Appl Microbiol Biotechnol (2014) 98:9595–9608
DOI 10.1007/s00253-014-6165-x

MINI-REVIEW

Mathematical tools to optimize the design of oligonucleotide probes and primers

Daniel R. Noguera · Erik S. Wright · Pamela Camejo · L. Safak Yilmaz

Received: 28 July 2014 / Revised: 12 October 2014 / Accepted: 14 October 2014 / Published online: 31 October 2014

Abstract The identification and quantification of specific organisms in mixed microbial communities often relies on the ability to design oligonucleotide probes and primers with high specificity and sensitivity. The design of these oligonucleotides (or "oligos" for short) shares many of the same principles in spite of their widely divergent applications. Three common molecular biology technologies that require oligonucleotide design are polymerase chain reaction (PCR), fluorescence in situ hybridization (FISH), and DNA microarrays. This article reviews techniques and software available for the design and optimization of oligos with the goal of targeting a specific group of organisms within mixed microbial communities. Strategies for enhancing specificity without compromising sensitivity are described, as well as design tools well suited for this purpose.

D. R. Noguera · P. Camejo
Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI, USA

D. R. Noguera
Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI, USA

E. S. Wright
Systems Biology Theme, Wisconsin Institute for Discovery, University of Wisconsin-Madison, Madison, WI, USA

E. S. Wright
Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA

L. S. Yilmaz
Program in Systems Biology, University of Massachusetts Medical School, Worcester, MA, USA

D. R. Noguera (✉)
1415 Engineering Drive, Madison, WI 53706, USA
e-mail: noguera@engr.wisc.edu

Keywords Oligonucleotides · DNA probes · FISH · PCR · Microarrays · Mismatch stability · Microbial diversity · Primer design

Introduction

Investigations in microbial ecology often involve the detection and quantification of specific microorganisms within communities in which closely related organisms may also be present. This is the case, for example, when trying to assess the relative contribution of specific groups of microorganisms to global inorganic and organic nutrient cycling (Abell et al. 2010; Cottrell and Kirchman 2000; Lenk et al. 2011; Mosier and Francis 2008). Sometimes the target microbial group has unique phenotypic or genotypic characteristics that allow it to be clearly differentiated from other organisms in the community, but this aim is made more challenging when the group's distinguishing characteristics are similar to those of other nontargeted organisms. Therefore, it is often necessary to employ sophisticated design and experimental approaches to ensure accurate identification of the targeted group of organisms.

Identification methods based on phylogenetic gene markers, such as the small subunit (ssu) or large subunit (lsu) ribosomal RNA (rRNA), as well as functional genes are now commonplace (Jones et al. 2008; Klein et al. 2001; Loy et al. 2007; Pester et al. 2012). Techniques such as fluorescence in situ hybridization (FISH), polymerase chain reaction (PCR), quantitative PCR (qPCR), and phylogenetic microarrays are all based on successfully designing oligonucleotide probes or primers (collectively referred here as oligonucleotides, or oligos for short) that hybridize to nucleic acid templates from the targeted organisms. These oligos must also be designed to limit cross-hybridization with nontarget organisms in order to minimize the number of false positive

identifications. The ssu rRNA is the most widely used phylogenetic marker and is therefore the target gene for many oligonucleotide designs. In addition, the availability of sizeable rRNA repositories, such as the Ribosomal Database Project (RDP) (Cole et al. 2007), SILVA (Quast et al. 2013), and Greengenes (DeSantis et al. 2006), as well as databases of previously used oligos (Loy et al. 2007), facilitates the widespread use of the ssu rRNA as a target for oligo designs. As our awareness of microbial diversity continues to expand along with the size of rRNA databases, the design of high-specificity oligonucleotides has become more challenging (Pfeiffer et al. 2014).

With a limited number of sequences in early ssu rRNA databases, the task of finding target sites with sufficient specificity to the organisms of interest was relatively easy compared to the same task using current databases, which encompass much greater sequence variability amongst target and nontarget organisms. Likewise, oligos designed to target specific groups that were defined using a limited number of sequences can become outdated (i.e., have incomplete coverage of the intended target group) when additional sequences reveal a higher diversity in the targeted group. A classical example of this is in the design of FISH probes and PCR primers intended to target all cells within the Bacteria domain. The initial design of probe EUB338 (Amann et al. 1990) was updated when new sequences revealed bacterial cells that would not hybridize with the original probe (Daims et al. 1999). Similarly, the original design of primers 27f and 1492r (Weisburg et al. 1991) was later revisited to improve their coverage of the domain bacteria in PCR amplifications (Frank et al. 2008). Fortunately, as the diversity represented in sequence databases increases, oligos designed with large sequence sets may take enough diversity into consideration to withstand the test of time.

New advances in oligonucleotide design have been developed to harness the ever-increasing breadth of modern sequence repositories. These advances draw on recently developed experimental techniques and mathematical models to provide cutting-edge methods for the design of oligos that simultaneously maximize sensitivity and specificity to a target group of organisms. Some of these tools have the ability to accurately identify nontargets present in large databases with high-throughput. Here, we describe both software tools and experimental approaches that have been used for the design of oligonucleotides in recent years.

Principles of oligonucleotide design

Hybridization is the process of an oligonucleotide forming a duplex with a complementary sequence. Therefore, ideal oligos should (i) strongly hybridize to their intended target(s), and (ii) not hybridize to nontargeted sequences. Although this is a simple fundamental principle, it is a complex task in practice because of the immense diversity of potential target and nontarget sequences. On one hand, if the targeted group of organisms is narrowly defined (e.g., at the strain or species level), it may not be possible to find a suitable site in the targeted gene that is unique to the group. On the other hand, if the targeted group is broadly defined (e.g., class or phylum level), finding a target site that is common to all of the organisms within the group but absent in other organisms becomes challenging. These considerations define two of the key goals in the design of oligonucleotides. That is, the oligos should have high *specificity* to the targeted organisms while maintaining high *coverage* of the targeted group.

Besides specificity and coverage, a third concept considered in oligonucleotide design is *sensitivity*, which refers to maximizing the number of target molecules hybridized to oligos so as to enable the highest possible degree of detection when abundance of the target is low. The detection limit, defined as the minimum number of targets that need to be present for positive detection, is tightly linked to the concept of sensitivity. However, since maximizing hybridization with the target molecule can simultaneously increase hybridization to nontargets (hence reducing specificity), detection limits cannot be optimized independently of specificity and coverage considerations. In many cases, it is preferable to hybridize under stringent conditions (e.g., high annealing temperature or high denaturant concentrations) to minimize nontarget hybridizations, even though this will not result in the lowest possible detection limit. This trade-off can be seen in Fig. 1, where increasing stringency simultaneously decreases both target and nontarget signal. The ideal level of stringency occurs at the point where target signal is maintained at an easily detectable level, but nontarget hybridization is negligible. This

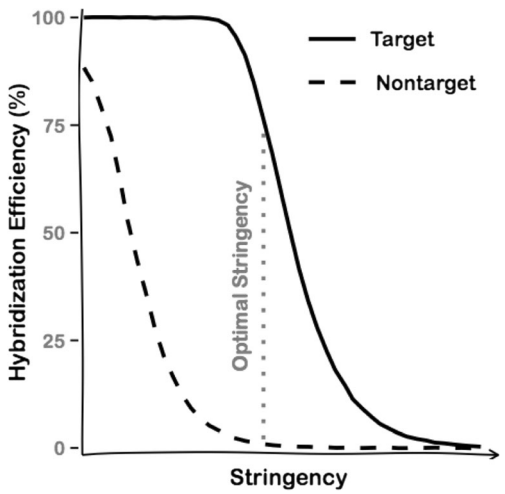

Fig. 1 Denaturation curves illustrating the trade-off between sensitivity and specificity. Increasing stringency results in less nontarget hybridization at the expense of decreasing target sensitivity. The optimal level of stringency occurs at the point where the nontarget signal is minimized (specificity is maximized) and the target signal (sensitivity) is maximized

optimal point can be determined experimentally by testing multiple levels of stringency and examining the effect on target and nontarget hybridizations. If the optimal stringency results in weak signals from the targeted organisms, then approaches to increasing specificity without compromising sensitivity can be implemented. In PCR, a method called Touchdown-PCR, where decreasing levels of stringency are used in each annealing step, ensures that the experiment reaches an optimal level of stringency with a high level of sensitivity (Korbie and Mattick 2008). For FISH, signal from the targeted organisms can be amplified through indirect approaches, such as in catalyzed reporter deposition (CARD)-FISH (Pernthaler et al. 2002).

An alternative approach is to increase specificity without compromising sensitivity by hybridizing under less stringent conditions while mitigating nontarget hybridizations. This can be achieved with competitor probes (Manz et al. 1992) or blocking primers (Wilcox et al. 2014) that prevent mismatched targets from hybridizing to the oligo. Yet, another approach to maximize specific identification involves the simultaneous hybridization of multiple oligos to the same target organism, such that the overlap between nontarget organisms cross-hybridizing to each oligo is minimized. This is chiefly used in PCR experiments where both forward and reverse primers must anneal for exponential amplification to occur. In qPCR, a third oligo can be used through the addition of an internal probe that will produce a fluorescent signal only from the targeted organism even if amplification of DNA from nontargeted organisms took place (Livak et al. 1995). Analogously, two or more probes can be used to target the same organism with FISH, and while each probe may hybridize to many nontargets, only the target organisms will hybridize to both probes simultaneously (Wright et al. 2014a).

