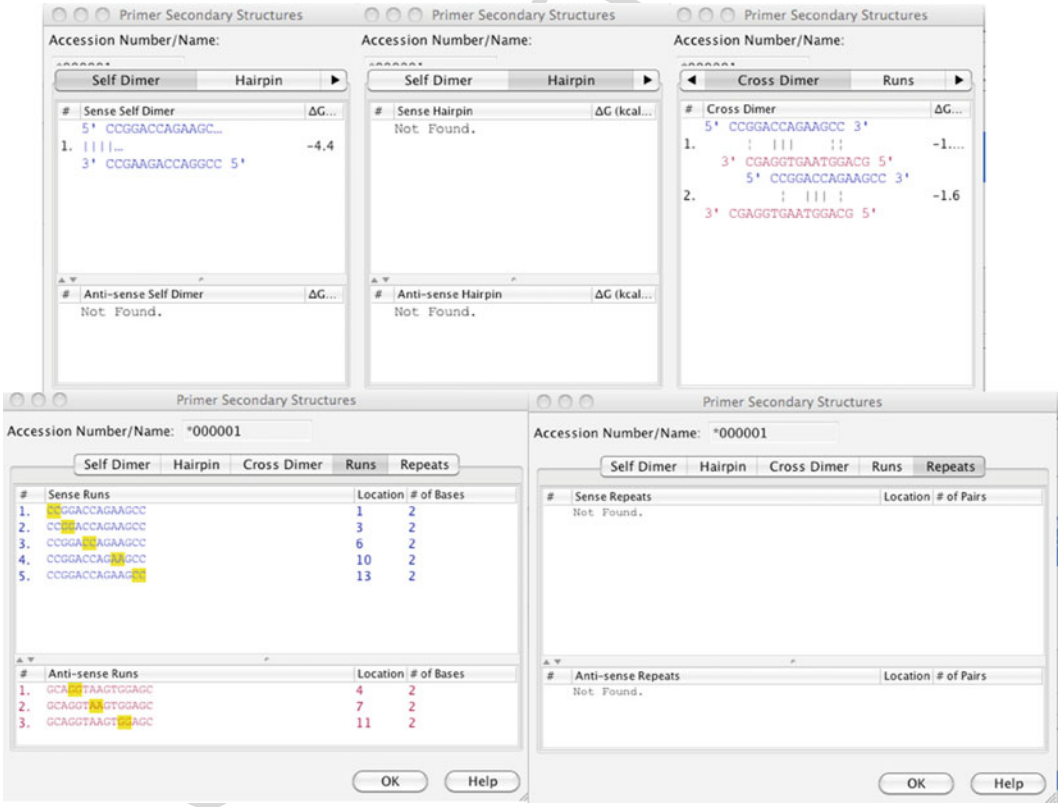
ACATTGCTTTGCTTGCCTCCCTCAATCCTCATAGCTTCTCTTTGGGAAGCCTTTGGGTCTGAAGTGTCTGTGAGACCTCACAGAAGAGCA

Fig. 18.4. Amplicon characteristics for variant 2-specific assay. The two primer sets described here are just two of many possible combinations. Alternatively it would have been possible to design exon-spanning primers or have the antisense primer in the common upstream sequence. There are no hard and fast rules and importantly, what gets the highest ratings in silico must be tested empirically as primer sequences that look very unpromising on paper can perform well and conversely, those that should result in optimal assays may not perform adequately.

Since it may be desirable to detect both transcripts simultaneously and determine the relative quantities of each, it would be useful to design a multiplex assay. The problem here is that the constraint of having to place the forward primer for variant 1 places severe limitations on assay design. Hence one option is to design a set of primers that detect both variants, with the second set being variant 2 specific. This provides no direct information with respect to variant 1 but allows some inference from a comparison of variant 2 and total mRNA copy numbers. This kind of assay is easily designed in BD, as shown in **Figs.** 18.5 and 18.6.

AlleleID generates a slightly different assay for variant 1, with the same probe but a different reverse primer (**Fig.** 18.7). In order to maximize the chances of obtaining an optimal assay, this sequence should also be tested empirically.

For anyone who would like the assurance of a well-designed assay but prefers not to tackle these programs, there are freely available design service options. Sigma Genosys offers a free consultancy-based option; scientists submit their sequence and



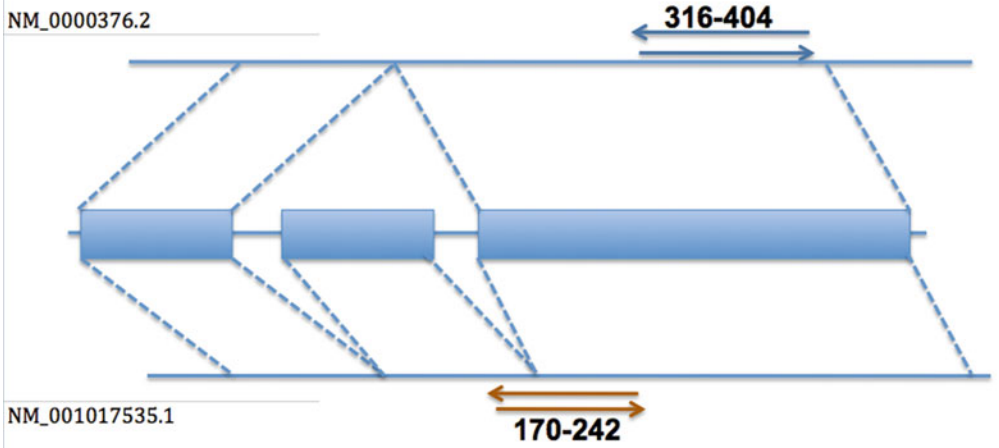
This figure will be printed in b/w

Fig. 18.5. Beacon Designer analysis of primer secondary structure. The program checks for self-dimers, hairpins, cross-dimers, single-nucleotide runs and repeats. These data, together with the ratings, help to base primer choice on more rational criteria.

