Molecular BioSystems

Cite this: *Mol. BioSyst.*, 2012, **8**, 1325–1338

www.rsc.org/molecularbiosystems

PAPER

Analysis of crucial factors resulting in microarray hybridization failure

Ting Wei,*[ab] Mike N. Pearson,[a] Karen Armstrong,[c] Dietmar Blohm[d] and Jue Liu*[e]

Received 19th July 2011, Accepted 12th January 2012
DOI: 10.1039/c2mb05300d

The factors that affect the formation and stability of DNA/DNA duplexes are complicated and still mostly unknown. In this study attempts were made to look for the crucial factor affecting hybridization failure in DNA microarray assays. A comprehensive range of factors were investigated simultaneously using a 25-mer oligonucleotide *Potyvirus* microarray. These included steric hindrance, direct/indirect labelling types, distance of a probe to the fluorescent labelling end, target (the DNA fragment used to hybridize with microarray probes) strand types either single strand or double strand, probes without mismatch and with different numbers of mismatch nucleotides (up to 36%) and different mismatch locations (5′ end, centre and 3′ end), probe GC content and T_m, secondary structures of probes and targets, different target lengths (0.277 kb to ~1.3 kb) and concentrations (0.1–30 nM). The results showed that whilst most of these known factors were unlikely to be the main causes of failed hybridization, there was strong evidence suggesting that the viral amplicon target structure is the most crucial factor. However, computing predicted target secondary structures by Mfold showed no correlation with the hybridization results. One explanation is that the predicted target secondary structures are different from the real structures. Here we postulate that the real target structure might be a combination of secondary structures resulting in a three-dimensional structure from exposure to three types of sub-structures: (1) a completely exposed linear structure to allow probes access for the successful hybridization and showing strong fluorescent signals; (2) a partially exposed structure to allow unstable binding and showing weak fluorescent signals; (3) a closed structure resulting in failed hybridization. These results are very important for microarray based studies as they not only provide an explanation for some current controversial results, but also provide potential resolution for the future studies. Due to the lack of available software for predicting the true target structure, development of microarrays should conduct an initial oligonucleotide probe selection procedure and those probes with capacity to hybridize with the target should be considered for the microarray development.

Introduction

Microarray technology is becoming a very powerful analytical tool in genomic and biomedical science research, including clinical diagnosis for humans, animals and plants,[1–3] functional genomics,[4–6] pharmaceutical industry,[7,8] food contamination[9] and microbial ecology.[10,11] However, non-specific hybridization[12–14] and failed hybridization, even between perfect match probes and targets,[15–17] are problems that limit wider adoption of this technology.

Research over the past decades has recognized many factors affecting microarray hybridization efficiency and appropriate solutions for some factors have been studied. These have included (1) steric hindrance caused by the high density of immobilized probes[18,19] or interference of the solid support,[19,20] resulting in lower hybridization efficiency that could be mitigated or removed by printing probes at optimal density[21] and addition of spacers to increase the distance between the probes and glass support;[22–24] (2) identical or similar T_m for all oligonucleotide probes is important for the stability of duplexes during microarray hybridization;[25,26] (3) different labeling methods, including direct and indirect labeling, affect microarray efficiency;[1,27,28] (4) the distance of a probe to the fluorescent labelled end significantly affects signal intensity;[29] (5) exclusion of probes with 70–80% or above sequence

[a] *School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142, New Zealand. Fax: +64 9373 7416; Tel: +64 9373 7599 ext. 88371*
[b] *HIV Molecular Virology, Division of Immunology and Infectious disease, Queensland Institute of Medical Research, Brisbane, Queensland 4006, Australia*
[c] *Bio-Protection Research Centre, Lincoln University, Canterbury, New Zealand*
[d] *Centre of Applied GenSensor Technology, University of Bremen, FB2-UFT, 283559 Bremen, Germany*
[e] *Institute of Animal Husbandry and Veterinary Medicine, Beijing Municipal Academy of Agriculture and Forestry Sciences, Beijing 100097, China. E-mail: liujue@263.net*

similarity to non-target sequences which might reduce non-specific hybridization;[12–14,30] (6) single stranded DNA (ssDNA) fragments provide better hybridization efficiency than double stranded DNA (dsDNA) fragments;[31] (7) the secondary structure of longer targets, which are more complex than those of shorter targets, can affect the accessibility of probes to the targets to form duplexes[31–33]and a target size of 0.250 to 0.8 kb is considered to be optimal;[34] (8) the intra- and inter-molecular structure of oligonucleotide probes can compete with probe–target duplex formation resulting in a low hybridization efficiency;[35] (9) the middle position of a probe is important for stabilizing the duplex formation,[36] therefore a probe with most of its GC content in the middle position should bind more strongly to the target than a probe with high GC content at other locations,[19,20,25] and probes with higher GC content can provide higher signal intensities than those with lower GC content.[10]

However, some controversial results on factors relevant to microarray hybridization efficiency have also been reported. For example, it is generally considered that the type of nucleotide (*i.e.* A, C, G or T) and position of the miss-match on a probe can affect the stability of the DNA duplexes, especially at the centre of a probe,[37–39] but other studies have shown that a single mismatch with different probe nucleotide types did not have a destabilizing effect[16,17,40,41] and could even yield higher hybridisation signals than those of the corresponding perfect-match probes.[42] In addition, longer hybridization periods (10 to 19 hours) are generally considered to ensure the complete hybridization and thus increase the microarray efficiency,[21,43–46] but elsewhere a shorter period of 1–2 hours has been reported to give higher signals than a 16 hour hybridization.[47]

In this study, a range of factors affecting microarray hybridization efficiency were investigated to determine which were the crucial factors leading to hybridization failure and therefore inconsistent fluorescent signals. These factors were assessed using a 25-mer oligonucleotide *Potyvirus* microarray including four different viral species from New Zealand: *Dasheen mosaic virus* (DsMV), *Leek yellow stripe virus* (LYSV), *Potato virus Y* (PVY) and *Zucchini yellow mosaic virus* (ZYMV). The microarray consisted of 33 perfect-match (PM) and 52 mismatch (MM) viral probes with up to 9 (36%) MM nucleotides located at different positions of the probes, five human origin negative control probes and two positive control probes. Except the positive control probes, each viral and negative control probe was designed either without any extra spacer linker or with one of four different spacer linkers, either poly-cytosine (C) or poly-thymine (T). The different target (PCR product) sizes ranged from 0.277 kb to ~1.3 kb with direct or indirect fluorescent labelling.

Nomenclature of each viral perfect-match probe was the combination of viral species name + an Arabic number (representing the order of the probe within a species), mismatch probe as viral species + an Arabic number (represents the order of the probe within a species) + m (represents mismatch) + an Arabic number (represents the number of mismatch nucleotides). For each species, an Arabic number was given to each probe as the probe identification number, this would be helpful to present data in figures.

Results

Impact of target size, labelling type and target origin

Microarray hybridization was first carried out using individual dsDNA PCR products (targets) of different sizes (~0.335 kb, ~0.7 kb, and ~1.3 kb) for each of the four New Zealand potyviruses, and each with two different labeling types: (1) direct (Cy5 dye labeled reverse primers), and (2) indirect labeling (one type of nucleotide on the target was modified with a cyanine fluorescent dye). The results showed that only one out of three targets from each of the four *Potyvirus* isolates (a ~1.3 kb fragment from DsMV [D13], a ~1.0 kb fragment from LYSV [L10], a ~0.8 kb fragment from PVY [P8] and a ~0.7 kb fragment from ZYMV [Z7]) gave a positive reaction with three to ten species-specific probes on each single target regardless of the labeling type (Fig. 1). None of the short fragments (~0.335 kb) from each of the four viral species gave positive signals (data not shown). The number of probes showing positive signals from the four targets, D13, L10, P8 and Z7, were 11 out of 16, 8 out of 11, 3 out of 3 and 6 out of 9, respectively (Table 1), and the positive signals for the different probes showed different intensities (Fig. 1). The direct and indirect labeling provided nearly the same positive patterns except that the indirect labeling provided one extra probe showing a weak positive signal to the D13 fragment (Fig. 2) and a higher fluorescent intensity signal to most of the probes with positive signals (Fig. 2) except those with pixel saturation (in white color rather than red). Therefore direct labeling was selected for the rest of the experiment due to the convenience and low cost. The probes showing positive signals were consistently showing positive signals under different spacer modifications. When the four fragments containing probes with positive signals were hybridized to the array in a mixture, all the obtained probes showing positive signals were exactly the same as those obtained from the individual fragment hybridization (Fig. 1). Non-specific positive hybridization was not observed.

The ten species-specific shorter fragments (one for DsMV and three for each of LYSV, PVY and ZYMV) of different sizes (0.277 to 0.736 kb) that covered the regions containing probes showing negative signals (nine out of ten targets) or containing probes with both negative and positive signals (one out of ten targets) were generated by PCR and labelled indirectly with fluorescent dye. The hybridization results showed that, with the exception of two fragments (0.285 kb and 0.367 kb) that provided the same negative results as those from hybridization using longer fragments previously, all other fragments gave one to five extra probe(s) with positive signal(s) (Table 2).