Thermodynamic principles of hybridization

Successful optimization of oligonucleotide sensitivity, specificity, and coverage requires accurate predictive tools during the design step. The basis for such tools is the thermodynamics of nucleic acid interactions (reviewed in Turner (2000)). Oligonucleotide sensitivity is addressed by the calculation of the thermodynamic affinity of the designed oligo to the target molecule. Although experimental optimization can partially solve problems arising from low target abundance, an oligo with very low affinity cannot yield sufficient signal under any condition, due simply to the lack of hybridization. Historically, semiquantitative formulations based on the GC content and oligo length have been employed to evaluate oligo affinity (Stahl and Amann 1991). Modern approaches use the entire sequence context to quantify the thermodynamic affinity in the form of a Gibbs free energy change (ΔG°) of duplex formation. This is achieved with the nearest-neighbor model, which sums the free energy contributions of nearest-neighbor base-stacking interactions between adjacent base pairs in a duplex (for an oligo of length L, there are L-1 nearest neighbors), and adds an initiation penalty to this sum to obtain a total ΔG°. The free energy parameters associated with all possible nearest neighbor sequences (there are 10 possibilities for DNA/DNA and RNA/RNA duplexes, and 16 for DNA/RNA duplexes) have been determined using *in solution* hybridizations with tens of different oligonucleotide pairs and are available for DNA/DNA (SantaLucia 1998), RNA/RNA (Turner and Mathews 2010), and DNA/RNA (Sugimoto et al. 1995) interactions.

Generally, the calculation of a free energy change of duplex formation is not sufficient to quantify the overall thermodynamic affinity of an oligonucleotide to its target. Both the target and oligo may form structures through self-folding interactions that compete with duplex formation. Therefore, free energy penalties for the unfolding of these structures need to be calculated and incorporated into the calculation of overall affinity. Thermodynamics of nucleic acid interactions in structured single strands have been studied for both DNA (SantaLucia and Hicks 2004) and RNA (Mathews et al. 1999b) to determine the free energy changes of formation for different structural motifs. With specialized software that uses these thermodynamic parameters (Markham and Zuker 2008; Zuker 2003), the free energy changes for unfolding the oligonucleotide and target molecules can be calculated. In addition, oligonucleotide dimerization due to intermolecular interactions within the oligonucleotide pool may also compete with duplex formation when the dimers have a melting temperature (temperature where half of the oligos are dimerized) similar to the hybridization temperature. Such intermolecular interactions are the cause of primer-dimer artifacts in PCR and must therefore be taken into account during primer design (Wright et al. 2014b). The ΔG° of duplex formation and ΔG° penalties for folding and dimerization can be combined in an equilibrium model to calculate an overall free energy change for hybridization, which describes the thermodynamic affinity of the oligonucleotide to its target (Mathews et al. 1999a).

Thermodynamics is useful also for evaluating oligonucleotide specificity and coverage. In particular, the prediction of the likelihood of oligonucleotide hybridization with mismatched sequences depends on the cumulative effect of mismatches on the thermodynamic stability of the duplex. Mismatch types that are commonly encountered during oligonucleotide design are shown in Table 1. The destabilizing effect of a mismatch can be quantified as a positive (unfavorable) difference in ΔG° of duplex formation, or $\Delta\Delta G^\circ$ (Sugimoto et al. 2000; Yilmaz et al. 2008), which has two components. First is the ΔG° lost due to the nearest-neighbor base pairs that are interrupted by the inclusion of a mismatch. For instance, a single mismatch eliminates favorable base-stacking interactions between the original perfect matching base pair and its

Table 1 Typical mismatch patterns in probe/target sequences

Mismatch type	Perfect match[a]	Mismatch loop[a]	NN lost[b]	Permutations[c]
Single	●●●/●(●)●	●$^{\Lambda}$●/●$_{v}$●	2	104/192
Bulge (insertion)	●●/●()●	●●/●$_{v}$●	1	64/64
Bulge (deletion)	●●●/●(●)●	●$^{\Lambda}$●/●●	2	64/64
Tandem	●●●●/●(●●)●	●$^{\Lambda\Lambda}$●/●$_{vv}$●	3	1176/2304
Multiple[d]	●●●●●●/●(●)●●(●)●	●$^{\Lambda}$●●$^{\Lambda}$●/●$_{v}$●●$_{v}$●	4	na[e]

[a] Patterns represent short segments of nucleotides excerpted from a longer probe/target duplex. Dot (●) indicates a perfect match, **v** and $^{\Lambda}$ indicate a mismatch. Parentheses encompass positions on the target that change to form the mismatches

[b] Number of perfect match nearest neighbor stacking pairs that are eliminated upon insertion of the mismatches

[c] Number of distinct mismatch loops considering all possible matching and mismatching bases. The first number is for DNA/DNA or RNA/RNA duplexes followed by that for DNA/RNA duplexes

[d] Two separate single mismatches are shown as an example

[e] Not applicable. The free energy changes for mismatches separated by more than one base are added to calculate an overall effect. When there is only one base pair left between mismatches, loops are merged to eliminate isolated base pairs

two nearest-neighbors, one to either side of the mismatch position (Table 1). The second component of ΔΔG° is the ΔG° of formation for the internal loop created by mismatch insertion. This is a function of the type of mismatch and the nearest neighboring base pairs, which define the loop altogether (Table 1). Thus, the prediction of mismatch stability requires not only the nearest-neighbor parameters for perfect matches but also the estimates of ΔG° values for all possible permutations of internal loops. For single mismatches in DNA/RNA duplexes, there are 192 unique permutations of internal loops (4 possibilities for each one of the nearest neighbors and 12 for the mismatch type), whereas for DNA/DNA or RNA/RNA, there are only 104 possibilities due to symmetrical occurrences. While parameters for all permutations are available for DNA/DNA (SantaLucia and Hicks 2004) and RNA/RNA (Turner and Mathews 2010) hybridizations, only a subset has been evaluated in DNA/RNA duplexes (Sugimoto et al. 2000; Watkins et al. 2011). The multitude of loop parameters in oligo design highlights how even a moderate set of single mismatches have ΔG° that vary over a wide range of values (Yilmaz et al. 2008). For this reason, discrete categorization of mismatch stability based solely on the type of mismatched base pair (Ludwig et al. 2004; Pozhitkov et al. 2006) is oversimplified.

Other simple mismatch motifs (Table 1) include bulges, which form upon insertion or deletion of a base, and tandem mismatches. Bulges generally have a highly destabilizing loop, but in some cases, they may have an overall effect comparable to a single mismatch, especially in the case of insertions since only one nearest-neighbor is lost (Yilmaz et al. 2012). Tandem mismatches are the equivalent of two adjacent mismatches but are generally less destabilizing than two separate mismatches due to one less nearest neighbor being eliminated (Table 1). For bulges and tandem mismatches, loop ΔG° values have not been measured exhaustively, but systematic studies in RNA/RNA and DNA/DNA duplexes allow reasonable estimations for all permutations (SantaLucia and Hicks 2004; Turner and Mathews 2010). The only exception is tandem mismatches in DNA/DNA, for which ΔG° values are mostly approximated using theoretical rules (SantaLucia and Hicks 2004). When there are multiple mismatches, as in two separate mismatches (Table 1), their cumulative effect can easily be approximated by summing up the ΔΔG° of individual components. In some cases, complex patterns may occur, such as bulges with multiple insertions or deletions and three adjacent mismatches making a large internal loop. Due to the sheer number of possibilities in such cases, theoretical free energy extrapolations are used rather than systematic measurements for different permutations (Jacobson and Stockmayer 1950; SantaLucia and Hicks 2004; Turner and Mathews 2010). All of these free energy rules are incorporated into the software UNAFold (Markham and Zuker 2008) making it an ideal tool for predicting the ΔG° of DNA/DNA and RNA/RNA duplexes. However, none of the motifs other than single mismatches have been systematically studied in DNA/RNA duplexes, and therefore, approximations are necessary for hybridizations with RNA targets.

Another complication with mismatch stability is positional effects. Typically, a central mismatch in a duplex is more destabilizing than one closer to the terminals (Pozhitkov et al. 2006; Urakawa et al. 2002). A mechanistic model that explained positional effects observed in a microarray platform was recently published by Yilmaz et al. (2012). According to this model, a mismatch near one of the terminals creates an alternative, thermodynamically more favorable state of the duplex, wherein the bases at and after mismatch position toward the nearest terminal remain unpaired. In cases where the mismatch is within four nucleotides from a terminal, this dangling state may have a more negative overall free energy change than the zipped but mismatched duplex, explaining the lessened effect of near-terminal mismatches on duplex stability. Computational tools that are capable of searching the duplex state with the minimum free energy, such as UNAFold (Markham and Zuker 2008) and ProbeMelt (Yilmaz et al. 2012), readily take this effect into account. However, it remains unknown whether the positional effect of mismatches on free energy can be generalized to all possible applications.