NM_0000376.2									
Sense Primer	Length	Tm	GC%	Rating	Pair Quality	Pair Rating	Product Length	Product Tm	Product TaOpt
GACCCTGGAGACTTTGAC	18	63.2	55.6	65.6	Good	69	89	81.3	60.8
Anti-sense Primer	**Length**	**Tm**	**GC%**	**Rating**					
CAGGTCATAGCATTGAAGTG	20	62.6	45	72.5					
Probe Sequence	**Length**	**Tm**	**GC%**	**Rating**	**Quality**				
TCCACACACCCCACAGATCC	20	69	60	64.9	Good				

NM_001017535.1									
Sense Primer	Length	Tm	GC%	Rating	Pair Quality	Pair Rating	Product Length	Product Tm	Product TaOpt
TCCCTCAATCCTCATAGC	18	61.7	50	74.2	Good	69.8	72	76	56.8
Anti-sense Primer	**Length**	**Tm**	**GC%**	**Rating**					
GCTCTTCTGTGAGGTCTC	18	62.9	55.6	68.9					

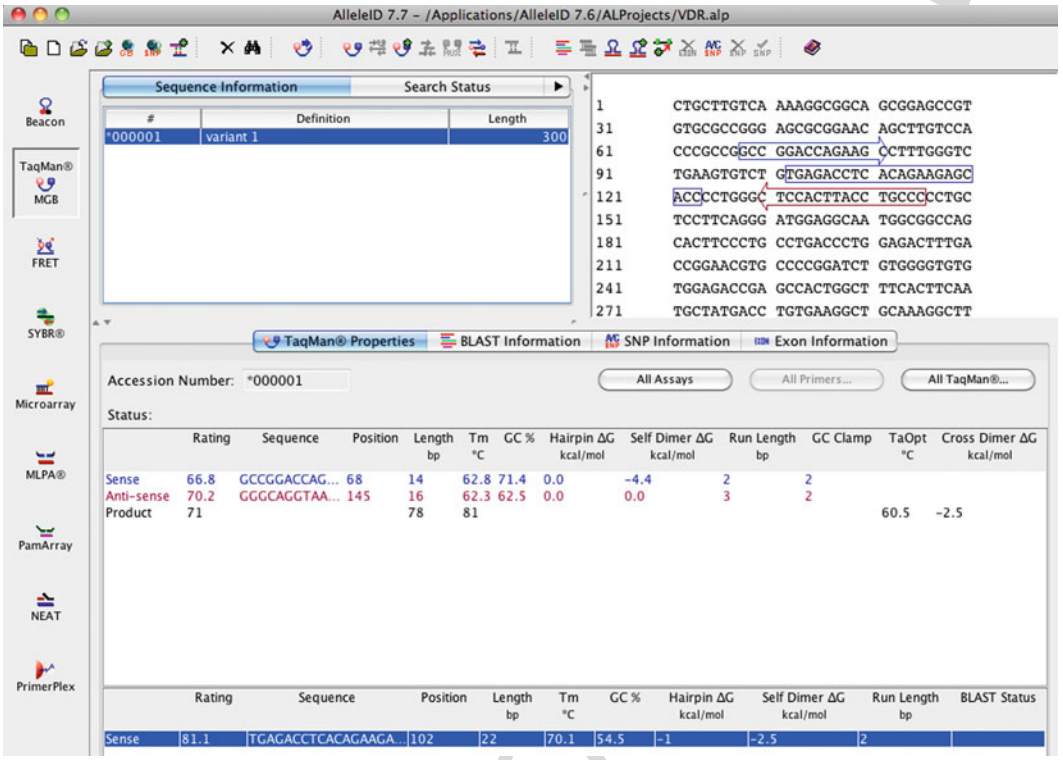
Probe Sequence	Length	Tm	GC%	Rating	Quality
TTCAGACCCAAAGGCTTCCCA	21	69.1	52.4	77.1	Best



This figure will be printed in b/w

Fig. 18.6. Multiplex assay for general VDR/variant 2-specific assay.

experimental requirements via the website (www.sial.com/designmyprobe) and a design specialist will apply the most appropriate design protocol from a suite of possibilities including Beacon Designer, AlleleID (as described above) and Roche FRET design software. All of the options described for these packages are therefore available from this service. A dedicated bioinformatics scientist team can also manage huge alignment projects of several hundred sequences. Alternatively, commercial companies such as Roche and Biosearch offer real-time design tools (e.g. http://www.biosearchtech.com/realtimedesign) that are useful for the less experienced user who does not require the levels of control offered by the above two heavyweight programs. The Biosearch option can be used to design assays for both standard quantification and SNP genotyping and can be used to design up to 10 different assays simultaneously. It offers default and moderate



This figure will be printed in b/w

Fig. 18.7. Amplicon characteristics for variant 1-specific assay (AlleleID). This program suggests a different downstream primer, confirming that there are a large number of possible primer sites on a target, depending on user-defined settings, making it essential that assays are published with primer and probe sequences.

user-modifiable modes (amplicon length, GC content, mononucleotide run length and dimers) and links directly to NCBI to retrieve sequences according to accession number and to confirm priming specificity. For our current assay, the need to place the forward primer for variant 1 at a very specific site precludes its use for the design of a variant 1-specific assay, but it can be used to design a variant 2-specific assay. Using the custom mode, the program offers a wide range of primer options, from which we initially choose three (**Fig. 18.8**).