Based on the previous hybridization results, two targets containing probes showing positive and negative signals (P8 and Z7) and two targets containing probes all showing negative signals (P13 and Z13) were selected for ssDNA isolation. Four ssDNA targets labeled directly with Cy5 fluorescent dyes were isolated and then hybridized with arrays. All of these targets provided the same positive and negative patterns as those of the direct-labeled dsDNA targets obtained from previous hybridization: the two targets P13 and Z13 which showed previously negative signals to all the probes included using dsDNA for the hybridization still showed

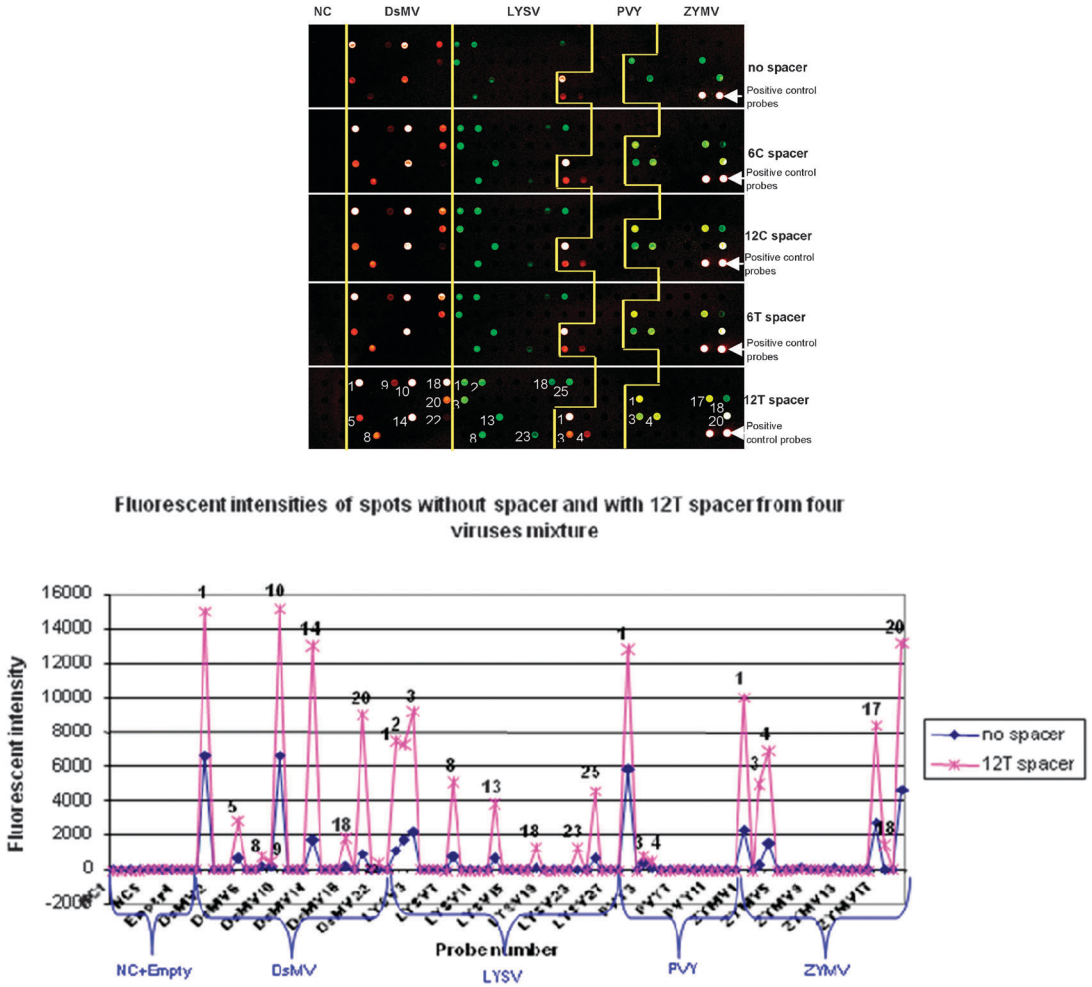

Fig. 1 Results from hybridization using a mixture containing four dsDNA targets with probes showing both positive and negative signals from four potyvirus species. Top image showed the negative and positive fluorescent probes and bottom figure indicated the fluorescent intensity value of probes without a spacer and with the 12T spacer. The four targets used for the hybridization included: D13 from DsMV labeled directly with Cy5 dye, L10 from LYSV labeled directly with Cy3 dye, P8 from PVY labeled directly with Cy5 dye and Z7 from ZYMV labeled directly with Cy3. The five columns from left to right in the top image represent negative control (NC) probes area and the four viral (DsMV, LYSV, PVY and ZYMV) probes areas. The five rows from top to bottom represent the probes with exactly same sequences but without any spacer modification (no spacer) and with four different spacer modifications of 6C, 12C, 6T and 12T spacers. The numbers to the left or lower-left side of the positive spots in the top image represent probe identification numbers (see Table 1) of relevant virus which corresponded to those in the bottom figure. Fluorescent intensity values for probes with other spacer modifications are similar to those of the 12T spacer which are omitted to avoid the mess peaks created by five sets data.

negative results when using ssDNA for the hybridization; two targets P8 and Z7 which showed previously both positive and negative signals to the probes included using dsDNA for the hybridization provided the same probes with positive signals when using ssDNA as targets (Fig. 3). Therefore ssDNA were not considered for the remaining experiments for convenience and cost effective purposes.

All the targets from each of the four phylogenetically closely-related potyviruses and the four healthy plants provided negative results (data not shown).

Impact of the spacers and distance of a probe to the fluorescent labelled end

It was observed from the image results that the same probe showing positive signals with spacers (four rows with 6C, 12C, 6T and 12T spacers in Fig. 1 image) appeared brighter than those without spacers (first row from top in Fig. 1 image), except where there was pixel saturation. As expected, the associated fluorescent intensity value for each positive probe with or without different spacer modifications corresponded with the image results (Fig. 1). In general, probes showing positive signals with any spacer type gave higher fluorescent intensities than those without spacers, except those with pixel saturation. Probes with 12 nt spacers showed higher fluorescent intensity than with 6 nt spacers in most cases, although there were no significant differences between probes without spacers and with different spacer treatments ($P > 0.05$), the absolute fluorescent intensity values of probes with spacers were 1.05–27.29 fold higher than those without spacers (data not shown). At the same time, the distance of each probe with a positive signal to the fluorescent labelled end was also investigated. The probes shown in Fig. 4 from the left to

Table 1 The details of probes with positive signals

No*[a]	Probe Name*[b]	Probe identification No*[c]	Probe Sequence*[d]
1	DsMV1^{++++}	1	TGAGCAGATGCACATCGTAATGAAT
2	DsMV5^{++++}	5	ACCACAGGAGGAGGTAATAATACCA
3	DsMV8^{++++}	8	CACAGATGAGTCTGCACAAAAGAGC
4	DsMV9^{+}	9	AAAGCATTGGCCTCTGAAGGAAGAG
5=1	DsMV10m1^{++++}	10	**C**GAGCAGATGCACATCGTAATGAAT
6=1	DsMV14m2^{++++}	14	**C**GAGCAGATGCACAT**T**GTAATGAAT
7	DsMV18m2^{++++}	18	CAGCGGGACATCTTTAA**T**AC**C**AGAG
8=1	DsMV20m3^{++++}	20	**C**GA**T**CAGATGCACAT**T**GTAATGAAT
9=5	DsMV22m4^{+}	22	**C**GATGATACAGTTGATGCAGGA**AAT**
1	LYSV1^{++++}	1	AACCCGTAGTATCCTATCCTTACCT
2	LYSV2^{++++}	2	AAGTTGGAACCCGTAGTATCCTATC
3	LYSV3^{++++}	3	AGTTGGAACCCGTAGTATCCTATCC
4	LYSV8^{++++}	8	TTCTGATACCCCACCTTTACCCTAC
5=2	LYSV13m2^{++++}	13	AA**A**CTGGAACCCGTAGTATCCTATC
6=2	LYSV18m3^{++}	18	AA**T**T**G**G**A**AACCCGTAGTATCCTATC
7=2	LYSV23m4^{++}	23	A**C**T**T**G**G**AAACCCGTAGTATCCTATC
8=12	LYSV25m4^{++++}	25	**TTT**G**G**AAACCCGTAGTATCCTATCC
1	PVY1^{++++}	1	GTTGCGATTCTGTCGTAGCAGTGAC
2	PVY3^{++++}	3	CATTGAAATCAGCCCAACCTCGACT
3	PVY4^{+}	4	GAAATCAGCCCAACCTCGACTTTTC
1	ZYMV1^{++++}	1	CGTAGTCCTGTCGGAAGCTTTAGTG
2	ZYMV3^{++++}	3	GCCTCTCACGAATAAGCTCGAGATT
3	ZYMV4^{++++}	4	TAGGTCGCCTACCTAGGTTATCGAT
4=3	ZYMV17m2^{++++}	17	GCCTCTCACG**G**ATAAGCTCGAG-TTA
5	ZYMV18m2^{++}	18	G**T**TTCGCT-CCGACGTAATTCTAATA
6	ZYMV20m1^{++++}	20	TTATGTTAGTTGTCCAG-AGTGCCGT

[a] This column shows the number of probes with positive signals within each viral species and "=" indicates that two probes were the same but with different mismatch nucleotide(s). Total probes for DsMV, LYSV, PVY and ZYMV are 16, 11, 3 and 9, respectively. [b] The probe name was constructed with viral name + Arabic number representing probe order number + an "m" (represents mismatch) followed by an Arabic number represents the number of mismatch nucleotides on the probe. The symbols "++++, +++, ++ and +" represent the fluorescent strength from high to low. [c] "Probe identification No" means the probe order number within each viral species. [d] The letters in bold and red colour represent mismatched nucleotides.

the right side of the *x* axis correspond to their position on the target fragments from the 5′ terminal (the fluorescent labeled end in the direct labeling system) to the 3′ terminal. The location of probes with positive signals, either close-to or away-from the fluorescent labelled end in direct-labeling type (Fig. 4 and Table 3), did not show a consistent relationship with fluorescent intensity. However in most cases, probes close to the fluorescent labelled end showed higher fluorescent intensity than those away from the fluorescent labeling end (first two probes on the DsMV 1.3 kb fragment and the LYSV 1.0 kb fragment, first probe on PVY in Fig. 4). There were still some exceptions. For example, the perfect-match probes ZYMV4 and ZYMV1, and the mis-match probe ZYMV20m1 which are 230 nt, 89 nt and 112 nt away from the fluorescent labelled end, respectively, showed higher fluorescent intensity than the perfect-match probe ZYMV3 which is closer to the fluorescent labelled end (61 nt away) (Fig. 4 and Table 3).

