Development of a DNA-based Microarray for the Detection of Zoonotic Pathogens in Rodent Species

Giles, Timothy[1]; Yon, Lisa[1]; de Bree, Freddy[2]; Bossers, Alex[2]; Hannant, Duncan[1]; Barrow, Paul[1]; Abu-Median, Abu-Bakr[1].

[1]University of Nottingham, Nottingham, United Kingdom

2 Central Veterinary Institute, Wageningen University and Research Centre, The Netherlands

Abstract

The demand for diagnostic tools that allow simultaneous screening of samples for multiple pathogens is increasing because they overcome the limitations of other methods, which can only screen for a single or a few pathogens at a time. Microarrays offer the advantages of being capable to test a large number of samples simultaneously, screening for multiple pathogen types per sample and having comparable sensitivity to existing methods such as PCR. Array design is often considered the most important process in any microarray experiment and can be the deciding factor in the success of a study. There are currently no microarrays for simultaneous detection of rodent-borne pathogens. The aim of this report is to explicate the design, development and evaluation of a microarray platform for use as a screening tool that combines ease of use and rapid identification of a number of rodent-borne pathogens of zoonotic importance. Nucleic acid was amplified by multiplex biotinylation PCR prior to hybridisation onto microarrays. The array sensitivity was comparable to standard PCR, though less sensitive than real-time PCR. The array presented here is a prototype microarray identification system for zoonotic pathogens that can infect rodent species.

Keywords

Microarray, Development, Rodent, Zoonoses, ArrayTube, ArrayStrip

Highlights

- We have developed a microarray to detect zoonotic pathogens in rodent species.
- The design stage of a microarray experiment is crucial for a successful experiment.
- We examined the difference between amplification methods prior to hybridisation.

1.1 Introduction

Prompt detection of pathogens is a significant issue in diagnostic testing for both human and veterinary health. This is particularly relevant when slow-growing or fastidious organisms are involved and the limitations of some existing diagnostic tools are driving researchers to consider alternative methods, as demands on quantity and rapidity of testing methods are increasing [1]. Serological methods provide an indication of exposure to a pathogen and are best used for screening populations. However, they also require an adequate time post-infection/exposure for antibodies to develop and may be unable to distinguish between different strains or antigenic types of pathogen. Zoonotic pathogens make up the majority (75%) of emerging diseases and wildlife are a major source of these pathogens [2]. Early detection of pathogens in wild animals would be useful in identifying risk factors associated with disease transmission to humans or domestic animals, and this could help prevent a possible outbreak. It has also been suggested that prevention of disease, which could be aided by an effective surveillance system, is better than reacting to an outbreak, or to finding a cure [3]. Microarrays offer the advantage of testing large numbers of samples simultaneously, coupled with screening a single sample for multiple pathogens. Use of this technology would enable timely, accurate and inexpensive detection of pathogens, which could lead to more effective control of these infectious diseases, which has positive implications for public health [4]. There are a wide ranging number of potential applications for pathogen detection arrays; and have been used for the detection of novel pathogens, as in the case of severe acute respiratory syndrome (SARS) [5], simultaneous detection of

Newcastle disease virus and avian influenza virus in birds [6] and detection of viruses that can cause vesicular or vesicular-like lesions in livestock [7].

Although microarrays are used widely, the fluorescence-based glass slide arrays are relatively expensive. Alternatives to the glass slide microarray are the ArrayTube™ (AT) and ArrayStrip™ (AS) platforms from Alere Technologies GmbH (Jena, Germany). These are much less expensive, and can be used without highly specialised equipment [8]. The AT (up to 225 spots) and AS (up to 600 spots per well) platforms make the use of a small array surface of size 4 x 4 mm placed on the bottom of a plastic vial or well. Hybridisation and analysis are simple and rapid, using standard laboratory methods, and hybridisation signals are detected following an enzyme-catalysed precipitation reaction [9]. The use of plastic tube-integrated arrays and fast non-fluorescent labelling and hybridisation protocols results in a system that is cost-effective, time saving, and allows high sample throughput, in a 96 well format [1].

There are currently no microarrays for the detection of multiple rodent-borne pathogens. The aim of this report is to explicate the design, development and evaluation of a microarray platform for use as a screening tool, which combines ease of use and rapid identification of a number of rodent-borne pathogens of zoonotic importance.