3.3. In Silico Validation

Primer specificity and target accessibility are two important parameters that must be reviewed prior to commencing empirical assay optimization and validation. When using Beacon Designer or AlleleID, these functions can be performed within the process. Primer specificity is most easily and rapidly checked using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) (*see* **Chapter 6** for more discussion on the application of Primer-BLAST). The program can analyse an amplicon, or one or both primer sequences. However, if a single primer is entered, the template sequence is also required. For the

0032 62 164 81.03

Primer Detail

Oligo	Tm	GC%	Length	5' Pos	3' Pos	5'-Sequence-3'
Forward	66.5	61	18	164	181	GCTTGCCTCCCTCAATCC
Reverse	65.4	48	21	225	205	CACAGACACTTCAGACCCAAA

Identified Mis-Alignments

Self Align	Dimer	3'Align	Hairpin	Total ΔG°	Pair Align	Dimer	3' Align	Bi-ΔG°
Forward	3	2	3	-11.07	Forward/Reverse	3	3	.00
Reverse	2	2	2	-11.04				

Amplicon Detail

Amplicon	Rank	AT Run	GC Run	GC%	Length	Variation
Allele 1	82.31	3	3	52	62	

GCTTGCCTCC CTCAATCCTC ATAGCTTCTC TTTGGGAAGC CTTTGGGTCT GAAGTGTCTG TG

0039 64 164 80.97

Primer Detail

Oligo	Tm	GC%	Length	5' Pos	3' Pos	5'-Sequence-3'
Forward	66.5	61	18	164	181	GCTTGCCTCCCTCAATCC
Reverse	67.5	48	23	227	205	CTCACAGACACTTCAGACCCAAA

Identified Mis-Alignments

Self Align	Dimer	3'Align	Hairpin	Total ΔG°	Pair Align	Dimer	3' Align	Bi-ΔG°
Forward	3	2	3	-11.07	Forward/Reverse	3	3	.00
Reverse	2	2	2	-11.10				

Amplicon Detail

Amplicon	Rank	AT Run	GC Run	GC%	Length	Variation
Allele 1	84.11	3	3	52	64	

GCTTGCCTCC CTCAATCCTC ATAGCTTCTC TTTGGGAAGC CTTTGGGTCT GAAGTGTCTG TGAG

0332 63 163 80.85

Primer Detail

Oligo	Tm	GC%	Length	5' Pos	3' Pos	5'-Sequence-3'
Forward	67.4	55	20	163	182	TGCTTGCCTCCCTCAATCCT
Reverse	65.4	48	21	225	205	CACAGACACTTCAGACCCAAA

Identified Mis-Alignments

Self Align	Dimer	3'Align	Hairpin	Total ΔG°	Pair Align	Dimer	3' Align	Bi-ΔG°
Forward	3	2	3	-11.10	Forward/Reverse	3	3	.00
Reverse	2	2	2	-11.04				

Amplicon Detail

Amplicon	Rank	AT Run	GC Run	GC%	Length	Variation
Allele 1	83.75	3	3	51	63	

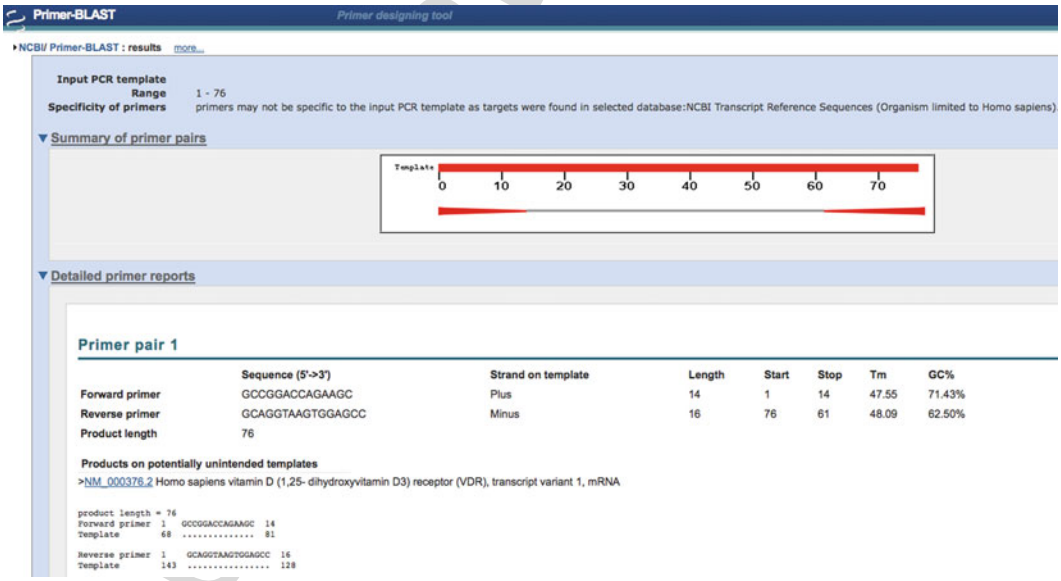
TGCTTGCCTC CCTCAATCCT CATAGCTTCT CTTTGGGAAG CCTTTGGGTC TGAAGTGTCT GTG

This figure will be printed in b/w

Fig. 18.8. Primers and probes for variant 2 suggested by the Biosearch online program. These differ from those obtained using Beacon Designer and AlleleID but are similar to each other. Since primers are cheap, it is best to have these assays synthesized without the probe, determine which one works best and only then order the probe. This has the added advantage of not jeopardizing the probe with primer contamination.

highest level of sequence specificity, the source organism and the smallest database that is likely to contain the target sequence are selected; for broadest coverage, it is best to choose the nr database and not to specify an organism. The results show potential additional targets, with sequence information to allow the researcher to decide whether the chosen primer pairs are acceptable. For VDR variant 1, primer choice is very limited due to the need to target a GC-rich splice junction. The Primer-BLAST results (**Fig. 18.9**) reflect the potential for mis-amplification.