Finally, all ΔG° information discussed in this section is available together with the corresponding enthalpic (ΔH°) and entropic (ΔS°) components, allowing the thermodynamic

predictions to be made at different hybridization temperatures based on the fundamental definition of Gibbs free energy change ($\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$). Because of specificity requirements, hybridizations are generally performed at a relatively high stringency. If a denaturant like formamide is used for stringency, instead of a temperature increase, then linear free energy models that predict ΔG° values as a linear function of denaturant concentration can be employed (Courtenay et al. 2000; Schellman 1978; Yilmaz and Noguera 2007).

Probe design for FISH applications

In the most widely used application of FISH, a DNA oligonucleotide probe is designed to target the rRNA molecule, which may be present in thousands of copies per cell (Bakshi et al. 2012). Most probes target the ssu rRNA, although there are some commonly used probes that target the lsu rRNA. In general, hybridizations in FISH are carried out with a denaturant and relatively long incubation times (>1 h), and therefore, equilibrium between probes and targets is obtained. Many variations in FISH protocols (for a recent review, see Wagner and Haider (2012)) have been developed to improve the signal strength from hybridized probes, such as for detecting microorganisms with low ribosomal content (Fazi et al. 2008; Zwirglmaler et al. 2003) or to apply FISH to the detection of low-copy messenger RNA (mRNA) instead of rRNA (Pilhofer et al. 2009; Pratscher et al. 2009). Regardless of the FISH application, the fundamental challenge of achieving high specificity, coverage, and sensitivity remains the same.

The most widely used software for probe design for FISH applications is ARB (Ludwig et al. 2004). Using ARB, designers can interactively select the sequences to be used in the design, cluster them to define target and nontarget phylogenetic groups, search for sequence regions that are unique to the targeted organisms, and identify potential target sites with high coverage of the targeted group. To assess probe specificity, prediction of mismatch stability is provided in ARB as a weighted mismatch (WM) score based on the type of mismatch and the corresponding perfect match (Strunk 2001). This weighed approach essentially classifies mismatches into two groups, having either weak or strong destabilizing effects, but this qualitative classification is obviously oversimplified given the complications of mismatch effects described in the previous section. Yilmaz et al. (2008) compared ARB's WM approach with quantitative thermodynamic indicators of mismatched duplex stability and found a low correlation between the default WM scores and experimental observations. Based on these results, an updated set of WM scores were suggested specifically for FISH probe design using ARB. However, it is difficult to determine whether ARB users modify the WM parameters or commonly use the default values. To address sensitivity, ARB designs probes to reach a sufficiently high dissociation temperature predicted based on probe length and GC content. In addition, for each probe candidate, the likelihood of a successful hybridization is evaluated using accessibility maps of the target molecule, which show the observed accessibility of FISH probes to specific regions of ssu and lsu rRNA in model organisms (Behrens et al. 2003; Fuchs et al. 2001; Fuchs et al. 1998). Thus, both specificity and sensitivity are optimized by ARB based on empirical approaches rather than on thermodynamic principles.

A significant advantage of ARB is that it seamlessly integrates with SILVA, a database that offers validation tools for oligo design (Quast et al. 2013). For instance, SILVA offers a TestProbe tool that is based on the ARB software. The user can test a given probe against the SILVA database, while configuring the number of allowed mismatches to the probe, and also using weighted mismatches, similar to the WM scores in ARB. The output offers a visualization of target group coverage, as well as lists of sequences with mismatched hybridizations showing the number and position of mismatches. A similar tool is available from the RDP website, which allows the user to search for one or two probes for matching sequences in the RDP database. The ProbeMatch output shows the number and position of mismatches in each sequence in a taxonomic hierarchy. These tools are particularly useful for quickly finding potential nontargets in the most up-to-date rRNA databases.

An alternative design approach uses thermodynamics-based mathematical models that simulate DNA/RNA hybridizations in silico (Yilmaz et al. 2008; Yilmaz and Noguera 2004; Yilmaz and Noguera 2007; Yilmaz et al. 2011). These models are useful to evaluate the strength of matched and mismatched duplexes, thus providing a uniform approach to predicting probe sensitivity and specificity. In 2004, Yilmaz and Noguera (2004) described the first thermodynamics-based model of FISH (Fig. 2a). This model represented in situ hybridizations as the equilibrium between three different reactions, namely the probe-template hybridization (DNA/RNA), the self-folding of the probe (DNA/DNA), and the self-folding of the target molecule (RNA/RNA). This initial model helped explain observations of varying accessibility of probes across the length of the ssu rRNA (Behrens et al. 2003; Fuchs et al. 2001; Fuchs et al. 1998). Rather than supporting the hypothesis that probe accessibility was a function of factors external to the hybridization reactions, the thermodynamic model showed that it was possible to effectively target all sites of the ssu rRNA by designing a probe with adequate binding strength (Okten et al. 2012; Yilmaz et al. 2006). Notably, this discovery meant that probe design did not need to be limited to specific regions of the ssu rRNA. These findings also suggested that the thermodynamic concepts for probe design should be applicable when targeting other RNA molecules in situ (e.g., lsu rRNA, mRNA), as the level of the

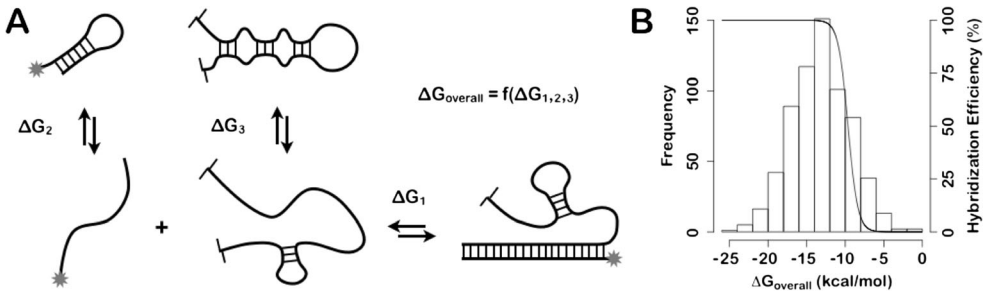

Fig. 2 **a** Mechanistic model of hybridization in FISH showing the three competing reactions of probe-folding, target-folding, and probe-target duplex formation. **b** Histogram of $\Delta G°_{overall}$ at 0 % formamide for the 658 probes used for FISH available from probeBase. $\Delta G°_{overall}$ was calculated for each probe against all the perfect match targets in the RDP using DECIPHER (Wright et al. 2014a). The median $\Delta G°_{overall}$ for each probe was used to make the histogram. Most of the probes have high predicted hybridization efficiency based on their $\Delta G°_{overall}$

complexity of molecular interactions (RNA-RNA and RNA-protein) that could cause inaccessibility in these targets is similar to or less than those of the ssu rRNA.

With the in silico FISH model (Fig. 2), an overall thermodynamic affinity term ($\Delta G°_{overall}$) can be calculated based on the free energy change of the individual reactions ($\Delta G°_1$, the free energy change associated with probe-target duplex formation; $\Delta G°_2$, the free energy change associated with probe self-folding; and $\Delta G°_3$, the free energy change associated with target self-folding). For hybridization temperatures and probe concentrations typically used in FISH (i.e., 46 °C and 0.25 to 1.0 μM probe concentration), Yilmaz and Noguera (2004) proposed that a threshold of $\Delta G°_{overall} < -13$ kcal/mole should be maintained to achieve efficient hybridization. Probes with higher (less negative) $\Delta G°_{overall}$ values would result in lower hybridization efficiencies, and therefore dimmer signals, because the target sites would be less saturated with probes at equilibrium (Fig. 2b). An analysis of designed FISH probes catalogued in probeBase (Loy et al. 2007) validated both this threshold and the free energy-hybridization efficiency relationship, as the mean $\Delta G°_{overall}$ for 658 probes targeting the 16S rRNA was −13.2 kcal/mol, with most probes falling between −8 and −18 kcal/mol (Fig. 2b).

Two important updates to the mechanistic model were later provided. Yilmaz and Noguera (2007) extended the model to include the effect of formamide, the denaturant typically used to control stringency of hybridization in FISH, on $\Delta G°_{overall}$. The extension was based on establishing the linear free energy relationship (Courtenay et al. 2000; Schellman 1978) between formamide and $\Delta G°$ values in each reaction. Later, Yilmaz et al. (2008) proposed an approach and metrics to calculate the effect of mismatches on the hybridization efficiency of probes in FISH. Specifically, the destabilizing effect of mismatches was evaluated with $\Delta\Delta G°_1$, the increase in the free energy change of reaction 1, the probe-target duplex formation, upon the insertion of mismatches (see above). Since a comprehensive dataset of internal loop parameters for DNA/RNA mismatches does not exist, Yilmaz et al. (2008) suggested that the effect of mismatches in DNA/RNA hybridizations could be estimated from the average of the destabilizing effects of similar mismatches in DNA/DNA and RNA/RNA duplexes. Combining this calculation with the overall hybridization scheme including the two other reactions and the linear model for the effect of formamide, they proposed ΔFA_m as a key metric to estimate the destabilizing effect of mismatches. ΔFA_m can be calculated as the difference in the formamide melting point (FA_m) of the perfectly matched probe and the probe with mismatches (Yilmaz et al. 2008), which is analogous to the offset between the target and nontarget curves in Fig. 1. A value of ΔFA_m greater than 20 % would suggest a sufficient separation to distinguish the perfect match from the mismatched duplex, and therefore, that it would be possible to optimize the experimental conditions to eliminate signal from the mismatched nontargets while maintaining high signal strength from the perfectly matched target organism. Yilmaz et al. (2008) also proposed $\Delta\Delta G°_1$ alone as an alternative metric to evaluate the effect of a mismatch. The advantage of this metric is that it is only dependent of the probe/target duplex, and therefore, calculation of probe and target folding is not necessary, thus providing a computationally efficient alternative to ΔFA_m.