2.1Materials and Methods

2.1.1. Probe design

An initial literature search was performed to identify zoonotic pathogens which are transmissible by rodents. A microarray was then developed to screen for the presence of these pathogens. Table 1 shows the list of pathogens to be screened for, including the source of any reference material if available. Unfortunately, not all of the pathogens for which the array was designed to detect could be sourced. Therefore the probes for, Hepatitis E Virus, *Bartonella*, MRSA, *R. typhi* and *S. monilliformis* were not evaluated. RNA from an infected

Rattus norvegicus sample was supplied but several attempts at PCR proved unsuccessful and it was concluded that the RNA had degraded too much to be of use. A further literature search was conducted to identify particular genes or target regions which had been previously used for identification purposes in other diagnostic tests such as PCR. The DNA sequences were obtained from the NCBI database and aligned using ClustalX2 (http://www.ebi.ac.uk/Tools/phylogeny/clustalW2_phylogeny/help/faq.html#5) software. Oligonucleotide sequences (probes) were designed for each pathogen from regions targeted by species-specific or generic primers. Two freely available software packages were used for probe design: Unique Probe Selector (UPS) [10] and OligoWiz [11, 12]. Both types of software were used to compensate for any limitations in the other. OligoWiz, at present, can only be used to design probes for bacteria.

An optimal length of 60-nucleotide probes was assigned, and parameters for both OligoWiz and UPS included cross-hybridisation, delta-Tm, low-complexity, position and folding. An *in silico* analysis was performed on all the probes using the BLAST tool on the NCBI database to determine if cross-hybridisation would occur with any other known sequences. The selected probes were synthesised at Metabion International (Jena, Germany) with the following specifications: NH_2 modification at the 3' end, no modification at the 5' end, purification with HPLC, 0.04 µmol scale, and absolutely biotin-free.

The AT platform was used for initial evaluation for individual pathogens, and the best performing probes were transferred to the AS platform making a pool of probes from different pathogens. For both platforms, each probe was directly spotted onto the array surface at a 15 µM concentration with each probe printed in duplicate.

2.1.2. Primer design

Generic primers were designed from conserved flanking regions of the target sequence using Primer3 (http://primer3.ut.ee/). Species-specific primers were designed from more variable regions of a sequence that were specific to certain pathogens. The amplicon size was set between 250 and 750 bases, with an optimum of 500 bases. Primer sequences for both multiplex PCR and real-time PCR can be seen in Table 2.

Table 1.

A list of all the reference materials that were available for this study.

Table 2.

Primers used during the evaluation of the arrays. The majority were designed during the study but others were obtained either from the literature or colleagues.

2.1.3. Nucleic acid amplification

Several amplification methods were tested including sequence-independent amplification [13] using a random pentadecamer primer and a primer tag, and sequence-dependent amplification using pathogen-specific primers.

2.1.3.1. Sequence-independent amplification

Any RNA present in the sample was reverse transcribed into cDNA with 1.0 µl of primer A (GTT TCC CAG TCA CGA TCN NNN NNN NNN NNN NN) (40 µM), 1.0 µl of 10mM dNTP mix (Invitrogen), and variable amounts of water and template (minimum 50 ng/µl) were mixed in a PCR tube to a total volume of 13µl. The volume of water was variable to allow for different concentrations of template. This was then heated to 65°C for five min using a thermal cycler. The mixture was placed on ice for at least one minute. A separate mixture containing 4.0 µl of 5x Reverse Transcriptase Buffer (Invitrogen), 1.0 µl of 0.1M Dithiothreitol (DTT) (Invitrogen), 1.0 µl of RNase inhibitor, RNaseOUT (Invitrogen), and 1.0 µl of SuperScript III Reverse Transcriptase (Invitrogen) was added to the PCR tube