Target accessibility is the second important parameter that must be checked in silico as it is important that primers anneal to areas of minimal secondary structure. This can be quite difficult to achieve when reverse priming from RNA, as RNA is characterized by extensive secondary structure. DNA templates, on the other hand, are less structured, making it easier to identify primer/target combinations that are located in open structures. The most useful analysis tools for determination of optimal and suboptimal secondary structures of RNA or DNA molecules are found on the MFOLD web server (http://mfold.bioinfo.rpi.edu/) (7). There are mfold tools ("m" simply refers to "multiple"), allowing the folding of DNA (http://mfold.bioinfo.rpi.edu/cgi-bin/dna-form1.cgi) and RNA (*see* **Chapter 19** for RNA folding prediction). RNA folding can be carried out at a fixed temperature of 37°C (version 3.2, http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1.cgi) or at variable temperature, using the



This figure will be printed in b/w

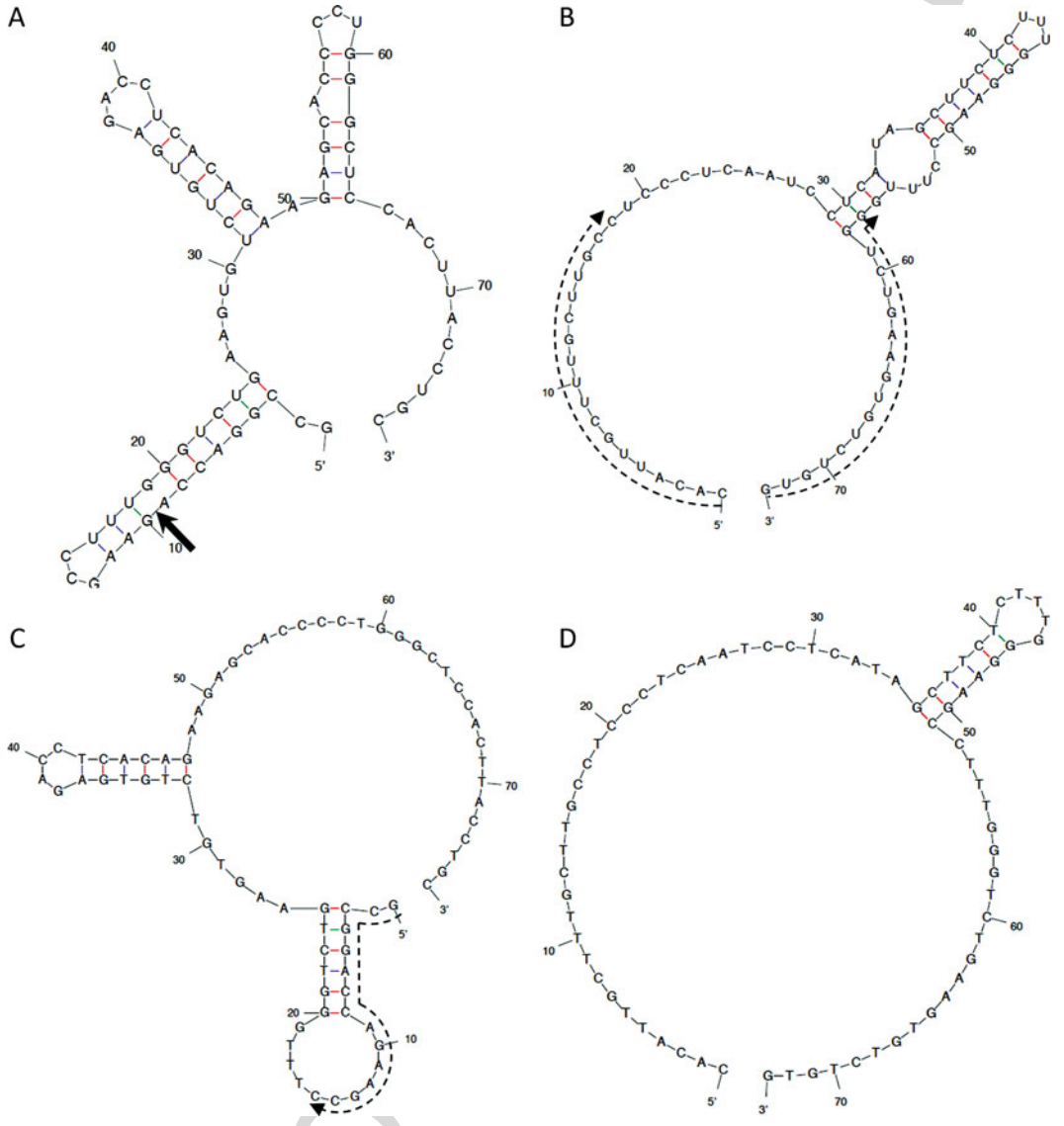
Fig. 18.9. Primer-BLAST output warning of potential non-specificity of primers. The designer needs to analyse the output in detail to decide whether his designs are sufficiently specific.

version that we recommend (version 2.3, http://mfold.bioinfo.rpi.edu/cgi-bin/rna-form1-2.3.cgi). User-definable parameters are significantly more flexible for DNA folding analysis, with an option to fine-tune folding temperature as well as ionic conditions (Na^+ and Mg^{2+}). RNA folding is carried out at 1 M Na^+, with no Mg^{2+}. Folding predictions are based on energy rules specific for DNA (8–14) or RNA (15–18) as well as unpublished parameters. Since any predicted optimal secondary structure for an RNA or a DNA molecule depends on the model of folding and the specific folding energies used to calculate that structure, different optimal folding may be calculated if the folding energies are changed. Because of uncertainties in the folding model and the folding energies, the "correct" folding may not be the "optimal" folding determined by the program. It is important to view several potential optimal and suboptimal structures within a few percent of the minimum energy and use the variation among these structures to determine which regions of the secondary structure can be predicted reliably (7).