Impact of different target concentration and hybridization period

The target Z7, which gave six out of nine probes with positive signals at a target concentration of 10 nM, was selected to evaluate the impact of target concentration and hybridization period on hybridization efficiency. Results showed that increasing the target concentration up to 30 nM did not provide any additional probe with positive signals but it did increase the signal intensity for all the positive probes (Fig. 5). The signal intensity of the probes with positive signals gradually diminished or disappeared with the reduction of the target concentration, especially for the probes without spacers, but even at the lowest concentration (0.1 nM), there were still three probes showing positive signals (Fig. 5). In addition, hybridization for 2 h or overnight showed the same probes with positive or negative signals (data not shown).

Impact of nucleotide match percentage and the location of mismatch nucleotide

Among the 85 viral probes, 26 (30.6%) gave positive signals including 14 PM (out of 33, 42.4%) and 12 MM (out of 52, 23.1%) probes. However the positive signal strength did not correspond with the nucleotide match percentage (Fig. 4 and Table 1). In some instances mismatch probes showed similar or higher fluorescent intensities than their perfect matched

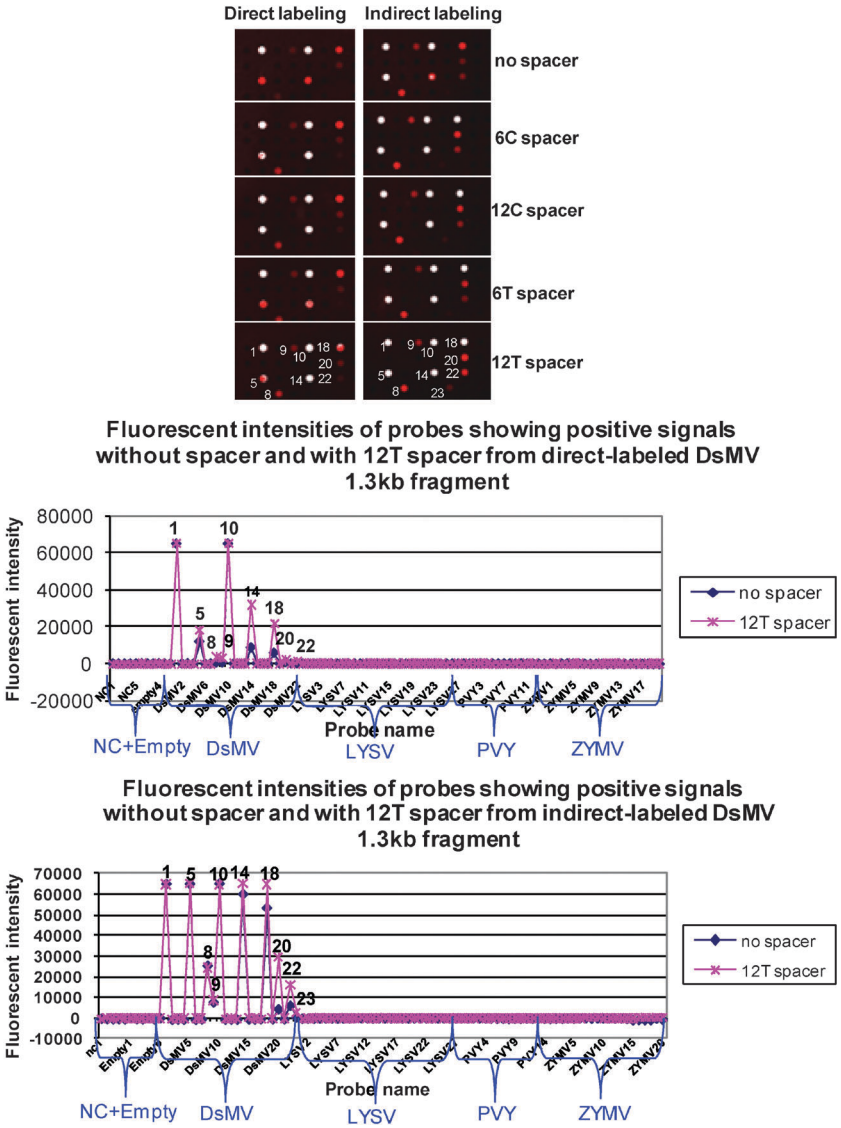

Fig. 2 Results from hybridization of the DsMV ~1.3 kb fragment with direct and indirect labeling. Top images showed the negative and positive fluorescent probes and bottom figures indicated the fluorescent intensity value of probes without a spacer and with the 12T spacer. Top-left image represents the hybridization using direct labelling target and top-right image using indirect labelling target. The numbers in the images to the left or lower-left side of the positive spots represent relevant DsMV probe identification numbers (see Table 1) which also correspond to those in bottom figures (not all probe names can be shown at the *x*-axis due to the large amount of probes included). Fluorescent intensity values for probes with other spacer modifications are similar to those of the 12T spacer which are omitted to avoid the mess peaks created by five sets data.

probes, such as DsMV1 and DsMV10m1, ZYMV3 and ZYMV17m2, respectively (Fig. 4 and Table 1). In addition, even mismatch nucleotides at the centre of the probe (ZYMV17m2) did not affect the fluorescent signal strength (Fig. 4 and Table 1).

Impact of probe thermodynamic properties on hybridization efficiency

The relationship between hybridization results and the GC content and T_m of all 85 probe sequences was analyzed using Chart Wizard in Microsoft EXCEL. Fluorescent intensities of probes without a spacer and with the 12T spacer were selected as examples for the comparison. The results showed that the positive and negative probes were distributed randomly within the GC content range of 36–52% and the T_m range of ~70 °C to 76 °C. There was no obvious correlation between GC content/T_m and the hybridization efficiency (Fig. 6). ΔG values of self-annealing and looping of each of the 85 probes without spacer and with 12T spacer modification were used as examples to analyze the relationship between hybridization efficiency and probe secondary structures. The more negative the ΔG value, the greater the energy in the self-annealing or looping configuration, resulting in a more stable secondary structure for a probe. There was no consistent relationship between hybridization success and the self-annealing or looping of probes. Probes showing both positive and negative signals were found distributed randomly within the self-annealing ΔG value range of "−11 to +2" and the self-looping ΔG value range of "−5 to +2", where the value of +2 represents non self-annealing and non self-looping (Fig. 7).

Table 2 Hybridization results using PCR fragments (dsDNA) amplified by specific primers

Virus	Fragment & size	Probe identification number*[a]	Results after hybridization*[b]
DsMV	Ds1, 0.457 kb	4, **8**$^{+++}$, **9**$^{+}$,17m2, 19m3	**4**$^{+}$, **8**$^{++++}$, **9**$^{++++}$,17m2, 19m3
LYSV	Ls1, 0.308 kb	11m1, 19m3, 24m4	**11m1**$^{+}$, 19m3, 24m4
LYSV	Ls2, 0.736 kb	11m1, 19m3, 24m4, 15m2, 6, 12m2, 17m2, 20m3	**11m1**$^{++}$, 19m3, 24m4, 15m2, **6**$^{++++}$, **12m2**$^{++++}$, **17m2**$^{++}$, **20m3**$^{++++}$
LYSV	Ls3, 0.4 kb	6, 12m2, 15m2, 17m2, 20m3	**6**$^{++++}$, **12m2**$^{+}$, 15m2, 17m2, **20m3**$^{++}$
PVY	Ps1, 0.285 kb	7m2, 8m2, 12m3	7m2, 8m2, 12m3
PVY	Ps2, 0.436 kb	2, 14m4, 11m3, 10m3, 7m2, 8m2, 12m3	**2**$^{++++}$, **14m4**$^{+}$, 11m3, 10m3, 7m2, 8m2, 12m3
PVY	Ps3, 0.277 kb	5, 6, 9m2	**5**$^{+++}$, **6**$^{++++}$, **9m2**$^{++}$
ZYMV	Zs1, 0.446 kb	2, 5, 6, 9m2, 15m3	**2**$^{++}$, **5**$^{+}$, **6**$^{+++}$, **9m2**$^{++}$, 15m3
ZYMV	Zs2, 0.392 kb	7, 8, 10m2, 12m3	**7**$^{++}$, **8**$^{++}$, 10m2, 12m3
ZYMV	Zs3, 0.367 kb	13m3, 14m3, 16m3, 19m6	13m3, 14m3, 16m3, 19m6

[a] "m" represents mismatch oligonucleotide, the Arabic number before an "m" represents probe identification number and after an "m" represents the number of mismatch nucleotides on the probe. Probe identification number means the order number of probe in each viral species. [b] This column indicated which probe has changed the signal status after hybridization. The Arabic number in bold and red color represents the relevant probe with positive signals from previous hybridization using dsDNA amplified by universal primers or from this hybridization using dsDNA amplified by specific primers. "++++, +++, ++, +" represent different visual fluorescent signal strengths from high to low.