All aspects of the Yilmaz and Noguera model (Yilmaz et al. 2008; Yilmaz and Noguera 2004; Yilmaz and Noguera 2007) were synthesized into mathFISH (Yilmaz et al. 2011), an online tool for the simulation of probe/target hybridizations that included the effects of formamide and mismatches. The web tool is useful to evaluate the thermodynamic affinity of probes designed based on sequence comparisons and empirical approaches (e.g., by ARB; see above) using potential target and nontarget RNA sequences and compare the probe-specific values with recommended thresholds for sensitivity ($\Delta G°_{overall} < -13$ kcal/mole), or specificity (e.g., $\Delta FA_m > 20$ %). Probe length may need to be readjusted, or a different target site used, if some of the thermodynamic criteria are not met. An important limitation of mathFISH is that it was designed to analyze only one probe and one target site. Thus, it can be used as a tool within a probe design process, but it is not a software that enables probe design.

Another online tool that can be used to evaluate designed probes is ProbeCheck (Loy et al. 2008), which integrates the

different aspects of probe design that have been developed throughout the years. The ProbeCheck website accepts user-defined probes and can query them against a user-selected database (e.g., RDP, SILVA, or Greengenes) using ARB's ProbeMatch tool. For specificity estimations, it uses the default WM parameters of ARB, allowing up to four mismatches to be present in the duplex. In addition, it calculates the free energy change of the matched and mismatched hybridizations using the two-state hybridization algorithm in UNAFold (Markham and Zuker 2008), which is based on free energy parameters in DNA/DNA hybridizations. Although this is conceptually similar to the $\Delta\Delta G^{\circ}_{1}$ term defined by Yilmaz et al. (2008), the use of DNA/DNA rules alone for DNA/RNA interactions in FISH may not always give accurate estimates since DNA/DNA and DNA/RNA interactions have significant thermodynamic differences (Turner 2000).

McIlroy et al. (2011) identified a particular case of mismatched hybridizations in FISH that were not adequately considered by either ARB or ProbeCheck. When the probe and the target form a mismatched duplex that contains a bulge, ARB and ProbeCheck interpret the duplex as having multiple mismatches instead of a single bulge. Therefore, during probe design, such potential hybridizations are not considered. McIlroy et al. (2011) demonstrated that such bulged mismatched hybridizations occur in practice, and using nearest neighbor data for RNA/RNA hybridizations with bulges (Znosko et al. 2002), they calculated the effect of the bulge (ΔG°_{bulge}) in the hybridization. They created the LOOPOUT program (McIlroy et al. 2011) to calculate ΔG°_{bulge} (also conceptually equivalent to $\Delta\Delta G^{\circ}_{1}$ defined by Yilmaz et al. (2008)) and suggested that a ΔG°_{bulge} of less than 4 kcal/mole is not sufficiently destabilizing to prevent a high chance of hybridization.

In the most recent update to the thermodynamic models of in situ hybridization, Wright et al. (2014a) have developed a comprehensive tool to design FISH probes. The Design Probes web tool, which is part of the DECIPHER (Wright 2012) software package, includes the algorithms of mathFISH to assess probe sensitivity, except that it allows users to design probes with a hybridization efficiency as low as 50 % (corresponds to $\Delta G^{\circ}_{overall} \sim -10$ kcal/mole). To address specificity, Design Probes simplifies the reaction scheme to having the probe-target hybridization as a single reaction but modifies the free energy (ΔG°_{1}) to take into account the effect of the competing self-folding reactions. This change increased the computational efficiency by three orders of magnitude without compromising accuracy, thus allowing a comprehensive search for all potential nontarget duplexes in large databases using thermodynamic principles (Wright et al. 2014a). The Design Probes tool accepts a user-defined set of target and nontarget sequences and finds the target sites that provide the highest coverage of the target group while allowing for degenerate probes. It then outputs the probes with the highest specificity based on ΔFA_m calculations for all potential nontargets, regardless of the number of mismatches present. Importantly, the presence of bulges in mismatched probes is taken into account in the calculation of ΔFA_m. Furthermore, Design Probes is the only software that allows the user to design dual probe sets by finding the best combination of two single probes to minimize nontarget cross-hybridizations. Since many single probes are not sufficient to discriminate targets from nontargets due to numerous potential false positives, a realistic approach for the accurate identification of some microbial targets with FISH is to design a complementary pair of optimized probes and hybridize them simultaneously (Wright et al. 2014a).

Primer design for PCR applications

Detection and quantification of microorganisms by PCR or qPCR requires the design of oligonucleotide primers that are specific to the organisms of interest, provide high coverage of the intended target templates, and allow the development of protocols with high sensitivity (Zhang and Fang 2006). Since primer design is a common task in many fields, several primer design programs that use thermodynamic approaches to evaluate primer/template hybridizations are readily available (Applied Biosystems 2004; Untergasser et al. 2012; Untergasser et al. 2007; Ye et al. 2012). The general criteria for PCR primer design is that both primers should have similar melting temperature, moderately balanced GC content, and avoid self-complementarity (Ye et al. 2012). Runs of a single base and di/tri-nucleotide repeats should be avoided as they may result in false priming due to entropic effects. Other major considerations are amplicon length, amplification efficiency, and avoiding cross-amplification of nontarget templates. Additional criteria are used when measuring transcript abundance after reverse transcription, to avoid amplification of genomic DNA (Arvidsson et al. 2008).

The most common procedure for detecting amplicons in qPCR is to use a dye that increases fluorescence in the presence of double-stranded DNA, such as SYBR Green (Gudnason et al. 2007; Wittwer et al. 1997). Alternatively, the amplification of templates can be monitored with a fluorescently labeled oligonucleotide probe that is designed to target the template region between forward and reverse primers. TaqMan probes will bind to amplified targets and generate a fluorescence response after being degraded by the 5′ nuclease action of *Taq* polymerase (Livak 1999). A modification of the TaqMan probe introduces a minor groove binder (Kutyavin et al. 2000), which is a tripeptide linked to the 3′-end of the TaqMan probe that folds back onto the probe and increases stability of the duplex. This increased stability offers the possibility of designing a shorter probe, thus improving specificity. Molecular beacons are another type of reporter molecule for qPCR, designed to form a hairpin structure that will only allow the probe

to fluoresce when bound to the template (Tyagi and Kramer 1996). The advantage of probe-based approaches to quantifying template abundances is that a third oligo allows for increased specificity in design, and different-colored dyes can be used to monitor the amplification of multiple targets in the same polymerase chain reaction (Baldwin et al. 2003; Weller et al. 2000).

One of the most widely used programs for primer design is Primer3 (Untergasser et al. 2012; Untergasser et al. 2007), which is available as a web tool (Untergasser et al. 2007), and has been incorporated into other primer design programs (Arvidsson et al. 2008; Ye et al. 2012). In the web interface, the user can specify a multitude of parameters, depending on the intended application. In addition, the program offers the possibility of designing an internal oligo along with a pair of primers, which is useful for protocols where reporter fluorescence comes from the internal probe. Primer3 only provides optimized primer pair designs for a single user-defined template and therefore does not assess the specificity of the primer pair. If qPCR is being used to determine transcript abundance for a single organism, then the specificity of the primer pair would need to be evaluated against similar regions within the genome. The software QuantPrime (Arvidsson et al. 2008) has been developed to address this issue, and it combines primer pair design using Primer3 with additional filters that remove design options with the potential to cause false positive amplifications, including the possibility of amplifying genomic DNA when measuring mRNA transcript abundance. However, the QuantPrime database only has sequences for a limited number of organisms.

The design of primers with high specificity when the template DNA contains homologous genes from other organisms is a challenging but necessary task in microbial ecology applications. To date, there is no unified strategy to achieve this goal, and researchers use different approaches that partially rely on using software to identify potential target sites that are conserved for the targeted organisms and have several mismatches to nontarget organisms. Then, time-consuming trial and error approaches are used to test the quality of different primer sets, often using samples containing the targeted organisms and controls without the targets. More recently, three different programs have been developed to more comprehensively address this challenge.

The Sequence-Specific Primer Design (SSPD) algorithm of PRIMEGENS3 (Kushwaha et al. 2011) uses a user-specified database of sequences in conjunction with Primer3 and Megablast (Morgulis et al. 2008) to identify potential primer pairs and then filter them to eliminate potential cross-hybridizations identified by sequence similarity. Likewise, Primer-BLAST (Ye et al. 2012) identifies candidate primer pairs using Primer3 and then uses a combination of local and global alignment algorithms to identify potential cross-hybridizations in reference databases available from NCBI. In this case, the specificity determination is based on a user-defined number of mismatches in the primer/template duplex and the number of mismatches in the 3′-end region.