Both amplicons chosen for the VDR variants are analysed using the RNA and DNA folding prediction programs. The problem with the assay is that the splice junction dictates the choice of target sequence resulting in determination of a relatively poor secondary structure prediction for the mRNA (**Fig. 18.10a**), which precludes a re-design. Instead, it may be possible to use a higher temperature for the reverse transcription (RT) protocol to optimize the efficiency and reliability of the RT step. The prediction for the folding of the DNA structure looks reasonable (**Fig. 18.10c**) and so it is acceptable to predict that the PCR should be efficient. The secondary structure prediction for variant 2 is somewhat better (**Fig. 18.10b**) and the DNA folding looks good (**Fig. 18.10d**), and should result in an acceptable assay.

3.4. qPCR Set-up

Reliable and reproducible qPCR experimental set-up constitutes a serious technical challenge for the routine use of qPCR, especially in a diagnostic setting. It is not surprising that an important, and obviously fundamental, source of errors associated with qPCR experiments arises from mistakes made during the set-up process. Furthermore, pre-analysis considerations such as testing for template quality, PCR inhibition and optimizing the assay considerations are important contributors to a successful assay but are not generally implemented, and certainly not automated. Prexcel-Q (P-Q) is an interesting tool that addresses these issues by providing a template to allow the user to carry out these analyses systematically and comprehensively. Prexcel-Q is not a data analysis program, it is an extensive qPCR validation, set-up and protocol printout program for each step of the qPCR experimental set-up process. It comprises 35 inter-linked Excel files and can be obtained by contacting Dr. Dario Valenzuela at Iowa State



This figure will be printed in b/w

Fig. 18.10. Mfold RNA/DNA folding of VDR variants. (**a**) The folding structure of mRNA variant 1 is severely handicapped by the need to place one of the primers over the splice junction (indicated by an *arrow*). The Beacon Designer software warns of this limitation and designs this assay only after manual override of basic primer design guidelines. However, the downstream primer-binding site is reasonably accessible and should be accessible to the primer for reverse transcription. (**b**) Both upstream and downstream primer-binding sites in mRNA variant 2, indicated by *dashed lines* with an *arrow*, are accessible, and it would have been easy to design slightly longer amplicons to reduce the secondary structure at the 3′-end of the amplicons even further. (**c**) The upstream primer-binding site in DNA amplicon 1, indicated by the *dashed line* with an *arrow*, lies within a hairpin, although there is a large loop that will help destabilize the secondary structure at high temperatures. The optimal annealing temperature for this assay has been kept high at 60°C, and it may be possible to empirically determine assay conditions that overcome the secondary structure problem. (**d**) There is no secondary structure at either end of DNA amplicon 2, with a very minor potential hairpin in the middle. This assay would be predicted to be efficient.

University Research Foundation (ISURF) at dariov@iastate.edu. P-Q automatically calculates amounts of all reagents required for nuclease treatments, RT and qPCR reactions, provides an estimate of total sample material needed, assists directly with calculations involving standard curve designs, identifies the dynamic dilution range of sample material within which qPCR inhibition is absent and target amplification efficiencies are highest and even estimates the total cost of the assay.

3.5. Data Analysis

The successful generation of high-quality qPCR experimental runs and acquisition of C_q values is followed by the critical data analysis step. The increasing penetration of 384-well and higher throughput qPCR instruments, together with PCR array methods such as the Biotrove and Fluidigm systems, has resulted in data analysis becoming one of the biggest bottlenecks in qPCR experiments. Data analysis is an essential component of the complete assay and this step needs to be handled with care and precision. Many biologists are statistically challenged and may not be able to analyse and interpret the quantitative data from their qPCR assays appropriately. In this context, dedicated data analysis software provides valuable guidance, although it is no substitute for expert consultations with knowledgeable biostatisticians. A wide range of analysis packages can be accessed from http://www.gene-quantification.de/download.html.

One of the first tools was REST (relative expression software tool) (19), which was developed to address the problem of comparing target gene expression levels relative to reference genes in different samples. Importantly, these calculations also accommodated differences in the efficiency of each of the PCR assays. The use of ratios for gene expression measurements makes it complex to perform traditional statistical analysis, as ratio distributions do not have a standard deviation. The latest REST (http://www.gene-quantification.de/rest-2009.html) is a stand-alone software package and has been extended with advanced algorithms using randomization and bootstrapping techniques to take into account issues of multiple reference gene normalization and quantification of the uncertainty inherent in any measurement of expression ratios. The program also allows measurement of not only the statistical significance of deviations but also their likely magnitude, even in the presence of outliers. A visually attractive graphing ability provides a visual representation of variation for each gene, highlighting potential issues such as distribution skew.

In terms of power and sophistication, two commercial software tools stand out: GenEx and qbase PLUS.