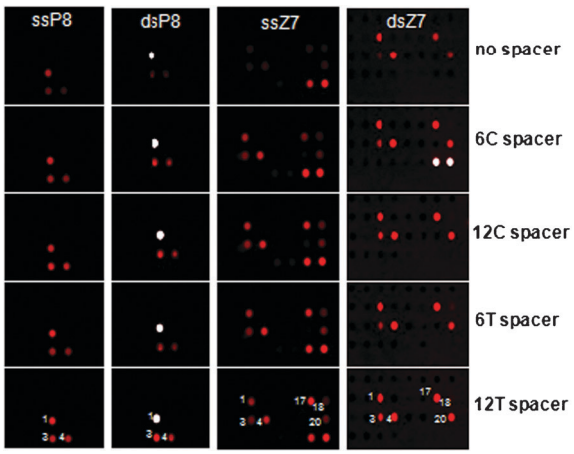

Fig. 3 Comparison of hybridization using ssDNA and dsDNA targets of P8 and Z7. The Arabic numbers at the left side of the positive spots are the relevant probe identification numbers of target P8 and Z7 (see Table 1) and those without an Arabic number are positive control spots (there are five sets of positive control spots on the ssZ7 image and only one set on the dsZ7 image).

Impact of target secondary structure on hybridization efficiency

All the targets from all four potyviruses, including those with positive and negative signals, were used for investigating the relationship between hybridization efficiency and target secondary structure. Secondary structure of each target fragment was generated at three different temperatures, 20, 37 and 55 °C, as some fragments may form secondary structures while handling the mixture on the slide at room temperature (~20 °C) before hybridization at 55 °C. The most stable structure from the Mfold output file at each temperature was selected from the software generated results for the analysis. Hybridization was carried out using probes with the 12T spacer and indirect labelling as this modification provided the best hybridization efficiency with most of the positive probes. The secondary structure results showed that the secondary structures of targets at lower temperatures (20 to 37 °C) are much more complex than those at higher temperature (55 °C), while at all investigated temperatures the secondary structures of longer fragments are more complex than those of shorter fragments. Analysis showed that even for the relatively simple structures at 55 °C, hybridization efficiency of different probes on different sites of the fragments showed no obvious correlation to the secondary structures of the targets. This was supported by evidence from two overlapping targets (DsMV ~1.3 kb and DsMV 0.457 kb) with far different structure complexity but providing probes with both positive and negative signals with different intensity values (Fig. 8). In the two selected target structures shown in Fig. 8, positive probes could be seen at complex structure sites while probes showing negative signals could be found at open linear sites. These results were also observed on other longer or shorter targets (data not shown). Upon hybridization at lower temperatures, probes with positive signals could be observed at much more complicated structure sites which was demonstrated by the ~0.7 kb ZYMV target Z7 (Fig. 9).

Discussion

Factors affecting microarray hybridization efficiency are numerous, and their relative importance remains largely unclear. In this study, some factors with the potential to affect microarray hybridization efficiency were mitigated or excluded by *in silico* design of the experiments. Firstly, commonly recognised factors including T_m (~72 ± 2 °C), GC content (40 to 60%) and length of the probes (25-mer) were taken into account. Secondly, to minimize non-specific hybridization, the initial probes generated from probe design software had been selected by eliminating those with ~80% or higher homology to 549 sequences representing 80 potyvirus species and four complete plant genomes. The probes were further selected by individually blasting each probe sequence against GenBank to eliminate probes with high homology to the other pathogens or plant tissues. Thirdly, MM probes (52 out of 85 probes) were

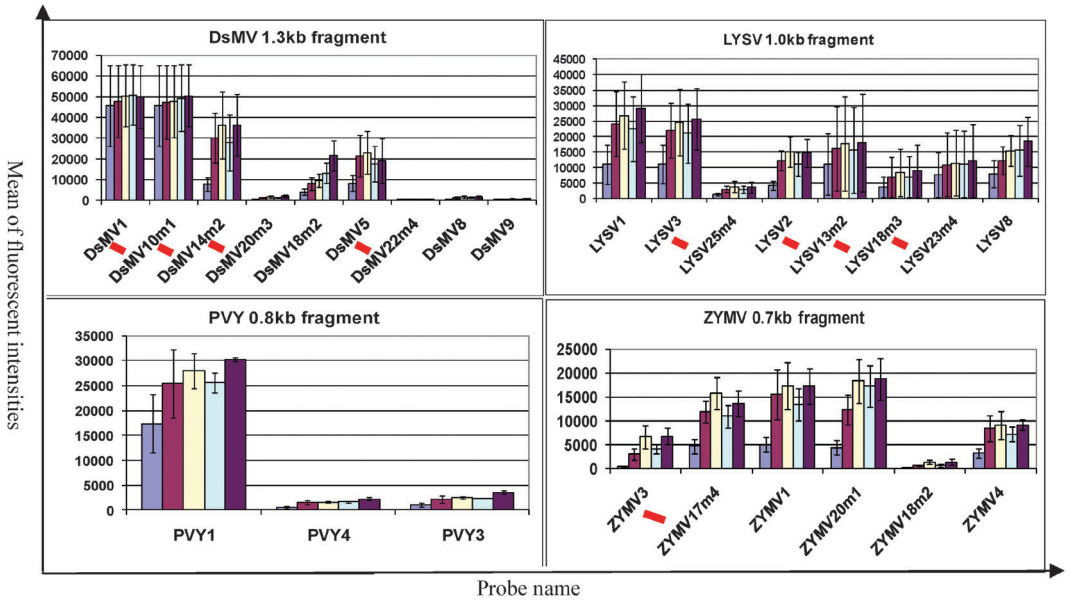

Fig. 4 Fluorescent intensity comparison of all probes showing positive signals with different spacer modifications and distance to the fluorescent labeling terminal. These data were from the hybridization using direct labeling dsDNA targets. Probes without spacers (first left bar in aqua color) and with different spacers (6C [light purple], 12C [light yellow], 6T [light blue] and 12T [dark purple]) were from the hybridizations using four targets: D13, L10, P8 and Z7. The probe names on the *x* axis were constructed by viral name + number of the probe + an "m" following an Arabic number representing the number of mismatch nucleotides on the probe. The height of the histograms represents the mean fluorescent pixel intensity of positive signals in triplicate with background subtracted. The error bars represent standard error of the mean. The probes sharing the same location but with different mismatch nucleotides are linked by a red bar.

Table 3 The distance of each probe showing a positive signal to the fluorescent labeling end in a direct labeling model

Probe name	Distance	Probe no.	Distance	Probe no.	Distance	Probe no.	Distance
DsMV1	66 nt	LYSV1	180 nt	PVY1	116 nt	ZYMV3	61 nt
DsMV10m1	66 nt	LYSV3	186 nt	PVY4	475 nt	ZYMV17m2	61 nt
DsMV14m2	66 nt	LYSV25m4	186 nt	PVY3	479 nt	ZYMV1	89 nt
DsMV20m3	66 nt	LYSV2	187 nt			ZYMV20m1	112 nt
DsMV18m2	152 nt	LYSV13m2	187 nt			ZYMV18m3	201 nt
DsMV5	509 nt	LYSV18m3	187 nt			ZYMV4	223 nt
DsMV22m4	612 nt	LYSV23m4	187 nt				
DsMV23m5	612 nt	LYSV8	316 nt				
DsMV8	707 nt						
DsMV9	769 nt						

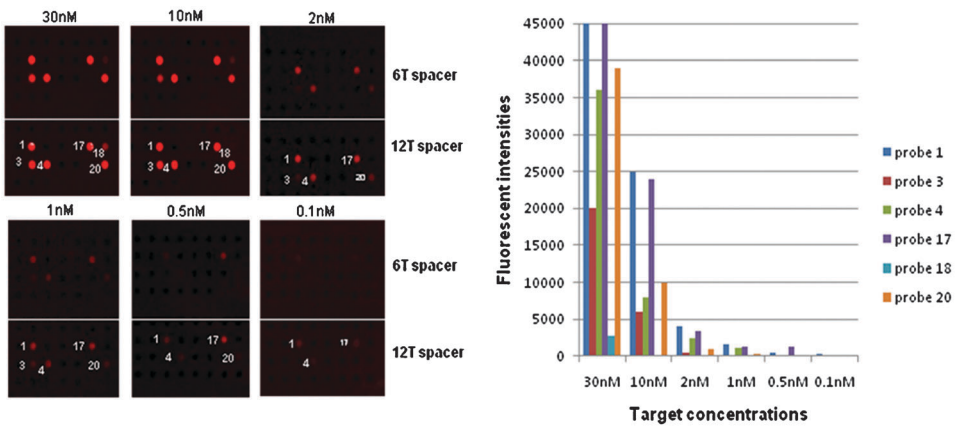

Fig. 5 Sensitivity of the microarray measured by dsDNA target Z7 at six different concentrations. The numbers to the left or lower-left side of the positive spots in the images represent relevant ZYMV probe identification numbers (see Table 1). The fluorescent intensities corresponding to each positive spot with the 12T spacer were shown in the right figure.

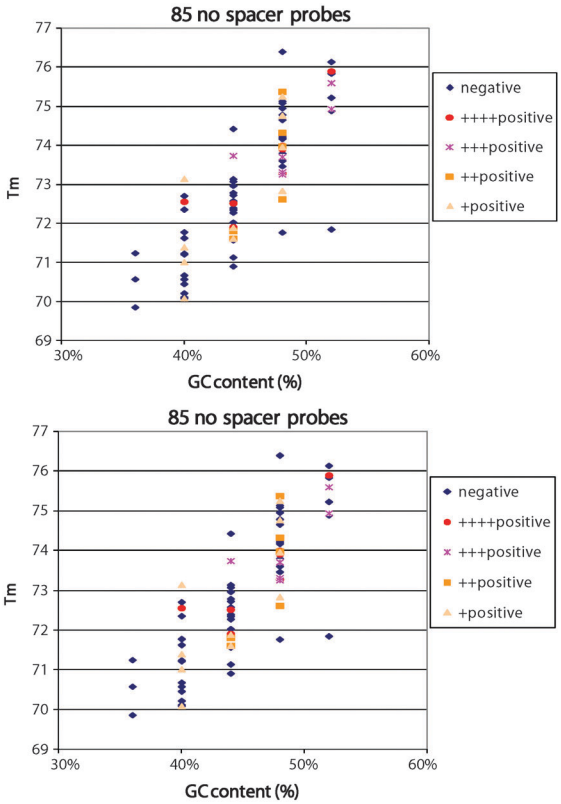

Fig. 6 Relationship between hybridization results and GC content and T_m of each probe. The data were from probes without a spacer and with the 12T spacer (data for other spacer modification probes were not included). ++++, +++, ++ and + are visual judgments based on the signal intensity of each probe.