contents, and mixed by pipetting. The 20 µl reaction was incubated using a thermal cycler at 25°C for five min, then at 50°C for one hour and finally at 70°C for 15 min to inactivate the reaction. The reaction was left at room temperature for five min, followed by one minute on ice. The mix was then heated to 94 °C for two min, and rapidly cooled to 10°C in the thermal cycler for five min. 10 µl of Klenow mix (1.0 µl 10x Klenow buffer (Promega UK), 8.7 µl water, 0.3 µl Klenow polymerase (Promega)) was then added. For any DNA already present in the sample primer extension was effected with 1.0 µl Primer A (40µM), 1.0 µl 10x Klenow buffer and variable amounts of water and template (minimum 50ng/µl) to make a total volume of 10 µl. This sample mixture was then heated to 94°C for two min and then allowed to cool to 10°C in a thermal cycler for five min. The following 5.05 µl reaction mix was added to the sample mixture during its incubation at 10°C: 0.5 µl 10x Klenow buffer, 1.5 µl 3mM dNTPs, 0.75 µl 0.1M DTT, 1.5 µl 500 µg/ml BSA, 0.3 µl Klenow polymerase (Promega UK), 0.5 µl water. The reaction was left at room temperature for five min, followed by one minute on ice. The mix was then heated to 94 °C for two min, and rapidly cooled to 10°C in the thermal cycler for five min. 10 µl of Klenow mix (1.0 µl 10x Klenow buffer, 8.7 µl water, 0.3 µl Klenow polymerase) was then added.

For both RNA and DNA sequence-independent steps the mixture was then heated to 37°C for 8 min, and then held at 37°C for a further 8 min. This was followed by a rapid increase to 94°C for two min after which the mix was cooled to 10°C for five min, during which 1.2 µl of diluted Klenow (1:4) was added. The temperature was again increased to 37°C for 8 min followed by a hold of 8 min at 37°C, and then the reaction was terminated by placing the mixture on ice for 5 min. Standard PCR was then conducted using Primer B (GTT TCC CAG TCA CGA TC) (100 µM) to amplify the round A product with the following cycle parameters one step at 95°C for 10 s; 35 cycles of 30 s at 94°C, 30 s at 40°C, 30 s at 50°C, 2 min at 72°C and one final extension step of 72°C for 2 min. A 50 µl reaction mix was

155 prepared from the following: 39.0 µl water, 1.5 µl 50mM Magnesium chloride ($MgCl_2$)
156 (Invitrogen UK), 5.0 µl 10x Mg-free buffer (Invitrogen UK), 0.5 µl 25mM dNTP mix, 0.5 µl
157 Primer B, 0.5 µl *Taq* polymerase (5 U/µl) (Invitrogen UK) and 3.0 µl Round A product.
158 2.1.3.2. Sequence-specific amplification
159 Sequence-specific PCR was performed using a 50-µl reaction containing 37.5 µl of nuclease-
160 free water, 2.0 µl of 50mM $MgCl_2$ (Invitrogen UK), 5.0 µl of 10x Mg-free Buffer (Invitrogen
161 UK), 1.0 µl of 25mM dNTP mix, 1.0 µl of 10µM forward primer, 1.0 µl of 10µM reverse
162 primer, 0.5 µl of *Taq* DNA polymerase (5 U/µl) and 2.0 µl of cDNA or DNA (optimal
163 concentration 50ng/µl). Cycling parameters were one step of 94°C for 2 min; 30 cycles of 30
164 s at 94°C, 1 min at 60°C and 1 min at 72°C and one final extension step of 10 min at 72°C.
165 In addition the Qiagen Multiplex PCR *Plus* kit was tested with the sequence-dependent
166 primer sets. This was carried out with both non-biotinylated and biotinylated primers.
167 Multiplex PCR was performed using a 50-µl reaction containing 25 µl Multiplex Master Mix,
168 5 µl 10 x primer mix (2µM each primer) and variable volumes of water and template (50
169 ng/µl). The recommended protocol in the Qiagen Multiplex PCR *Plus* handbook was
170 followed with cycling parameters of one step at 95°C for 5 min; 40 cycles of 30 s at 95°C, 90
171 s at 60°C and 90 s at 72°C and one final extension step of 10 min at 68°C.
172 Real-time PCR was carried out using the Applied Biosystems 7500 Fast Real-Time PCR
173 System. Real-time PCR was performed using a 10-µl reaction containing 5 µl of TaqMan®
174 Universal PCR Master Mix 2x (Life Technologies), 1 µl of 300nM forward primer, 1 µl of
175 300 nM reverse primer, 1 µl TaqMan® probe (2.5µM), 1 µl of nuclease-free water and 1 µl
176 of sample (or water as a negative control). The recommended protocol was followed with
177 cycling parameters of one step at 50°C for two min, another step at 95°C for 10 min; 40
178 cycles of 95°C for 15s and 60°C for 60 s. Each sample was run in triplicate.
179 2.1.4. Microarray hybridisation