GenEx (http://www.multid.se/) (*see* **Note 3**) is a powerful tool for processing and analysing qPCR data. It offers a vast range of tools, ranging from basic data editing and management to advanced, cutting-edge data analysis including selection

and validation of reference genes, classification of samples, gene grouping, monitoring of time-dependent processes and basic and sophisticated statistical and graphic capacity. Statistical features include parametric and non-parametric tests, clustering methods, principal component analysis and artificial neural networks. Most qPCR experimental processes have a nested design with technical replicates at each step, where each step introduces more variance to the data. The noise contribution from each step can be analysed and used to optimize the experimental design for future experiments. One interesting feature is the experimental design optimization module, which calculates how to minimize the total variation for the experimental design with consideration of the experimental budget. GenEx can be used for the analysis of both relative quantification ($\Delta\Delta C_q$) and "absolute" quantification, i.e. relative to a dilution curve, with powerful presentation tools generating visually attractive illustrations. For accurate and robust normalization of qPCR data, GenEx has the advantage of incorporating both NormFinder (20) and geNorm (21) in the software. geNorm (http://medgen.ugent.be/~jvdesomp/genorm/) is used to select an optimal set of reference genes from a larger set of candidate genes. It is critical that the tested candidate genes are not co-regulated. The geNorm process performs a reiterative calculation of the M value, which describes the variation of a potential reference gene compared with all other candidate genes. The gene with highest M value is eliminated until there are only two reference genes left. Usually the last two or three remaining genes are recommended as the optimum candidate reference genes. NormFinder (http://www.mdl.dk/publicationsnormfinder.htm) is also used to find the optimum reference genes but, in contrast to geNorm, takes information of expression in groupings of samples into account. The result is an optimum pair of reference genes that might have compensating expression so that one gene is slightly overexpressed in one group but the other gene is correspondingly underexpressed in the same group.

qbasePLUS (http://www.biogazelle.com/products/qbaseplus) is a relatively new software tool developed by members of the group that invented the geNorm algorithms (21, 22). It is the only non-web-based software tool that can run natively on Mac, Linux and Windows operating systems. The basic concept of qbasePLUS is similar to that of GenEx in that the emphasis is on validation of all technical parameters using appropriate statistical and quality control metrics. It contains four levels of quality control analysis: PCR replicate variation, assessment of positive and negative control samples, determination of reference gene expression stability and evaluation of deviating normalization factors.

Technical replicates (repeated measurements of the same sample in the same run) are useful as they allow for quality control of the technical reproducibility of the qPCR data, provide better accuracy and allow the generation of results when individual qPCR reactions fail. Low reproducibility of the qPCR reaction is often the first indication of an unstable assay that requires optimization. qbasePLUS and GenEx automatically deal with technical replicates or repeated measurements. Outliers can be excluded so that they do not contribute to a significant bias and large errors. However, it is important to note that biological variability is often much larger than RT technical variability which is also much larger than PCR technical variability.

The second type of quality control allows an evaluation of the positive and negative sample controls. An amplification signal in the NTC sample indicates a potential contamination issue or formation of primer dimers; hence both packages flag suspicious no-template control (NTC) samples based on a user-defined or default thresholds.

Reference gene or normalization factor stability is the third type of quality control. The user is able to choose a minimal acceptable normalization factor stability by defining a threshold value for two indicators of expression stability of the used reference genes: the geNorm expression stability value of the reference gene (M) (21) and the coefficient of variation of the normalized reference gene relative quantities (CV) (22). qbasePLUS has extended and improved the functionality of geNorm by allowing ranking of candidate reference genes up to the single most stable gene, by combining the calculation of relative quantities and geNorm analysis and by allowing the processing of experiments with missing data in a way that has the lowest impact on the overall analysis through intelligent retention of as many data points as possible.

Inspection of normalization factors is a very useful feature and constitutes a fourth type of quality control. qbasePLUS displays the geometric mean of selected reference genes (the normalization factor) for each sample. These factors should be similar for all samples if approximately equal amounts of equal quality input material were used. High variability of the normalization factors indicates large differences in starting material quantity or quality, or a problem with one of the reference genes, which may be either not stably expressed or not adequately measured.

There are alternative, usually web-based tools available: QPCR (https://esus.genome.tugraz.at/rtpcr/) allows storage, management and analysis of qPCR data (23). It comprises a parser to import data from qPCR instruments and includes a variety of analysis methods to calculate cycle threshold and amplification efficiency values. The analysis pipeline includes technical and biological replicate handling, incorporation of sample- or

gene-specific efficiency, normalization using single or multiple reference genes, inter-run calibration and fold change calculation. Moreover, the application supports an assessment of error propagation throughout all analysis steps and allows the conduction of statistical tests on biological replicates (*see* **Note 4**). Results can be visualized in customizable charts and exported for further investigation. Calculation of amplification efficiencies for RT-PCR experiments (CAmpER) (http://camper.cebitec.uni-bielefeld.de) is a limited, not very user-friendly web-based tool for the basic analysis of qPCR assay runs on Roche's LightCycler, the Biorad Opticon and the Qiagen Rotor-Gene.

Currently, transparent data interchange between instrument software and third-party data analysis packages, colleagues and collaborators, and between authors, peer reviewers, journals and readers is difficult. One solution to this problem is the development of the Real-time PCR Data Markup Language (RDML), a structured and universal data standard for exchanging qPCR data (24). The aim is to enable transparent exchange of annotated qPCR data by providing sufficient information for a third party to be able to understand the experimental set-up, re-analyse the data and interpret the results. A useful tool for converting instrument-specific data to the RDML format for convenient data interchange is found on the RDML website (http://medgen.ugent.be/rdml/chooseTool.php).

As qPCR moves from the realm of a cutting-edge technology to becoming a routine technique, there is the danger of losing the attention to experimental detail that is required to ensure that the C_q values generated accurately reflect the original biology under examination. When handling large numbers of samples, these details become even more difficult to manage. When sample sets are divided between several plates, or processed independently, additional factors such as variability between these individual runs must also be accommodated. While this can be carried out using some instrument packages, such as the Stratagene (Agilent) MxPro suite, it is only a small matter of scale before management of these data sets requires sophisticated data management and processing. The publication of the MIQE guidelines is testament to the drive towards improving the quality, uniformity and transparency of qPCR experiments. There is a clear requirement for open access to qPCR data in a system analogous to those developed for microarray users. This would promote a heightened awareness of the requirement for diligent experimental procedure and also make it possible for scientists to re-analyse combined data sets. These are also essential factors to consider when selecting software to use for assistance in designing experiments and analysing data. Software packages should be considered as scientific tools and applied intelligently, using processes that are absolutely transparent to the user.