The DECIPHER software (Wright 2012) uses its own primer design algorithm (Wright et al. 2014b), which is based on a model incorporating the thermodynamic effect of mismatches on hybridization and the effect of 3′-end mismatches on elongation. The authors define overall amplification efficiency as the product of hybridization and elongation efficiencies. In general, the hybridization efficiency of a perfect match duplex decreases as the annealing temperature is increased, which results in delayed amplification and a shallower amplification curve (Fig. 3a) since the hybridization efficiency remains lower throughout the amplification cycle. Mismatched primers will incur an initial decrease in hybridization efficiency (causing amplification delay), but as each cycle produces new perfectly matched templates, the overall slope of the amplification curve is not altered (Fig. 3b). Elongation efficiency is relatively independent of temperature and represents the fraction of mismatched templates that are elongated in each cycle of PCR. Lowering elongation efficiency causes further delays in amplification, but similar to central mismatches, it does not change the shape of the amplification curve (Fig. 3b). In this manner, a primer with a central mismatch will only incur an initial decrease in hybridization efficiency, whereas a primer with a mismatch located near the 3′-end will typically suffer both a decrease in hybridization efficiency and elongation efficiency. The DECIPHER software attempts to maximize amplification delays of cross-hybridizations by designing primers with 3′-end mismatches to nontarget sequences. Therefore, the combined hybridization and elongation efficiency approach allows the program to generate candidate primers for user-defined DNA sequences and then incorporate the effect of different mismatches located near the 3′-end of the primer to improve primer specificity. In this way, specificity is not only solely based on the number of mismatches or their location but also incorporates the observation that *Taq* polymerase responds differently to different types of mismatches located near the 3′-end of the primer (Ayyadevara et al. 2000; Stadhouders et al. 2010). The Design Primers tool of DECIPHER is accessible through a web interface, which allows the user to specify whether to use the effect of mismatches on elongation efficiency or only hybridization efficiency when determining primer specificity.

An approach known as *Taq*MAMA further harnesses the delay incurred by 3′-end mismatches by inducing an artificial mismatch at the penultimate position of the primer sequence (Cha et al. 1992). This approach takes advantage of the nonlinear effect of multiple mismatches located near the 3′-end on the extension rate of primers by *Taq* polymerase. The added mismatch often results in a small amplification delay for target sequences (1 mismatch) but causes a much greater delay for nontarget sequences (≥2 mismatches). This approach

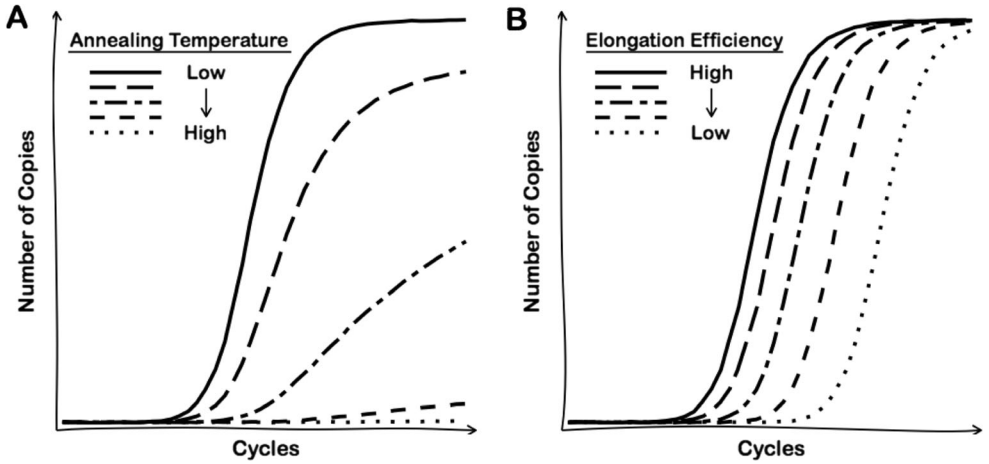

Fig. 3 Illustration of the effect of **a** increasing annealing temperature and **b** decreasing elongation efficiency on PCR amplification. **a** With perfect match primers, increasing the annealing temperature results in shallower slopes during PCR amplification as hybridization efficiency drops below 100 % and remains low throughout the amplification cycle. **b** In primers with mismatches, the unfavorable mismatches delay amplification but do not change the slope of the amplification curve. Although the initial templates have mismatches to the primer that decrease the hybridization efficiency, subsequent copies include the primer sequence and are therefore amplified with the same efficiency as the perfect match template. Unfavorable mismatches at the 3′-end cause increased delays in amplification as the probability of elongation decreases simultaneously with a decrease in hybridization efficiency

has been shown to allow discrimination of single nucleotide polymorphisms (SNPs) with PCR (Li et al. 2004). Wright et al. (2014b) adapted this approach by moving the mismatch further away from the 3′-end, which had the same effect on nontargets, but lessened the effect on target amplification. The DECIPHER software is also able to design primers with induced mismatches at the 3′-end, which may be useful in cases where adequate specificity cannot be obtained with standard perfect match primer sequences.

Probe design for phylogenetic microarray applications

The design of probes for phylogenetic microarrays shares many of the same principles as FISH probes and PCR primers, where the goal is to detect specific targets with high specificity and sensitivity. Phylogenetic microarrays have the added advantage that tens to hundreds of probes can be used to target a single organism, enabling new design and analysis strategies not available in other oligo applications in which specificity and sensitivity depend on the performance of only one to three oligos. Furthermore, as a high throughput platform, microarrays may contain oligos targeting many different groups, and therefore, minimizing false positive identifications also requires additional considerations, which can be implemented at the oligo design stage or in the post-hybridization processing of the data.

Another important difference between microarray applications and FISH or PCR is that the use of high-density microarrays is highly dependent on the manufacturer's overall strategies and instrumentation. For instance, the phyloChip (DeSantis et al. 2007), a microarray that uses 16S rRNA targeted oligos, is specifically designed for the Affymetrix platform and uses Affymetrix-specific approaches to differentiate true and false positive identifications. Until recently, NimbleGen manufactured high-density custom arrays with user-defined probes, a platform that has also been used to create phylogenetic chips with 16S rRNA oligos (Wright et al. 2013). The literature also has many examples of custom-designed phylogenetic microarrays that are manufactured by spotting or printing the oligos onto glass slides (Kelly et al. 2005; Liu et al. 2001; Loy et al. 2002; Loy et al. 2005).

Many programs have been developed for oligonucleotide probe design (Ilie et al. 2013; Li et al. 2005; Nielsen 2003; Nordberg 2005; Rouillard 2003), although not all of them have been used for the design of phylogenetic microarrays. A comprehensive review and analysis of different probe design programs have been recently performed (Lemoine et al. 2009). In their review, Lemoine et al. (2009) discuss the advantages and disadvantages of these programs, as well as the different approaches used to consider sequence identity and potential cross-hybridizations. The majority of the programs calculate sequence identities using a local alignment algorithm, such as BLAST (Nielsen 2003; Rouillard 2003), although drawbacks to this approach have been recently described (Ilie et al. 2013). Other programs use different algorithms for more accurate calculations of sequence identity, such as global alignments or multiple spaced seeds (Ilie et al. 2013; Li et al. 2005).

Oligonucleotide probes on phylogenetic microarrays are typically designed to have similar melting temperatures based on length and GC content. With spotted microarrays, where

targets may be readily available, experimental optimization of hybridization and wash conditions is performed to find the right balance of sensitivity and specificity. In some cases, the optimization relies on understanding the kinetics of denaturation of mismatched probes, which could be denatured under constant stringent conditions or with increasing stringency during the denaturation process (Liu et al. 2001; Urakawa et al. 2002). With high density microarrays possibly having thousands to millions of probes in the array, experimental optimization for each oligo is not an option, and therefore, approaches to design millions of probes with similar hybridization characteristics are desirable.

Interestingly, given the complexity of the problem, many attempts to model and predict the signal intensity from hybridized probes in phylogenetic microarrays concluded that hybridization characteristics (i.e., signal intensity) of probes designed for species discrimination could not be predicted (Pozhitkov et al. 2006), suggesting that the actual mechanism of probe-target hybridizations when the probes are attached to a surface is not well understood (Halperin et al. 2005; Peterson et al. 2001; Pozhitkov et al. 2007). Modeling approaches that were reported to be successful used the concept of adsorption isotherms to describe the relationship between target concentration and the fraction of target-bound probes. In particular, the Langmuir isotherm model, which assumes that saturation of the probes by the targets can be reached, is the most commonly used model to predict signal intensity from hybridized probes (Binder et al. 2005; Burden et al. 2004; Held et al. 2003; Hooyberghs et al. 2009).