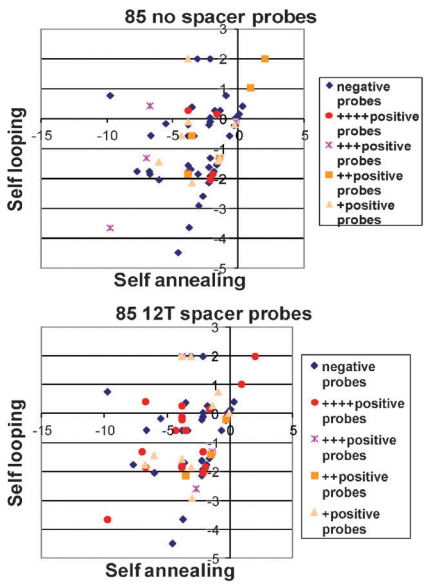

Fig. 7 Relationship between hybridization results and self-annealing and self-looping of each probe. The data were from probes without a spacer and with the 12T spacer (data for other spacer modification probes were not selected in order to avoid the mess colorful spots). ++++, +++, ++ and + are visual judgments based on the signal intensity of each probe.

designed with one to nine nucleotide mismatches (out of 25) to the targeted viral sequences, but these MM probes are perfect-match to some other isolates belonging to the same *Potyvirus* species from GenBank. These MM probes were primarily included to detect isolates with low homology to the type sequence, but they also provided an opportunity to test the potential of cross-hybridization. Fourthly, the steric hindrance mediated by the solid support which affects probe–target hybridization efficiency[18,38] was mitigated by adding spacer molecules to each probe.[18,19] Steric hindrance caused by the high density of immobilized probes[18,19] was avoided by printing the probes at a lower density of 4.5×10^{11} molecules cm^{-2}, which is slightly lower than the recommended density of 2×10^{12} molecules cm^{-2}.[21] Fifthly, to reduce the hybridization background and stabilize the DNA duplexes, higher salt concentration (150 mM NaCl)[45,48] was used in both hybridization and washing buffer, and a Liquid Blocking Reagent (Amersham Biosciences) was added to the hybridization buffer to reduce the background and increase the hybridization efficiency. Finally, multiple probes were designed from different genes/regions of each virus genome to further increase the reliability of the array.

Non-specific cross-hybridization in microarrays is a well known problem for microarray hybridization with as little as 70–80% sequence similarity having the potential to cause the cross-hybridization.[12–14,30] However in our study non-specific cross-hybridization was not observed, whether using targets from investigated potyviruses, closely related potyviruses or healthy plants, even at high target concentration, which indicated that the probes designed in this study are highly specific to the targeted potyviruses.

The density of the probe is a key factor in the capability of immobilized probes to capture the targets, as the crowding of the immobilized probes can cause the steric hindrance and consequently result in the lower hybridization efficiency.[18,19] It has been demonstrated that lower probe density can provide more effective hybridization and the kinetics of binding are comparatively faster.[38,49] Peterson *et al.*[21] recommended less than 2×10^{12} molecules cm^{-2} as the optimal density of probes. The density of the probes in this study was 4.5×10^{11} molecules cm^{-2} which matches the normal requirement, and therefore the possible steric hindrance caused by the crowding of the probes should be mitigated or removed.

Hybridization failure was observed between some perfectly matched targets and probes as well as some mismatched probes which further highlights the importance of investigating a range of factors that contribute to hybridization failure. It is commonly recognized that perfectly matched targets and probes (from *in silico* design) in microarrays sometimes fail to form duplexes and result in a negative signal.[15,43,50] Naef *et al.*[16] and Zhou and Abagyan[17] both reported that around 30% of probes consistently failed to generate positive signals in microarray hybridization. In this study, 18 out of 33 perfect-match probes (55%) failed to hybridize with their target fragments which was similar to results presented by Naef and Magnasco[42] when discussing hybridization failure between perfectly matched probes.

Schmidt[34] reported that target fragments with sizes of 0.25–0.8 kb provided the best hybridization efficiency for

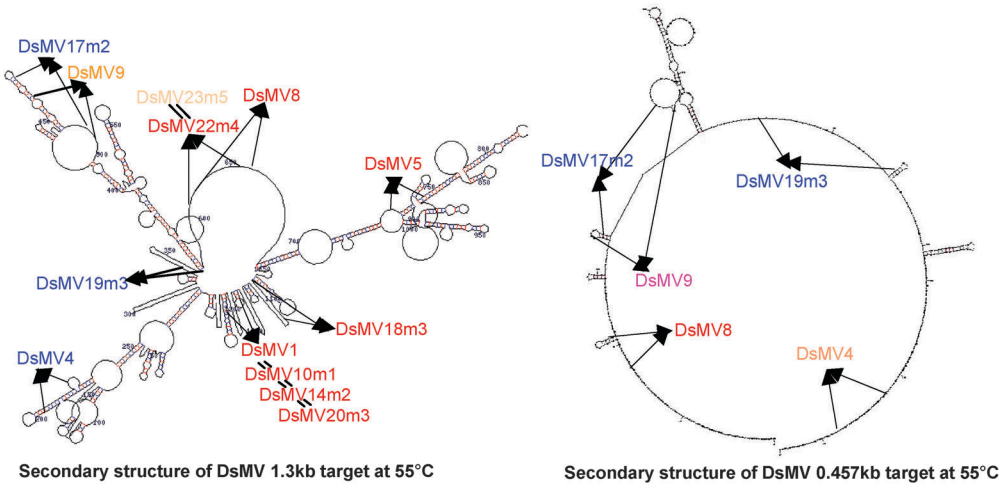

Fig. 8 Relationship between the hybridisation efficiency and the target secondary structure at 55 °C demonstrated by two overlapped dsDNA targets. DsMV 0.457 kb target (right) was overlapped with DsMV ~1.3 kb target (left) with the sharing of five probes: DsMV4, 8, 9, 17m2 and 19m3. The location of each probe is indicated by an arrow and the probe identification number and the Arabic number after an "m" represents the number of mismatch nucleotide on that probe. Hybridization results for each probe are indicated by different colours: red = ++++positive, orange = ++positive, light orange = +positive, blue = negative. The probes with the same location but with different mismatch nucleotide(s) are indicated by linking two probe identification numbers with two parallel straight lines.

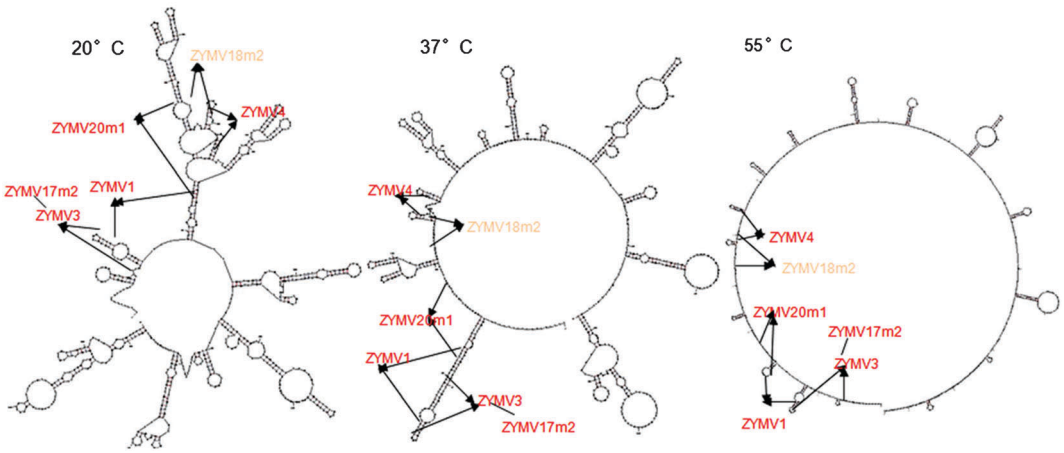

Fig. 9 Comparison of hybridisation efficiency and secondary structures at 20 °C, 37 °C and 55 °C using dsDNA target Z7. The location of each probe is indicated by two arrows and the probe identification number (see Table 1). Hybridization results for each probe are indicated by different colours: red = ++++positive, orange = ++positive, light orange = +positive, blue = negative. The probes with the same location but with different mismatch nucleotide(s) are indicated by linking two probe identification numbers with a single straight line.

oligonucleotide microarrays. In our study, targets of ~0.335 kb, ~0.7 kb or ~1.3 kb were initially generated from the four viral species and compared. However only one out of the three targets, excluding the shortest ~0.335 kb target, gave positive reactions. In addition, when ten shorter virus-specific fragments (0.277 to 0.736 kb; one fragment covering probes with both previous positive and negative signals and nine fragments covering probes with all negative signals) were generated and hybridized with the array, eight of the fragments gave at least one extra positive probe. However two fragments of shorter size (0.285 and 0.367 kb) remained negative reactions to all the included probes, this supports the concept that shorter targets can increase the hybridization efficiency in some but not all instances.