4. Notes

1. Arguably the most common biostatistical trap for many users is failing to distinguish between exploratory and confirmatory biostatistical studies. For a study to arrive at a statistically significant conclusion, its hypothesis needs to be defined *before* data collection and analysis; this is a confirmatory study. In an exploratory study, data are collected and analysed for various patterns and trends. The exploratory study *may* produce hypotheses, but due to conscious or subconscious manipulation of the data, multiple testing issues will be raised and statistical significance will risk being unreliable at best. A correct approach for a complete statistical analysis involves first an exploratory study to generate one or a few hypotheses that are subsequently validated or discarded in a confirmatory study based on a freshly collected data set. An exploratory study may be of value on its own, but authors should stress that the results produced by such a study have not been verified statistically and only constitute hypotheses pending statistical confirmation.
2. Beacon Designer is our preferred assay design program, and there are a few considerations we would like to share: Designs of primer or probe oligonucleotide sequences often involve trade-off issues between potentially conflicting parameters such as melting temperature, self-hybridization and cross-hybridization risks, and GC content. The Beacon Designer program does a very good job at finding optimal balances between these parameters for a list of different detection chemistries. However, in many applications, primer or probe sequences are restricted to specific locations on the target sequence. The primer or probe location is a parameter that no predesigned algorithm can take into account; it is something that depends on the specific target of the assay. One example is a desire to have a primer or a probe covering an exon–exon junction to avoid amplification of genomic DNA. Another example is a desire to have a primer or a probe in a conserved region of a sequence alignment to allow detection of all species in a group of species. Conversely, it may be desired to have a primer or a probe in a unique region of a sequence alignment to allow detection of only one species in the group of species. The "Alternate Assays" setting in Beacon Designer is a useful feature as it allows the user to specify the number of assay designs that Beacon Designer will generate. The default setting is 5 (for normal assays) or 2 (for LNA assays) and the maximum setting is 50. By using the maximum 50, the list

of additional design alternatives often allows users to find designs at desired locations.

It is worth noting that Beacon Designer has a limit of sequence length range of 1200 bp for calculations of target secondary structure features that may inhibit PCR efficiency. For longer target sequences, the first 1200 bp are used by default for the calculation. However, the location and the length of the calculated interval can be adjusted. By shortening the interval and locating it around a desired location, the algorithm is driven to produce assay designs near the desired location.

One feature available in Beacon Designer is an algorithm to identify assay designs for SNP classification. This feature drives the assay design algorithm to the location of the SNP. Although this is the intended purpose of the feature in Beacon Designer, users can also use this feature to drive the assay design algorithm to any location on the target sequence by introducing fake SNPs at desired locations.

To some extent, several of the unique features available in AlleleID can also be accomplished in Beacon Designer. Sequence alignment performed in a sequence alignment program such as CLC viewer or ClustalW can be used to identify conserved or unique target sequence regions. Then by limiting assay design locations in Beacon Designer, it may be possible to find suitable designs at desired locations such that they do target a multitude of species or a unique subset of the sequence alignment.

Under circumstances that an assay design fails to produce any results, this may not necessarily be cause to abandon the attempt to create an assay design for this particular target. Beacon Designer reports which parameter limit was crossed and relaxation of the limiting parameter may enable a new calculation to come up with results. The introduced trade-off may very well be reasonable, given the specifics of the target and the assay. However, this procedure requires detailed understanding of each parameter's consequence on assay performance and a careful balance between opposing desires. The report from Beacon Designer on failing parameters may thus be a useful item to look out for.

3. One of the advantages with the GenEx software is that two reference gene validation algorithms have been made easily available through the common user interface of GenEx. Using several reference gene validation algorithms enables the methods to be cross-validated against each other for more confident identification of preferred reference genes.

   A nice feature of GenEx is that the NormFinder algorithm has been supplemented with an analysis of accumulated

standard deviations. The analysis thus not only provides a ranking of preferred reference genes but also estimates the optimal number of reference genes for the assay. Multivariate data are inherently difficult to visualize due to the high dimensionality of the data. GenEx has capabilities to produce all of 1D, 2D and 3D graphs. Grouping subsets of the data and assigning specific colours and symbols to them may add additional visual cues to facilitate visual understanding of the data. A specific tip is to use variation in colour to distinguish categories of one data set dimension (for example, time after induction) and variation in plot symbol to distinguish categories of another data set dimension (for example, dose).

Replicates can be used to reduce confounding variability from technical handling steps in the assay. Taking advantage of a pilot study, GenEx has tools to determine optimal distribution of replicates in a nested experimental design under conditions of economics and limits to total number of samples. Another important use of replicates is to evaluate statistical significance. The larger the number of biological replicates, the smaller the treatment effect verified at a given statistical significance. Based on estimates of standard deviations within each sample group, which may have been obtained from a pilot study, GenEx has a feature to calculate the necessary number of biological replicates for validation of a given observed treatment effect and level of statistical significance. This is useful to ensure that enough samples are collected to validate the studied hypothesis.

Optimal statistical tests often depend on the underlying sample distribution. Choices include the parametric *t* test and non-parametric tests. GenEx provides automatic tests of the underlying sample distribution and provides recommendations of the most suitable statistical test depending on sample distribution features. It should be kept in mind, though, that the tests of the underlying sample distribution become unreliable for small sample sizes and the recommendations provided by GenEx should only be considered tentative pending deeper analysis by a trained biostatistician. The easy access of several different types of statistical tests within GenEx is particularly useful during exploratory studies. Careful consideration of underlying sample distributions and choice of suitable statistical test is critical only for confirmatory statistical studies.