In a departure from the Langmuir model and inspired by the approaches used in FISH for discrimination of targets and nontargets by increasing the stringency of the hybridization buffer (Fig. 1), Yilmaz et al. (2012) developed a linear free energy model (LFEM) of probe hybridization targeting the prediction of equilibrium denaturation profiles when the stringency of the hybridization buffer was increased. In their model, rather than attempting to predict signal intensity, the focus is on predicting the melting point of duplexes when formamide is used as the denaturant in the hybridization buffer. Using high-density NimbleGen custom microarrays, they obtained thousands of formamide curves for perfectly matched and mismatched probes and obtained basic thermodynamic parameters for nearest-neighbors and the different types of mismatches that could be expected when designing probes for differentiation of closely related phylogenetic groups (i.e., single mismatches, tandem mismatches, bulges, and the effect of position of the mismatch on the duplex), which then enable a prediction of potential cross-hybridizations for every probe in the microarray. All of these parameters are used in tools on the DECIPHER website to calculate the $\Delta G°$ for any probe/target duplex with or without mismatches (ProbeMelt tool) and to design multiple probes targeting a set of user-defined groups of sequences while minimizing nontarget hybridizations (Design Array tool).

The massively parallel nature of DNA microarrays enables more sophisticated analysis methods than possible with single oligos alone. For example, Wright et al. (2013) represented each probe with an equation, whereby the amount of probe signal was a combination of the hybridization efficiencies of each target with that probe. Hybridization of the entire array can therefore be represented by a system of equations representing the hybridization of every probe to each potential target. This system of equations can be solved using standard linear algebra methods, such as nonnegative least squares, to determine the contribution of each target to the observed hybridization signal. This approach resulted in far fewer false positives than simpler analysis approaches when analyzing the signals obtained from a custom DNA microarray (Wright et al. 2013). This approach is detailed in the vignette "Designing Group-Specific Microarray Probes" available with the DECIPHER package for the R programming language (Wright 2012).

Conclusions and future directions

The selective detection and quantification of specific organisms within a microbial community that contains closely related nontarget organisms is a prevalent challenge in microbial ecology research. FISH, qPCR, and microarrays are common approaches requiring specific detection of organisms of interest for accurate identification and quantification. In all of these cases, oligonucleotides that are phylogenetically specific to the target organisms need to be designed. The recent creation of thermodynamic-based models for the design of oligos for these applications has transformed the design process from an empirical approach requiring extensive trial and error to a more mechanistic and uniform design. In turn, the new models have accelerated the evolution of these applications to take advantage of continuously expanding databases of phylogenetic markers.

Critical in the assessment of oligo specificity is the ability to predict potential cross-hybridizations with mismatched templates that result in false positive identifications. All oligo design applications could continue to benefit from improvements in predictive parameters for mismatched hybridizations. Although current models are predictive enough to be useful in practice, it is our hope that new high-throughput techniques will offer a means of obtaining mismatched thermodynamic parameters in masse. In particular, there is room for improvement in predicting hybridization efficiencies of mismatched FISH probes, since DNA/RNA loop parameters are incomplete and FISH hybridizations employ formamide instead of temperature as a denaturant. Similarly, PCR applications may benefit from more systematic studies on DNA motifs for

which *in solution* free energy rules are still oversimplified, such as in tandem mismatches, and from more complete datasets to predict the effect that mismatches near the 3′-end of a primer have on elongation efficiency.

Furthermore, most of the modeling advances that were described in this review are based on hybridizations with DNA oligos and are therefore not necessarily applicable when using alternative oligos such as peptide nucleic acids (PNA), locked nucleic acids (LNA), or even RNA oligos for particular applications (Karkare and Bhatnagar 2006; Perry-O'Keefe et al. 2001; Rosenthal et al. 2013; Silahtaroglu et al. 2004).

Finally, as novel and more creative applications of nucleic acid probes and primers are developed, mathematical models that help in the process of oligo design will also need to evolve to embrace the new challenges. For instance, the recent description of the application of RNA probes for in situ hybridizations to mRNA (Rosenthal et al. 2013) using linear amplification of fluorescent signal by cascade reactions (Dirks and Pierce 2004), the use of forced intercalation probes in FISH (Bethge et al. 2008), or high resolution melt analyses for qPCR applications (Rasmussen et al. 2007; Schwartz et al. 2009) highlight the need to continue the quest for better predicting interactions of oligos with matched and mismatched templates. In turn, improvements in thermodynamic models and oligo design will continue to extend the boundaries of how oligos can be applied in the future.

Acknowledgments The authors would like to thank Alexander Steinbüchel for the invitation to write this review. This research was partially supported by a fellowship from Becas Chile (Conicyt) to Pamela Camejo.

Conflict of interest The authors declare that they have no conflict of interest.

References

Abell GCJ, Revill AT, Smith C, Bissett AP, Volkman JK, Robert SS (2010) Archaeal ammonia oxidizers and nirS-type denitrifiers dominate sediment nitrifying and denitrifying populations in a subtropical macrotidal estuary. ISME J 4(2):286–300. doi:10.1038/ismej.2009.105

Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol 56(6):1919–1925

Applied Biosystems (2004) Primer express software version 3.0. getting started guide

Arvidsson S, Kwasniewski M, Riano-Pachon DM, Mueller-Roeber B (2008) QuantPrime—a flexible tool for reliable high-throughput primer design for quantitative PCR. BMC Bioinforma 9. doi:10.1186/1471-2105-9-465

Ayyadevara S, Thaden JJ, Reis RJS (2000) Discrimination of primer 3′-nucleotide mismatch by Taq DNA polymerase during polymerase chain reaction. Anal Biochem 284(1):11–18. doi:10.1006/abio.2000.4635

Bakshi S, Siryaporn A, Goulian M, Weisshaar JC (2012) Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. Mol Microbiol 85(1):21–38. doi:10.1111/j.1365-2958.2012.08081.x

Baldwin BR, Nakatsu CH, Nies L (2003) Detection and enumeration of aromatic oxygenase genes by multiplex and real-time PCR. Appl Environ Microbiol 69(6):3350–3358. doi:10.1128/aem.69.6.3350-3358.2003

Behrens S, Ruhland C, Inacio J, Huber H, Fonseca A, Spencer-Martins I, Fuchs BM, Amann R (2003) In situ accessibility of small-subunit rRNA of members of the domains bacteria, archaea, and eucarya to Cy3-labeled oligonucleotide probes. Appl Environ Microbiol 69(3):1748–1758. doi:10.1128/aem.69.3.1748-1758.2003

Bethge L, Jarikote DV, Seitz O (2008) New cyanine dyes as base surrogates in PNA: forced intercalation probes (FIT-probes) for homogeneous SNP detection. Bioorg Med Chem 16(1):114–125. doi:10.1016/j.bmc.2006.12.044

Binder H, Preibisch S, Kirsten T (2005) Base pair interactions and hybridization isotherms of matched and mismatched oligonucleotide probes on microarrays. Langmuir 21(20):9287–302. doi:10.1021/la051231s

Burden CJ, Pittelkow YE, Wilson SR (2004) Statistical analysis of adsorption models for oligonucleotide microarrays. Stat Appl Genet Mol Biol 3:Article35 doi:10.2202/1544-6115.1095

Cha RS, Zarbl H, Keohavong P, Thilly WG (1992) Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. PCR Methods Appl 2(1):14–20

Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM, Bandela AM, Cardenas E, Garrity GM, Tiedje JM (2007) The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res 35:D169–D172. doi:10.1093/nar/gkl889

Cottrell MT, Kirchman DL (2000) Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. Appl Environ Microbiol 66(4):1692–1697. doi:10.1128/aem.66.4.1692-1697.2000

Courtenay ES, Capp MW, Saecker RM, Record MT (2000) Thermodynamic analysis of interactions between denaturants and protein surface exposed on unfolding: interpretation of urea and guanidinium chloride m-values and their correlation with changes in accessible surface area (ASA) using preferential interaction coefficients and the local-bulk domain model. Proteins Struct Funct Genet: 72–85

Daims H, Bruhl A, Amann R, Schleifer KH, Wagner M (1999) The domain-specific probe EUB338 is insufficient for the detection of all bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol 22(3):434–444

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72(7):5069–5072. doi:10.1128/aem.03006-05

DeSantis TZ, Brodie EL, Moberg JP, Zubieta IX, Piceno YM, Andersen GL (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microb Ecol 53(3):371–383. doi:10.1007/s00248-006-9134-9

Dirks RM, Pierce NA (2004) Triggered amplification by hybridization chain reaction. Proc Natl Acad Sci U S A 101(43):15275–15278. doi:10.1073/pnas.0407024101

Fazi S, Aulenta F, Majone M, Rossetti S (2008) Improved quantification of *Dehalococcoides* species by fluorescence in situ hybridization and catalyzed reporter deposition. Syst Appl Microbiol 31(1):62–67. doi:10.1016/j.syapm.2007.11.001

Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ (2008) Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol 74(8):2461–2470. doi:10.1128/aem. 02272-07

Fuchs BM, Wallner G, Beisker W, Schwippl I, Ludwig W, Amann R (1998) Flow cytometric analysis of the in situ accessibility of *Escherichia coli* 16S rRNA for fluorescently labeled oligonucleotide probes. Appl Environ Microbiol 64(12):4973–4982

Fuchs BM, Syutsubo K, Ludwig W, Amann R (2001) In situ accessibility of *Escherichia coli* 23S rRNA to fluorescently labeled oligonucleotide probes. Appl Environ Microbiol 67(2):961–968. doi:10.1128/AEM. 67.2.961-968.2001

Gudnason H, Dufva M, Bang DD, Wolff A (2007) Comparison of multiple DNA dyes for real-time PCR: effects of dye concentration and sequence composition on DNA amplification and melting temperature. Nucleic Acids Res 35(19):e127. doi:10.1093/nar/gkm671