The labeling system is a major factor affecting the microarray efficiency.[27] Xiang *et al.*[28] and Bystricka *et al.*[1] considered cDNA with indirect labeling more effective and less expensive than direct labeling. However, in our study using PCR products, direct labeling was considered cheaper, easier and quicker than indirect labeling. Both labeling systems provided almost identical hybridization results for all the tested targets. In addition to the form of labelling, target strand type is also important. Single stranded DNA (ssDNA) fragments have been reported to provide better hybridization efficiency than double stranded DNA (dsDNA) fragments.[31] However, the results from our study showed no obvious difference between hybridizations using ssDNA and dsDNA fragments, and dsDNA is much more convenient and less expensive to generate.

Steric hindrance is an important factor caused by the interference of the solid support and affects the hybridization efficiency.[19,20]

The addition of spacers to extend the distance between the probes and glass support has been shown to mitigate the impacts of steric hindrance and improve the hybridization efficiency.[19,22,23,31] A previous study[43] concluded that a C12 (carbon12) linker plus a five thymine residue spacer at the 5′ end of the probe provided optimal spacing to mitigate the steric impacts. In our study, we compared probes with a C6 linker (AmC6) but without an additional nucleotide spacer, or with a six or twelve cytosine or thymine residue spacer at the 5′ terminus of all the probes. Results showed that all probes with spacers gave a greater number of positive signals and higher signal intensities than those without a spacer. In most cases, probes with twelve cytosine or thymine residues provided higher signal values than those with six residue spacers. However in most instances, the differences between with spacers and without spacers, and between six residue and twelve residue spacers were not significant ($p > 0.05$). This may be due to the variable fluorescent intensity values obtained from the triplicate experiments using different slides which often resulted in a high standard error (Fig. 4).

The position of the fluorescent label is also considered a crucial factor affecting the signal intensity in microarray hybridization, with "shorter distances between the hybridization site and the fluorescent label providing higher signal intensity to perfect-match probes".[29] However, this phenomenon was not consistently observed in this study as some probes showing higher positive signals were with longer distance to the fluorescent labelled end than those with shorter distance (Fig. 4).

All assays have a detection limit and there is no exception to microarrays. Nagino *et al.*[44] reported that microarray signal strength decreased with decreasing target DNA concentration. Not surprisingly the same trend was observed in our study (Fig. 5). However, when higher target concentration was used, there were no extra probes with positive signals observed although there was an increase in positive signal intensity (Fig. 5). Khadijah *et al.*[51] demonstrated that a microarray for the detection of white spot syndrome virus in asymptomatic shrimps could detect the presence of three viral transcripts with the combination of the *in vitro* transcription technology. Wilson *et al.*[33] reported that the sensitivity of a microarray combined with PCR can be as low as 500 fg of pathogen DNA. In this study, our array could detect as little as ~3.4 ng PCR product which represents ~3.2 fg of original virus cDNA.

The hybridization periods used in different microarray studies vary widely but normally range from 2 hours to 16 hours.[52] Longer hybridization periods (10 to 19 hours) are generally used to ensure complete hybridization to increase the microarray efficiency.[21,43–46] On the other hand, short hybridization periods (2 to 6 hours) are also reported to provide satisfactory hybridization efficiency.[10,33] A recent study[47] showed that 1–2 hours gave optimal and higher signals than 16 hour hybridization in the detection of bacteria using a 50- to 70-mer microarray. In this study, we compared the short (2 h) and long (overnight) hybridization period and found that these two hybridization periods resulted in similar positive and negative patterns. Therefore, 2 hour hybridization was considered optimal for the current study.

Mismatch nucleotides are commonly used to discriminate sequence variants of the target species. Varied results have been reported for the hybridization of mismatch probes as the number, type (*i.e.* A, C, G or T) and position of the mismatch nucleotides on a probe can affect the stability of the DNA duplexes, especially at the centre of a probe.[31,37–39,53] In some instances, single mismatches with different nucleotides on the same probes did not provide identical destabilizing effects[16,17,40] and the effects of the mismatch position could be ignored.[41] On the other hand, some mismatch probes yield higher signal intensities to the target than those of the corresponding perfect-match probes.[42] In our study, a total of 52 mismatch probes, with 1–9 different mismatch nucleotides, of different types (A, C, G or T) and location (3′ terminus, centre and 5′ terminus) were designed. Approximately 25% (13 out of 52) of the mismatch probes provided positive signals with the number of mismatched nucleotides ranging from 1 to 5. Although the signal strength varied, there was no obvious correlation between the number of mismatched nucleotides or their position on the probes (Table 1). It was also observed that the fluorescent intensity values of some mismatch probes were higher than some with fewer mismatched nucleotides and even some perfect match probes (Fig. 4 and Table 1). Our results indicate that there are no absolute rules for predicting the influence of mismatch type, position or number of mismatched nucleotides on duplex formation and stability.

Base composition on a probe is reported to have important effects on DNA duplex formation and stability due to the lower stability of A:T pairs compared to G:C pairs. Since the central portion of the probe is more important in stabilizing the duplex formation,[36] a probe with most of its GC content in the central region should bind more strongly to the target than a probe with high GC content at other locations.[19,20,25] In addition, probes with higher GC content are often considered to provide higher signal intensities than probes with lower GC content.[10] However the 85 potyvirus probes developed in this study, which differed in both their GC content (36–52%) and location, showed no obvious correlation between the hybridization efficiency and the GC content/T_m (Fig. 7). This might indicate that the GC content and T_m of the probes are not critical factors which can obviously affect the hybridization efficiency.

The intra- and inter-molecular structure of oligonucleotide probes can compete with probe–target duplex formation resulting in a low hybridization efficiency.[35] It was reported that 40-mer probes with $\Delta G \leq -65$ kcal mol^{-1} could provide satisfactory hybridization results.[1] In our study, both self-annealing and self-looping of each probe together with the hybridization results were analyzed. The probes with positive and negative signals were found distributed randomly within a ΔG range of "−11 to +2" for self-annealing and "−5 to +2" for self-looping (Fig. 7), indicating that probe self-annealing and self-looping with the described range did not have an obvious and consistent effect on the duplex formation.

It is well known that the secondary structures of the target fragments play a significant role in hybridization efficiency because the complicated secondary structures of the targets, especially long target fragments, can affect the accessibility of probes to the targets, the stability of the duplex, and consequently its binding efficiency.[18,31] The MFold[54] software for predicting

secondary structure of RNA or DNA molecules uses the most recent energy minimization model and since its publication in 2003, it has been widely used (cited 2893 times up to 2010, http://mfold.rna.albany.edu/?q=mfold) and considered a popular software tool for RNA or DNA secondary structure prediction. In this study, secondary structures of targets predicted by Mfold did not correlate with the hybridization efficiency (Fig. 8). No consistent rule could be found from the secondary structure to assist in predicting the hybridization efficiency.

From the combined results of our study, none of the investigated factors could be identified as the main cause of hybridization failure since positive hybridizations were typically consistent across a range of different conditions for both perfectly matched probes and targets as well as those with mismatch oligonucleotides. This was found regardless of probe type (PM or MM), the distance from a probe to the fluorescent labeling end, direct or indirect labelling, target length, target strand type (dsDNA or ssDNA), hybridization period (2 h or overnight), secondary structure of probes and targets generated by software, or steric hindrance between probes and targets.

It is clear that a range of factors affect the specificity and sensitivity of microarray hybridization and many of these can be modified by the experimenter to optimise the reaction. However, from this and other studies it is evident that even probes with a perfect match to the target often fail despite manipulation of the hybridization conditions. For example in this study only 42.4% of perfect match probes successfully hybridized. Although various modifications can be made to the hybridization reaction the one critical factor over which the experimenter has little control is the inherent secondary and tertiary structure of the target DNA. This can be mitigated to some extent by using short target fragments as was demonstrated when the use of shorter PCR products (ranging from 0.277 to 0.736 kb) resulted in hybridization from probes that did not react with longer targets (Table 2). Even so, some probes still did not hybridize. Although target secondary structures generated by Mfold were analysed none correlated to the hybridization efficiency in this study, our evidence strongly suggests that microarray hybridization failure is mainly due to the target structure inaccessible regions. Firstly, it was found that once a probe gave a positive signal, its MM probe(s) all gave positive signals regardless of the length of targets, probe status and hybridization conditions. For example, probes DsMV10m1, DsMV14m2 and DsMV20m3 mismatched to DsMV1; DsMV23m5 mismatched to DsMV22m4 (Fig. 8). This suggests that once target sequences can be accessed, even the mismatched probes can provide positive signals. Secondly, PM probes sometimes showed lower fluorescent intensity than MM probes suggesting that completely and partially accessible sites might exist; the former might provide better hybridization opportunity for even mismatched probes while the latter might provide difficulty for even perfectly matched probes. Thirdly when the ten shorter virus-specific target fragments were generated and tested, two still resulted in all probes giving a negative signal, while the other eight targets result in one or more additional probe(s) giving positive signals. These results demonstrated that most of the shorter fragments formed more simple secondary structures which resulted in more accessible sites for successful hybridization, while some of the shorter targets still maintained some or all non-accessible sites. Fourthly, when four different targets giving both positive and negative signals were hybridized in a mixture with the microarray, positive patterns were still the same when only a single target was used which suggests that each target forms its own specific structure not affected by the presence of the other targets. These results indicate that each target has its own specific and consistent structure, presumably including three sub-structures within it: (1) completely exposed regions complementary to PM or MM probes that enable the target to hybridize stably and show strong fluorescent signals, (2) partially exposed regions only partially complementary to the probes resulting in less stable hybridization and showing weak fluorescent signals, and (3) inaccessible regions that probes are unable to hybridize resulting in complete hybridization failure. Therefore it is hypothesised that successful hybridization between probes and targets will depend mainly on the target structure. In other words, target structure is the crucial factor resulting in microarray hybridization failure. Based on our results, it appears that accurate prediction of the secondary and tertiary structure of the target is an essential prerequisite for the successful design of microarray probes. For the present we can continue to design more probes than required and experimentally select those that work, but the challenge for the future is to develop new software that better predicts the target structure for successful probe–target hybridization. Our data might be helpful for combining bioinformatic technology to develop new software.