Visualization of expression profiles and hierarchical clustering are recommended first steps in exploratory studies of multivariate data sets. The modular design of GenEx's user interface allows for easy transitions between different analysis methods on the same data set. The data set can thus,

for example, be tentatively analysed with hierarchical clustering to identify related groups, these groups assigned specific colours and the groups analysed in one of the more advanced multivariate data analyses methods available in GenEx. An interesting trick is to transpose the data in order to toggle analysis between the perspective of genes and the perspective of the sample groups.

4. Biological replicates are often the most expensive among the steps in a nested design, but they are necessary for development and confirmation of statistical tests. More biological replicates are needed if high statistical significance is desired; the assay has a high confounding variability and/or the studied effects are small. Technical replicates are valuable to reduce confounding variability due to the technical handling. However, the optimal distribution of replicates will depend on the sensitivities of each of the technical handling steps to the introduction of confounding variabilities and the amplitude of the studied effect in the assay. It is therefore recommended that a *pilot study* be performed and used as an exploratory study. The pilot study is then used to
   - evaluate the amplitude of confounding variabilities within each technical handling step so that the optimal distribution of replicates can be determined;
   - estimate the amplitude of the studied effect in the assay so that the number of biological replicates necessary for the desired degree of statistical significance is attainable;
   - optimize assay experimental conditions and
   - validate reference genes for data normalization.

Acknowledgements

S.A.B. would like to thank the charity B&CR (Charity Number 1119105) for support.

References

1. Bustin, S. A., Benes, V., Garson, J. A., et al. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55, 611–622.
2. Bustin, S. A. (2010) Why the need for qPCR publication guidelines? – The case for MIQE. *Methods* 50, 217–226.
3. Crofts, L. A., Hancock, M. S., Morrison, N. A., and Eisman, J. A. (1998) Multiple promoters direct the tissue-specific expression of novel N-terminal variant human vitamin D receptor gene transcripts, *Proc Natl Acad Sci USA* 95, 10529–10534.
4. Lefever, S., Vandesompele, J., Speleman, F., and Pattyn, F. (2009) RTPrimerDB: the

portal for real-time PCR primers and probes. *Nucleic Acids Res* 37, D942–D945.

5. Pattyn, F., Robbrecht, P., De Paepe, A., et al. (2006) RTPrimerDB: the real-time PCR primer and probe database, major update 2006. *Nucleic Acids Res* 34, D684–D688.
6. Pattyn, F., Speleman, F., De Paepe, A., and Vandesompele, J. (2003) RTPrimerDB: the real-time PCR primer and probe database. *Nucleic Acids Res* 31, 122–123.
7. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31, 3406–3415.
8. SantaLucia, J., Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. 95, 1460–1465.
9. Bommarito, S., Peyret, N., and SantaLucia, J., Jr. (2000) Thermodynamic parameters for DNA sequences with dangling ends. 28, 1929–1934.
10. Peyret, N., Seneviratne, P. A., Allawi, H. T., and SantaLucia, J., Jr. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. 38, 3468–3477.
11. Allawi, H. T., and SantaLucia, J., Jr. (1998) Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects 37, 9435–9444.
12. Allawi, H. T., and SantaLucia, J., Jr. (1998) Thermodynamics of internal C.T mismatches in DNA. 26, 2694–2701.
13. Allawi, H. T., and SantaLucia, J., Jr. (1998) Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. 37, 2170–2179.
14. Allawi, H. T., and SantaLucia, J., Jr. (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. 36, 10581–10594.
15. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* 288, 911–940.
16. He, L., Kierzek, R., SantaLucia, J., Jr., et al. (1991) Nearest-neighbor parameters for G.U mismatches: [formula; see text] is destabilizing in the contexts [formula; see text] and [formula; see text] but stabilizing in [formula; see text]. 30, 11124–11132.
17. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) Stabilities of consecutive A.C, C.C, G.G, U.C, and U.U mismatches in RNA internal loops: Evidence for stable hydrogen-bonded U.U and C.C.+ pairs. 30, 8242–8251.
18. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1990) Effects of GA mismatches on the structure and thermodynamics of RNA internal loops. 29, 8813–8819.
19. Pfaffl, M. W., Horgan, G. W., and Dempfle, L. (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30, e36.
20. Andersen, C. L., Jensen, J. L., and Orntoft, T. F. (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64, 5245–5250.
21. Vandesompele, J., De Preter, K., Pattyn, F., et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3, 0034.0031–0034.0011.
22. Hellemans, J., Mortier, G., De Paepe, A., et al. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* 8, R19.
23. Pabinger, S., Thallinger, G. G., Snajder, R., et al. (2009) QPCR: Application for real-time PCR data management and analysis. *BMC Bioinformatics* 10, 268.
24. Lefever, S., Hellemans, J., Pattyn, F., et al. (2009) RDML: structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic Acids Res* 37, 2066–2069.

AQ8

Chapter 18

Q. No.	Query
AQ1	Please provide email address for "Anders Bergkvist and Tania Nolan".
AQ2	Please check the edited affiliation.
AQ3	Please check if the sentence "Hence we have organized this chapter..." says the intended meaning.
AQ4	Please check "...in the customization both programs permit the user" for clarity.
AQ5	"Mfold", "MFOLD", and "mfold" are given. Please advise if usage should be made consistent.
AQ6	Please check if the edit made to the sentence "Then by limiting assay design locations in Beacon Designer..." is OK.
AQ7	Please check if the sentence "The data set can thus, for example, be tentatively..." says the intended meaning.
AQ8	Please provide journal title for Refs. 8, 9, 10, 11, 12, 13, 14, 16, 17, and 18.