Halperin A, Buhot A, Zhulina EB (2005) Brush effects on DNA chips: thermodynamics, kinetics, and design guidelines. Biophys J 89(2):796–811. doi:10.1529/biophysj.105.063479

Held GA, Grinstein G, Tu Y (2003) Modeling of DNA microarray data by using physical properties of hybridization. Proc Natl Acad Sci U S A 100(13):7575–80. doi:10.1073/pnas.0832500100

Hooyberghs J, Van Hummelen P, Carlon E (2009) The effects of mismatches on hybridization in DNA microarrays: determination of nearest neighbor parameters. Nucleic Acids Res 37(7):e53. doi:10.1093/nar/gkp109

Ilie L, Mohamadi H, Golding GB, Smyth WF (2013) BOND: basic oligonucleotide design. BMC Bioinforma 14(69):1–8. doi:10.1186/1471-2105-14-69

Jacobson H, Stockmayer WH (1950) Intramolecular reaction in polycondensations. I. The theory of linear systems. J Chem Phys 18:1600–1606

Jones CM, Stres B, Rosenquist M, Hallin S (2008) Phylogenetic analysis of nitrite, nitric oxide, and nitrous oxide respiratory enzymes reveal a complex evolutionary history for denitrification. Mol Biol Evol 25(9):1955–1966. doi:10.1093/molbev/msn146

Karkare S, Bhatnagar D (2006) Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino. Appl Microbiol Biotechnol 71(5):575–586. doi:10.1007/s00253-006-0434-2

Kelly JJ, Siripong S, McCormack J, Janus LR, Urakawa H, El Fantroussi S, Noble PA, Sappelsa L, Rittmann BE, Stahl DA (2005) DNA microarray detection of nitrifying bacterial 16S rRNA in wastewater treatment plant samples. Water Res 39(14):3229–3238. doi:10.1016/j.watres.2005.05.044

Klein M, Friedrich M, Roger AJ, Hugenholtz P, Fishbain S, Abicht H, Blackall LL, Stahl DA, Wagner M (2001) Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes. J Bacteriol 183(20):6028–6035. doi:10.1128/jb.183.20.6028-6035.2001

Korbie DJ, Mattick JS (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat Protoc 3(9):1452–1456. doi:10.1038/nprot.2008.133

Kushwaha G, Srivastava GP, Dong X PRIMEGENSw3 (2011) A web-based tool for high-throughput primer and probe design. In 2011 I.E. International Conference on Bioinformatics and Biomedicine (BIBM), 12–15 Nov. 2011 2011. p 345–351 doi:10.1109/bibm.2011.43

Kutyavin IV, Afonina IA, Mills A, Gorn VV, Lukhtanov EA, Belousov ES, Singer MJ, Walburger DK, Lokhov SG, Gall AA, Dempcy R, Reed MW, Meyer RB, Hedgpeth J (2000) 3′-minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures. Nucleic Acids Res 28(2):655–661. doi:10.1093/nar/28.2.655

Lemoine S, Combes F, Le Crom S (2009) An evaluation of custom microarray applications: the oligonucleotide design challenge. Nucleic Acids Res 37(6):1726–1739. doi:10.1093/nar/gkp053

Lenk S, Arnds J, Zerjatke K, Musat N, Amann R, Mussmann M (2011) Novel groups of Gammaproteobacteria catalyse sulfur oxidation and carbon fixation in a coastal, intertidal sediment. Environ Microbiol 13(3):758–774. doi:10.1111/j.1462-2920.2010.02380.x

Li BH, Kadura I, Fu DJ, Watson DE (2004) Genotyping with TaqMAMA. Genomics 83(2):311–320. doi:10.1016/j.ygeno.2003.08.005

Li X, He Z, Zhou J (2005) Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. Nucleic Acids Res 33(19):6114–23. doi:10.1093/nar/gki914

Liu WT, Mirzabekov AD, Stahl DA (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. Environ Microbiol 3(10):619–629

Livak KJ (1999) Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genet Anal-Biomol E 14(5–6):143–149. doi:10.1016/s1050-3862(98)00019-9

Livak KJ, Flood SJA, Marmaro J, Giusti W, Deetz K (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Meth Appl 4(6):357–362

Loy A, Lehner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer KH, Wagner M (2002) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Appl Environ Microbiol 68(10):5064–5081. doi:10.1128/AEM. 68.10.5064-5081.2002

Loy A, Schulz C, Lucker S, Schopfer-Wendels A, Stoecker K, Baranyi C, Lehner A, Wagner M (2005) 16S rRNA gene-based oligonucleotide microarray for environmental monitoring of the betaproteobacterial order "Rhodocyclales". Appl Environ Microbiol 71(3):1373–1386. doi:10.1128/AEM. 71.3.1373-1386.2005

Loy A, Maixner F, Wagner M, Horn M (2007) probeBase—an online resource for rRNA-targeted oligonucleotide probes: new features 2007. Nucleic Acids Res 35:D800–D804. doi:10.1093/nar/gkl856

Loy A, Arnold R, Tischler P, Rattei T, Wagner M, Horn M (2008) probeCheck—a central resource for evaluating oligonucleotide probe coverage and specificity. Environ Microbiol 10(10):2894–2898. doi:10.1111/j.1462-2920.2008.01706.x

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar BA, Lai T, Steppi S, Jobb G, Forster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, Konig A, Liss T, Lussmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer KH (2004) ARB: a software environment for sequence data. Nucleic Acids Res 32(4):1363–71. doi:10.1093/nar/gkh293

Manz W, Amann R, Ludwig W, Wagner M, Schleifer KH (1992) Phylogenetic oligodeoxynucleotide probes for the major subclasses of Proteobacteria—problems and solutions. Syst Appl Microbiol 15(4):593–600

Markham NR, Zuker M (2008) UNAFold: software for nucleic acid folding and hybridization. Methods Mol Biol 453:3–31. doi:10.1007/978-1-60327-429-6_1

Mathews D, Burkard M, Freier S, Wyatt J, Turner D (1999a) Predicting oligonucleotide affinity to nucleic acid targets. RNA 5:1458–1469. doi:10.1017/S1355838299991148

Mathews D, Sabina J, Zuker M, Turner D (1999b) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288:911–940. doi:10.1006/jmbi.1999.2700

McIlroy SJ, Tillett D, Petrovski S, Seviour RJ (2011) Non-target sites with single nucleotide insertions or deletions are frequently found in 16S rRNA sequences and can lead to false

positives in fluorescence in situ hybridization (FISH). Environ Microbiol 13(1):33–47. doi:10.1111/j.1462-2920.2010.02306.x

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schaffer AA (2008) Database indexing for production MegaBLAST searches. Bioinformatics 24(16):1757–1764. doi:10.1093/bioinformatics/btn322

Mosier AC, Francis CA (2008) Relative abundance and diversity of ammonia-oxidizing archaea and bacteria in the San Francisco Bay estuary. Environ Microbiol 10(11):3002–3016. doi:10.1111/j.1462-2920.2008.01764.x

Nielsen HB (2003) Design of oligonucleotides for microarrays and perspectives for design of multi-transcriptome arrays. Nucleic Acids Res 31(13):3491–3496. doi:10.1093/nar/gkg622

Nordberg EK (2005) YODA: selecting signature oligonucleotides. Bioinformatics 21(8):1365–70. doi:10.1093/bioinformatics/bti182

Okten HE, Yilmaz LS, Noguera DR (2012) Exploring the in situ accessibility of small subunit ribosomal RNA of members of the domains Bacteria and Eukarya to oligonucleotide probes. Syst Appl Microbiol 35(8):485–495. doi:10.1016/j.syapm.2011.11.001

Pernthaler A, Pernthaler J, Amann R (2002) Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. Appl Environ Microbiol 68(6):3094–3101. doi:10.1128/AEM. 68.6.3094-3101.2002

Perry-O'Keefe H, Rigby S, Oliveira K, Sorensen D, Slender H, Coull J, Hyldig-Nielsen JJ (2001) Identification of indicator microorganisms using a standardized PNA FISH method. J Microbiol Methods 47(3):281–292. doi:10.1016/s0167-7012(01)00303-7

Pester M, Rattei T, Flechl S, Grongroft A, Richter A, Overmann J, Reinhold-Hurek B, Loy A, Wagner M (2012) amoA-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic regions. Environ Microbiol 14(2):525–539. doi:10.1111/j.1462-2920.2011.02666.x

Peterson AW, Heaton RJ, Georgiadis RM (2001) The effect of surface probe density on DNA hybridization. Nucleic Acids Res 29(24):5163–5168. doi:10.1093/nar/29.24.5163

Pfeiffer S, Pastar M, Mitter B, Lippert K, Hackl E, Lojan P, Oswald A, Sessitsch A (2014) Improved group-specific primers based on the full SILVA 16S rRNA gene reference database. Environ Microbiol. doi:10.1111/1462-2920.12350

Pilhofer M, Pavlekovic M, Lee NM, Ludwig W, Schleifer KH (2009) Fluorescence in situ hybridization for intracellular localization of nifH mRNA. Syst Appl Microbiol 32(3):186–192. doi:10.1016/j.syapm.2008.12.007