Methods

Probes and arrays

Eighty five 25-mer probes, including 33 perfect-match (PM) and 52 mismatch (MM) probes, were designed based on the sense partial sequences of four New Zealand *Potyvirus* species: *Dasheen mosaic virus* (DsMV, AY994104), *Leek yellow stripe virus* (LYSV, AY842136), *Potato virus Y* (PVY, DQ217931) and *Zucchini yellow mosaic virus* (ZYMV, AY995216). Due to the variability of potyviruses, longer probes were unable to be designed for covering different viral species. The mismatch probes (with one to nine nucleotides mismatch and different types of nucleotide mismatches at different locations on the probe) were selected directly from sequences of virus isolates with different mismatch patterns. Among these probes, 23 (9 PM and 13 MM) are for DsMV, 28 (10 PM and 18 MM) for LYSV, 14 (6 PM and 8 MM) for PVY and 20 (8 PM and 12 MM) for ZYMV. Five negative control probes of the same length were designed based on the human sequences (accession no.: NM_000979 and NT_011109). Two positive probes (22-mer) and their complementary synthetic targets labeled with cyanine fluorescent dye (Cy5) were provided by the Centre of Applied Genesensor (CAG, University of Bremen, Germany). Polycytosine (C) and poly-thymine (T) spacers with different nucleotide numbers (6C, 12C, 6T and 12T) were added to the 5′ terminus of each of the viral and negative control probes. All probes were then modified at the 5′ end with an extra amino modifier consisting of a 6-carbon spacer arm. Probes were synthesized at ILLUMINA, Inc. (San Diego, California, USA)

and printed onto glass slides with a density of 4.5 × 10^{11} molecules cm^{-2} by the CAG. Printing used a non-contact printing robot equipped with a TopSpot microfluidic printhead. Aminosilane (3-aminopropyltrimethoxysilane) and a PDITC-linker (1,4-phenylendiisothiocyanate) coated glass slides were purchased from Asper Biotech (Tartu, Estonia now Genorama Ltd.). Probes were spotted at a concentration of 10 μM in 200 pL droplets of 150 mM Na_3PO_4 buffer, pH 8.5, resulting in a spot diameter of about 220 μM. For measurements using SYBR Gold (Applied Biosystems), the coupling efficiency was calculated to be about 10% leading to a probe density of about 10^{11} capture molecules per cm^2 (4.5 × 10^{11} molecules cm^{-2}). After 16 h incubation in a wet chamber the microarrays were stored under a nitrogen atmosphere for shipping from Germany to New Zealand. The printing density was suggested by CAG based on their more than five years experience resulting in arrays with lower steric hindrance and higher hybridization efficiency.

Viral targets

Three different sizes of target fragments (~0.335 kb, ~0.7–1.0 kb and ~1.3 kb; the ~0.335 kb fragment overlaps the 3′ and 5′ regions of the other two fragments, respectively) were PCR amplified using *Potyvirus* universal primers[55,56] from the recombinant plasmids containing sequences of the four New Zealand *Potyvirus* species. Based on the sequences of these products, specific primers were then designed for each species to amplify different sized fragments (0.277 to 0.736 kb) covering the different regions where probes with negative signals or negative/positive signals were located.

Targets with double stranded DNA (dsDNA) were used for hybridization individually or in a mixture. Single stranded DNA (ssDNA) targets were isolated from selected dsDNA targets which contained probes showing only negative signals or showing both positive and negative signals, and then hybridized with the microarray for the comparison. The ssDNA targets were prepared using forward primers labelled with biotin and then coupled with fluorescent dye labelled reversed primers to amplify from the dsDNA fragments. The anti-sense strands of targets containing fluorescent dye were then isolated using Streptavidin Magnetic Particles (Roche Diagnostics, Germany) following the manufacturer's protocol.

In addition, four New Zealand potyviruses which are phylogenetically close to the four investigated potyviruses: *Ornithogalum mosaic virus* (OrMV), *Ornithogalum virus* 2 (OrV2), *Onion yellow dwarf virus* (OYDV) and *Zantedeschia mild mosaic virus* (ZaMMV) plus four disease-free plant species: *Nicotiana benthamiana* (*N. benthamiana*), *Cucurbita pepo* subsp. *Pepo* (zucchini), *Colocasia esculenta* (Taro) and *Dactylis glomerata* (Cocksfoot) were used for testing the specificity of the array. Target fragments from closely related potyviruses were the same as those from investigated potyviruses described above. Negative control targets from healthy plants were cDNA synthesized by random hexamer primers (Invitrogen).

Labelling models

Direct and indirect labeling models using cyanine fluorescent Cy5 or Cy3 dye (Amersham Biosciences) were compared. For direct labeling, PCR products were generated using Cy5 dye labeled reverse primers and the resulting antisense strands (targets) contained the fluorescent dye at the 5′ terminus. For indirect labeling, the PCR for generating targets from potyviruses, and the reverse transcription for generating cDNA from healthy plants were performed with a dNTP mixture containing aminoallyl- and aminohexyl-modified nucleotides (SuperScript™ Indirect cDNA Labeling System, Invitrogen™ Life technologies). This enabled the resultant PCR products to couple with a cyanine fluorescent dye using a SuperScript™ Indirect cDNA Labeling System (Invitrogen™ Life technologies).

Hybridization

Microarray hybridizations were performed according to the following protocol: (1) hybridization mixtures in a total volume of 70 μL contained 10 nM of viral target, 1 nM positive control target, 1 × Liquid Blocking Reagent (Amersham Biosciences), and 1 × hybridization buffer (20 mM Tris-HCl, pH 7.3; 150 mM NaCl; 5 mM EDTA, pH 7.3; 0.05% Tween-20; 0.045% milk powder; 0.01 mg mL^{-1} salmon sperm DNA). (2) The hybridization mixture was denatured at 95 °C for 5 min and then immediately placed on ice for at least 3 min. Before loading hybridization mixtures onto slides, the corners of the area containing the probes were marked on the reversed side of the slides using a permanent marker pen and a 1.5 × 1.6 cm Gene Frame (ABgene®, UK) was then applied onto the slide to enclose the probe area so that hybridization mixtures would be held in the probe area. After mixtures were loaded onto the slides, a polyester coverslip (ABgene®, UK) was placed on each Gene Frame avoiding bubbles and leakage. Slides were then put in a normal incubator with around 70% humidity and incubated at 55 °C for 2 hours or overnight. (3) Slides were firstly washed twice in TETBS buffer (TBS buffer plus 5 mM EDTA and 0.05% Tween-20) and then twice in TBS buffer (20 mM Tris-HCl, pH 7.3 and 150 mM NaCl). (4) After drying by centrifugation at 1500 rpm for 3 min, slides were scanned using a GenePix™ 4000B microarray reader (Axon Instruments), with lasers of wavelengths 532 nm and 635 nm used to excite Cy3 dye (green color) and Cy5 dye (red color), respectively. The photomultiplier tube (PMT) in a GenePix scanner, which is used to detect the photons that are emitted from the laser excited fluorophores on the array, was set at 600 in most of the cases. But 700 and 800 were also used for visualizing the weak positive signals. The acquired images were initially analyzed by GenePix® Pro 5.0 software (Axon Instruments, Inc.) using the median fluorescent intensity value of each spot with the background subtracted. The fluorescent intensity data were normalized across all of the slides to allow comparisons among different slides from separated experiments.

The impact of target concentration on hybridization efficiency was evaluated using a selected target containing probes with both positive and negative signals at six different concentrations of 0.1, 0.5, 1, 2, 10 and 30 nM.

Data analysis

The impact of probe secondary structure (analyzed by Oligo Analyzer 1.0.2 [Teemu Kuulasmaa, http://molbiol-tools.ca/molecular_biology_freeware.htm]), T_m and GC content on

array hybridization efficiency was analyzed using Chart Wizard in Microsoft EXCEL. The potential impact caused by secondary structure of viral targets was analyzed by comparing the location of each probe sequence within the secondary structure of the target sequence (antisense strand) and the relative hybridization result. Secondary structures of antisense strands of targets were analyzed using Mfold online service (http://www.bioinfo.rpi.edu/applications/mfold)[54] under three different temperatures of 20 °C (room temperature), 37 °C (temperature between room temperature and hybridization temperature) and 55 °C (hybridization temperature). These three temperature points were selected based on the hybridization procedure starting from room temperature *via* 37 °C to 55 °C hybridization temperature which might result in the formation of different secondary structures. The secondary structure output file of each single target includes several different structures with different stable strengths. The most stable structure was presumed to predominate in the mixture and thus was selected for the comparison analysis. The results of the different spacer treatments were analyzed by one-way-analysis of variance (ANOVA) using Minitab Software (release 15, PA, USA).

Acknowledgements

We thank the National centre for Advanced Bio-Protection Technologies, New Zealand for providing fund for this project. We thank Dr Kelvin Lau, Associate professor Cristin Print, Dr Franz Pichler, Dr Lei Zhang, Dr Sasha Todt, Mr Frank Meyerjürgens and Mr Sven Roll for their comments and kind help while this project was in progress. We also thank Dr Haran Sivakumaran in Queensland Institute of Medical Research for his kind help on modifying figures.