Prior to hybridisation of the labelled sample onto the array, the AS was conditioned by washing with 150 µl of water for 20 min at 30°C. After the water was removed using a pipette, a pre-hybridisation buffer (5x saline-sodium citrate (SSC), 0.1% sodium dodecyl sulphate (SDS), 4x Denhardt's solution) was pipetted into each well for 30 min at 50°C. Both washes were performed using a thermomixer (BioShake iQ, QUANTIFOIL Instruments GmbH, Jena Germany) at 550 rpm, which was used in all subsequent incubation steps unless otherwise stated. A 10-µl aliquot of the biotin-labelled sample was added to 90 µl of hybridisation buffer (5x SSC, 1% SDS, 4x Denhardt's solution). The mixture was denatured at 95°C for 3 min and then kept on ice. The denatured sample (100 µl) was then pipetted into the AS well and allowed to hybridise for 30 min at 55°C at 550 rpm. The sample solution was then removed and the AS was washed successively for 20 min at 60°C at 550 rpm with 150 µl wash buffer 1 (1x SSC, 0.2% SDS), wash buffer 2 (0.1x SSC, 0.2% SDS), and wash buffer 3 (0.1x SSC). This buffer was then removed and vacant binding sites on the microarray were blocked by incubation with a blocking solution (100 µl) of 2% biotin-free milk in PBS containing 1% bovine serum albumin (BSA) and 0.1% Tween™ 20 for 60 min at 30°C at 300 rpm. The blocking solution was replaced with 100 µl conjugation solution (Streptavidin Poly-Horseradish peroxidase (HRP) diluted 1:100 in the blocking solution), and the array incubated for 15 min at 30°C at 300 rpm. Post-conjugation washes were performed using wash buffers 1-3 as described for the post-hybridisation washes. After removal of wash buffer 3, 100 µl of a tetramethylbenzidine (TMB)-hydrogen peroxide (H_2O_2) solution, in this instance TrueBlue™ (Insight BioTechnology LTD, UK), was added and incubated for 10 min at 25°C without shaking. After removing the solution, the AS was then inserted into the ArrayMate and the array image was recorded with raw data generated. The recorded image was analysed using Alere's integrated IconoClust software and analysis script. Iconoclust processes the signals and automatically normalises the signal value after an

algorithm processes the average intensity of the spot and the local background noise. The output range of the signals was between 0 and 1, with 0 being negative and 1 being the maximal possible signal value. The normalised intensity of the spots was automatically calculated by subtracting the local background noise from the average intensity of the automatically recognised spot.

3.1. Results

3.1.1. PCR amplification and hybridisation

A 327 probe ArrayStrip was produced, and the number of probes per pathogen are given in Table 3. During the evaluation stage, it was determined that sequence-independent amplification resulted in lower hybridisation signals on the array than sequence-dependent amplification. Some of the pathogens (e.g. Cowpox, *T. gondii* and *C. jejuni*), when amplified by their specific primers produced good quality, detectable, hybridisation signals, but when random amplification was used, they showed no or weak hybridisation. A DNA sample of *C. jejuni* for example was amplified using sequence-independent PCR and the product was then hybridised onto the array. A measurable signal was seen with 26.6% of the *C. jejuni* probes on the array. With specific amplification there was 100% probe hybridisation at significantly higher signal strength (data not shown).

Table 3.

The number of probes for each pathogen that were spotted on the WT_Rodent_Chip_03 ArrayStrip