Pozhitkov A, Noble PA, Domazet-Loso T, Nolte AW, Sonnenberg R, Staehler P, Beier M, Tautz D (2006) Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. Nucleic Acids Res 34(9): doi:10.1093/nar/gkl133

Pozhitkov AE, Tautz D, Noble PA (2007) Oligonucleotide microarrays: widely applied—poorly understood. Brief Funct Genomic Proteomic 6(2):141–8. doi:10.1093/bfgp/elm014

Pratscher J, Stichternoth C, Fichtl K, Schleifer KH, Braker G (2009) Application of recognition of individual genes-fluorescence in situ hybridization (RING-FISH) to detect nitrite reductase genes (nirK) of denitrifiers in pure cultures and environmental samples. Appl Environ Microbiol 75(3):802–810. doi:10.1128/aem. 01992-08

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41(D1):D590–D596. doi:10.1093/nar/gks1219

Rasmussen JP, Saint CP, Monis PT (2007) Use of DNA melting simulation software for in silico diagnostic assay design: targeting regions with complex melting curves and confirmation by real-time PCR using intercalating dyes. BMC Bioinforma 8. doi:10.1186/1471-2105-8-107

Rosenthal AZ, Zhang XN, Lucey KS, Ottesen EA, Trivedi V, Choi HMT, Pierce NA, Leadbetter JR (2013) Localizing transcripts to single cells suggests an important role of uncultured deltaproteobacteria in the termite gut hydrogen economy. Proc Natl Acad Sci U S A 110(40):16163–16168. doi:10.1073/pnas.1307876110

Rouillard JM (2003) OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res 31(12):3057–3062. doi:10.1093/nar/gkg426

SantaLucia J Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci U S A 95:1460–1465. doi:10.1073/pnas.95.4.1460

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

Schellman JA (1978) Solvent denaturation. Biopolymers 17:1305–1322

Schwartz SB, Thurman KA, Mitchell SL, Wolff BJ, Winchell JM (2009) Genotyping of *Mycoplasma pneumoniae* isolates using real-time PCR and high-resolution melt analysis. Clin Microbiol Infect 15(8):756–762. doi:10.1111/j.1469-0691.2009.02814.x

Silahtaroglu A, Pfundheller H, Koshkin A, Tommerup N, Kauppinen S (2004) LNA-modified oligonucleotides are highly efficient as FISH probes. Cytogenet Genome Res 107(1–2):32–37. doi:10.1159/000079569

Stadhouders R, Pas SD, Anber J, Voermans J, Mes THM, Schutten M (2010) The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5′ nuclease assay. J Mol Diagn 12(1):109–117. doi:10.2353/jmoldx.2010.090035

Stahl DA, Amann R (1991) Development and application of nucleic acid probes. In: Stackebrandt E, Goodfellow M (eds) Nucleic acid techniques in bacterial systematics. Wiley, Chichester, pp 205–248

Strunk O (2001) ARB: entwicklung eines programmsystems zur erfassung, Verwaltung und auswertung von nuklein- und aminosäuresequenzen. PhD Thesis, Technische Universität München

Sugimoto N, Nakano S, Katoh M, Matsumura A, Nakamuta H, Ohmichi T, Yoneyama M, Sasaki M (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. Biochemistry 34(35):11211–6. doi:10.1021/bi00035a029

Sugimoto N, Nakano M, Nakano S (2000) Thermodynamics-structure relationship of single mismatches in RNA/DNA duplexes. Biochemistry 39(37):11270–11281. doi:10.1021/bi000819p

Turner D (2000) Chapter 8: conformational changes. In: Bloomfield VA, Crothers DM, Tinoco IJ (eds) Nucleic acids: structures, properties, and functions. University Science, Sausalito

Turner DH, Mathews DH (2010) NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. Nucleic Acids Res 38:D280–D282. doi:10.1093/nar/gkp892

Tyagi S, Kramer FR (1996) Molecular beacons: probes that fluoresce upon hybridization. Nat Biotechnol 14(3):303–308. doi:10.1038/nbt0396-303

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35:W71–W74. doi:10.1093/nar/gkm306

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40(15):e115. doi:10.1093/nar/gks596

Urakawa H, Noble PA, El Fantroussi S, Kelly JJ, Stahl DA (2002) Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses. Appl Environ Microbiol 68(1):235–244. doi:10.1128/AEM. 68.1.235-244.2002

Wagner M, Haider S (2012) New trends in fluorescence in situ hybridization for identification and functional analyses of microbes. Curr Opin Biotechnol 23(1):96–102. doi:10.1016/j.copbio.2011.10.010

Watkins NE, Kennelly WJ, Tsay MJ, Tuin A, Swenson L, Lee HR, Morosyuk S, Hicks DA, SantaLucia J (2011) Thermodynamic contributions of single internal rA circle dot dA, rC circle dot dC, rG circle dot dG and rU circle dot dT mismatches in RNA/DNA duplexes. Nucleic Acids Res 39(5):1894–1902. doi:10.1093/nar/gkq905

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S Ribosomal DNA amplification for phylogenetic study. J Bacteriol 173(2):697–703

Weller SA, Elphinstone JG, Smith NC, Boonham N, Stead DE (2000) Detection of *Ralstonia solanacearum* strains with a quantitative, multiplex, real-time, fluorogenic PCR (TaqMan) assay. Appl Environ Microbiol 66(7):2853–2858. doi:10.1128/aem. 66.7.2853-2858.2000

Wilcox TM, Schwartz MK, McKelvey KS, Young MK, Lowe WH (2014) A blocking primer increases specificity in environmental DNA detection of bull trout (*Salvelinus confluentus*). Conserv Genet Resour 6(2):283–284. doi:10.1007/s12686-013-0113-4

Wittwer CT, Herrmann MG, Moss AA, Rasmussen RP (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. Biotechniques 22(1):130

Wright ES (2012) DECIPHER: database enabled code for ideal probe hybridization Employing R. R Package version 1.12.0. http://www.bioconductor.org/packages/2.12/bioc/html/DECIPHER.html

Wright ES, Strait JM, Yilmaz LS, Harrington GW, Noguera DR (2013) Identification of bacterial and archaeal communities from source to tap. Water Research Foundation, Denver

Wright ES, Yilmaz LS, Corcoran AM, Okten HE, Noguera DR (2014a) Automated design of probes for rRNA-targeted fluorescence in situ hybridization (FISH) reveals the advantages of dual probes for accurate identification. Appl Environ Microbiol 80(16):5124–5133. doi:10.1128/AEM. 01685-14

Wright ES, Yilmaz LS, Ram S, Gasser JM, Harrington GW, Noguera DR (2014b) Exploiting extension bias in polymerase chain reaction to improve primer specificity in ensembles of nearly identical DNA templates. Environ Microbiol 16(5):1354–1365. doi:10.1111/1462-2920.12259

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinforma 13. doi:10.1186/1471-2105-13-134

Yilmaz LS, Noguera DR (2004) Mechanistic approach to the problem of hybridization efficiency in fluorescent in situ hybridization. Appl Environ Microbiol 70(12):7126–7139. doi:10.1128/AEM. 70.12.7126-7139.2004

Yilmaz LS, Noguera DR (2007) Development of thermodynamic models for simulating probe dissociation profiles in fluorescence in situ hybridization. Biotechnol Bioeng 96(2):349–363. doi:10.1002/bit.21114

Yilmaz LS, Okten HE, Noguera DR (2006) Making all parts of the 16S rRNA of *Escherichia coli* accessible in situ to single DNA oligonucleotides. Appl Environ Microbiol 72(1):733–744. doi:10.1128/AEM. 72.1.733-744.2006

Yilmaz LS, Bergsven LI, Noguera DR (2008) Systematic evaluation of single mismatch stability predictors for fluorescence in situ hybridization. Environ Microbiol 10(10):2872–2885. doi:10.1111/j.1462-2920.2008.01719.x

Yilmaz LS, Parnerkar S, Noguera DR (2011) mathFISH, a web tool that uses thermodynamics-based mathematical models for in silico evaluation of oligonucleotide probes for fluorescence in situ hybridization. Appl Environ Microbiol 77(3):1118–1122. doi:10.1128/aem.01733-10

Yilmaz LS, Loy A, Wright ES, Wagner M, Noguera DR (2012) Modeling formamide denaturation of probe-target hybrids for improved microarray probe design in microbial diagnostics. PLoS ONE 7(8):e43862. doi:10.1371/journal.pone.0043862

Zhang T, Fang HHP (2006) Applications of real-time polymerase chain reaction for quantification of microorganisms in environmental samples. Appl Microbiol Biotechnol 70(3):281–289. doi:10.1007/s00253-006-0333-6

Znosko BM, Silvestri SB, Volkman H, Boswell B, Serra MJ (2002) Thermodynamic parameters for an expanded nearest-neighbor model for the formation of RNA duplexes with single nucleotide bulges. Biochemistry 41(33):10406–10417. doi:10.1021/bi025781q

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13):3406–15. doi:10.1093/nar/gkg595

Zwirglmaler K, Ludwig W, Schleifer KH (2003) Improved fluorescence in situ hybridization of individual microbial cells using polynucleotide probes: the network hypothesis. Syst Appl Microbiol 26(3):327–337. doi:10.1078/072320203322497356