References

1 D. Bystricka, O. Lenz, I. Mraz, L. Piherova, S. Kmoch and M. Sip, *J. Virol. Methods*, 2005, **128**, 176–182.
2 V. Martín, C. Perales, D. Abia, A. R. Ortíz, E. Domingo and C. Briones, *BMC Bioinf.*, 2006, **7**, 117.
3 T. Wei, M. N. Pearson, D. Blohm, M. Nölte and K. Armstrong, *J. Virol. Methods*, 2009, **162**, 109–118.
4 S. Biswas, J. Manikandan and P. N. Pushparaj, *Bioinformation*, 2011, **6**(4), 153–157.
5 P. S. Meltzer, *Curr. Opin. Genet. Dev.*, 2001, **11**, 258–263.
6 D. G. Wang, J. B. Fan, C. J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, T. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. P. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, R. Lipshutz, M. Chee and E. S. Lander, *Science*, 1998, **280**, 1077–1082.
7 C. Debouck and P. N. goodfellow, *Nat. Genet.*, 1999, **21**, 48–50.
8 L. C. Hsiung, C. L. chiang, C. H. Wang, Y. H. Huang, C. T. Kuo, J. Y. Cheng, C. H. Lin, V. Wu, H. Y. chou, D. S. Jong, H. Lee and A. M. Wo, *Lab Chip*, 2011, **11**, 2333.
9 K. Rudi, S. L. Flateland, J. F. Hanssen, G. Bengtsson and H. Nissen, *Appl. Environ. Microbiol.*, 2002, **68**, 1146–1156.
10 T. Z. DeSantis, C. E. Stone, S. R. Murray, J. P. Moberg and G. L. Anderse, *FEMS Microbiol. Lett.*, 2005, **245**, 271–278.
11 D. Y. Guschin, B. K. Mobarry, D. Proudnikov, D. A. Stahl, B. E. Rittmann and A. D. Mirzabekov, *Appl. Environ. Microbiol.*, 1997, **63**, 2397–2402.
12 J. C. Barrett and E. S. Kawasaki, *Drug Discovery Today*, 2003, **8**, 134–141.
13 Y. C. Hsu, T. J. Yeh and Y. C. Chang, *J. Virol. Methods*, 2005, **128**, 54–60.
14 M. D. Kane, T. A. Jatkoe, C. R. Stumpf, J. Lu, J. D. Thomas and S. J. Madore, *Nucleic Acids Res.*, 2000, **28**, 4552–4557.
15 A. Loy, A. Lehner, N. Lee, J. Adamczyk, H. Meier, J. Ernst, K. H. Schleifer and M. Wagner, *Appl. Environ. Microbiol.*, 2002, **8**, 5064–5081.
16 F. Naef, C. R. Hacker, N. Patil and M. Magnasco, *Genome Biol.*, 2002, **3**, research0018.
17 Y. Zhou and R. Abagyan, *BMC Bioinf.*, 2002, **3**, 3.
18 J. Peplies, F. O. Glöckner and R. Amann, *Appl. Environ. Microbiol.*, 2003, **69**, 1397–1407.
19 M. S. Shchepinov, S. C. Case-Green and E. M. Southern, *Nucleic Acids Res.*, 1997, **25**, 1155–1161.
20 Z. Guo, R. A. Guilfovle, A. J. Thiel, R. Wang and L. M. Smith, *Nucleic Acids Res.*, 1994, **22**, 5456–5465.
21 A. W. Peterson, R. J. Heato and R. M. Georgiadis, *Nucleic Acids Res.*, 2001, **29**, 5163–5168.
22 D. J. Duggan, M. Bittner, Y. Chen, P. Meltzer and J. M. Trent, *Nat. Genet.*, 1999, **21**, 10–14.
23 A. Relógio, C. Schwager, A. Richter, W. Ansorge and J. Valcárcel, *Nucleic Acids Res.*, 2002, **30**, e51.
24 S. Matysiak, N. Hauser, S. Wurtz and J. Hoheisel, *Nucleosides Nucleotides*, 1999, **18**, 1289–1291.
25 T. R. Hughes, M. Mao, A. R. Jones, J. Burchard, M. J. Marton, J. W. Shannon, S. M. Kefkowitz, M. Ziman, J. M. Schelter, M. R. Meyer, S. Kobayashi, C. Davis, H. Dai, Y. D. He, S. B. Stephaniants, G. Cavet, W. L. Walker, A. West, E. Coffey, D. D. Shoemaker, R. Stoughton, A. P. Blanchard, S. H. Friend and P. S. Linsley, *Nat. Biotechnol.*, 2001, **19**, 342–347.
26 A. Vainrub and B. M. Pettitt, *Biopolymers*, 2003, **68**, 265–270.
27 E. V. Grigorenko, *DNA arrays*, CRC press, Boca Raton, USA, 2002.
28 C. C. Xiang, O. A. Kozhich, M. M. Chen, J. M. Inman, Q. N. Phan, Y. Chen and M. J. Brownstein, *Nat. Biotechnol.*, 2002, **20**, 738–742.
29 L. Zhang, T. Hurek and B. Reinhold-Hurek, *Nucleic Acids Res.*, 2005, **33**, e166.
30 E. M. Evertsz, J. Au-Young, M. V. Ruvolo, A. C. Lim and M. A. Reynolds, *BioTechniques*, 2001, **31**, 1182–1184.
31 L. Zhang, *The faculty of Biology and Chemistry*, *PhD thesis*, University of Bremen, Bremen, Germany, 2004, p. 185.
32 R. N. Bhattacharjee, K. S. Park, K. Okada, Y. Kumagai, S. Uematsu, O. Takeuchi, S. Akira, T. Iida and T. Honda, *Biochem. Biophys. Res. Commun.*, 2005, **335**, 328–334.
33 W. J. Wilson, C. L. Strout, T. Z. DeSantis, J. L. Stilwell, A. V. Carrano and G. L. Andersen, *Mol. Cell. Probes*, 2002, **16**, 119–127.
34 M. P. Schmidt, *PhD thesis*, Bremen University, Bremen, Germany, 2004.
35 O. V. Matveeva, S. A. Shabalina, V. A. Nemtsov, A. D. Tsodikov, R. F. Gesteland and J. F. Atkins, *Nucleic Acids Res.*, 2003, **31**, 4211–4217.
36 L. Zhang, C. Wu, R. Carta and H. Zhao, *Nucleic Acids Res.*, 2007, **35**, e18.
37 A. V. Fotin, A. L. Drobyshev, D. Y. Proudnikov, A. N. Perov and A. D. Mirzabekov, *Nucleic Acids Res.*, 1998, **26**, 1515–1521.
38 E. M. Southern, K. U. Mir and M. Shchepinov, *Nat. Genet.*, 1999, **21**, 1–9.
39 H. Urakawa, S. E. Fantroussi, H. Smidt, J. C. Smoot, E. H. Tribou, J. J. Kelly, P. A. Nobel and D. A. Stahl, *Appl. Environ. Microbiol.*, 2003, **69**, 2848–2856.
40 N. Sugimoto, M. Nakano and S. Nakano, *Biochemistry*, 2000, **39**, 11270–11281.
41 J. W. Szostak, J. I. Stile, B.-K. Tye, P. Chiu, F. Sherman and R. Wu, *Methods Enzymol.*, 1979, **68**, 419–428.
42 F. Naef and M. O. Magnasco, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2003, **68**, 011906.
43 L. Bodrossy, N. Stralis-Pavese, J. C. Murrell, S. Radajewski, A. Weilharter and A. Sessitsch, *Environ. Microbiol.*, 2003, **5**, 566–582.
44 K. Nagino, O. Nomura, Y. Takii, A. Myomoto, M. Ichikawa, F. Nakamura, M. Higasa, H. Akiyama, H. Nobumasa, S. Shiojima and G. Tsujimoto, *J. Biochem.*, 2006, **139**, 697–703.
45 K. Rose, J. O. Mason and R. Lathe, *BioTechniques*, 2002, **33**, 54–58.
46 Y. Wang, Y. Li, S. Liu, W. Shen, B. Jiang, X. Xu and Y. Xie, *J. Nanosci. Nanotechnol.*, 2005, **5**, 1249–1255.

47 J. E. Burton, O. J. Oshota, E. North, M. J. Hudson, N. Polyanskaya, J. Brehm, G. Lloyd and N. J. Silman, *Mol. Cell. Probes*, 2005, **19**, 349–357.
48 P. Tijssen, *Laboratory techniques in biochemistry and molecular biology*, Elsevier Science, Amsterdam, 1993, vol. 24.
49 P. V. Riccelli, F. Merante, K. T. Leung, S. Bortolin, R. I. Zastawny, R. Janeczko and A. S. Benight, *Nucleic Acids Res.*, 2001, **29**, 996–1004.
50 E. M. Southern and U. Maskos, *J. Biotechnol.*, 1994, **35**, 217–227.
51 S. Khadijah, S. Y. Neo, M. S. Hossain, L. D. Miller, S. Mathavan and J. Kwang, *J. Virol.*, 2003, **77**, 10162–10167.
52 A. Halperin, A. Buhot and E. B. Zhulina, *Biophys. J.*, 2005, **89**, 796–811.
53 D. A. Stahl and R. Amann, *Development and application of nucleic acid probes*, John Wiley & Sons Ltd, Chichester, UK, 1991.
54 M. Zuker, *Nucleic Acids Res.*, 2003, **31**, 3406–3415.
55 S. A. Langeveld, J. M. Dore, J. Memelink, A. F. Derks, C. I. van der Vlugt, C. J. Asjes and J. F. Bol, *J. Gen. Virol.*, 1991, **72**, 1531–1541.
56 A. M. Mackenzie, M. Nolan, K.-J. Wei, M. A. Clements, D. Gowanlock, B. J. Wallace and J. Gibbs, *Arch. Virol.*, 1998, **143**, 903–914.