Figures 1A-1H show the images recorded after hybridisation with a variety of *Salmonella* species following sequence-dependent and sequence-independent amplification. The spots indicated by arrows are the biotin markers, which act as assay controls. Numerous probes showed cross-hybridisation in these images, albeit at low signal intensities, so they were

removed from the final version of the array. It is also apparent that the cross-hybridising probes were only visible in the images which show samples that had been amplified using sequence independent-amplification (Figures 1A-1F). The images which show hybridisation following sequence-specific amplification were much cleaner (Figures 1G-1H), and had the expected hybridisation profile. The three spots indicated by the rectangular box in all images except Figure 1G are probes that were designed to hybridise with a wide range of *Salmonella* species. Figure 1H shows the amplification of *S.* Typhimurium with a set of primers designed to amplify this region which is common to multiple *Salmonella* species. *S.* Gallinarum (Figure 1A), *S.* Dublin (Figure 1B), *S.* Pullorum (Figure 1C), *S.* Enteritidis (Figure 1D), *S.* Hadar (Figure 1E), and *S.* Typhimurium (Figures 1F-1H) were tested on the array. Although it was visible on the array, the signal strength is low in comparison to Figure 1H. This is particularly noticeable for *S.* Dublin and *S.* Hadar from Figures 1B and 1E.

Both sequence-independent and sequence-specific amplification were used for *S.* Typhimurium. The images produced after hybridisation can be seen in Figures 1F and 1G. Figure 1F shows sequence-independent amplification, and the circled probes were designed to be specific for *S.* Typhimurium. This set of probes did not show any detectable hybridisation in the other images so it appeared these were good probes for distinguishing *S.* Typhimurium from other *Salmonella* species. Figure 1G shows the sequence-specific amplification with a set of primers designed to amplify the *S.* Typhimurium-specific region. The probes for all of the pathogens tested which produced a hybridisation signal can be seen in Table A.1 in the Appendix.

Figure 1.

Images produced after hybridisation of various *Salmonella* species on WT_Rodents_2_1.0 array. The spots indicated by arrows are the biotin markers. The solid square and rectangular areas are the orientation markers.

A. *S.* Gallinarum hybridisation following random amplification
B. *S.* Dublin hybridisation following random amplification
C. *S.* Pullorum hybridisation following random amplification
D. *S.* Enteritidis hybridisation following random amplification
E. *S.* Hadar hybridisation following random amplification
F. *S.* Typhimurium amplification following random amplification
G. *S.* Typhimurium amplification with primers Salm/flag/1366055/F and Salm/flag/1366482/R (*S.* Typhimurium-specific)
H. *S.* Typhimurium amplification with primers Salm/CDP/2167279/F and Salm/CDP/2005357/R (Generic *Salmonella* species)

Table 4.

The negative control sample and reference pathogen samples in lanes 1-12 from Figure 2A.

3.1.2. Multiplex PCR amplification and hybridisation

The primer sets which performed well in singleplex PCR reactions were then tested in a multiplex reaction. As it is unlikely that a sample would contain all of the pathogens tested, the effectiveness of the primer mix in detecting a pathogen was tested using a sample of rodent liver DNA which was spiked with individual pathogen DNA (DNA concentration ranged from 1.66 – 112.5 ng/µL, and copy number from $2.33 \times 10^{9} – 2.09 \times 10^{11}$). Figure 2A shows a gel image of the result of amplification of individual pathogens from the spiked material when the multi-pathogen primer mix was used. Table 4 shows the pathogen detected in each lane from Figure 2A. As can be seen from the figure, the majority of lanes had a strong band. The two bands in Lane 9 represent the specific *S.* Typhimurium amplicons (663 bp) and the generic *Salmonella* amplicons (428 bp). This was expected, as the multi-pathogen primer mix had primers specific for *S.* Typhimurium, and also had primers to amplify a region common to many *Salmonella* species. Figure 2B shows the spots

283 that hybridised after using the multiplex primer mix (with biotinylated primers) on a pooled
284 nucleic acid sample from all pathogens for which reference samples were available.
285 Although the band seen in Lane 12 for *T. gondii* was quite faint, careful analysis of the
286 recorded image indicated that all of the pathogens in the sample, including *T. gondii*,
287 hybridised with the expected specific probes on the array.

288 Figure 2.

289 A. Agarose gel electrophoresis image produced after amplification of nucleic acid of
290 individual pathogens using the multiplex primer mix.

291 B. Profile produced after hybridisation of a mixture of all the pathogens following
292 amplification with the multiplex primer mix for which reference samples were available.

293 3.1.3. Array sensitivity testing

294 The sensitivity of the array was tested by performing real-time PCR using serially diluted *Y.*
295 *pestis* DNA. The pathogenic DNA in the sample was no longer detectable using real-time
296 PCR (Figure 3) at copy numbers less than 4.39×10^2. As expected, there was no
297 amplification for the negative control sample. Samples in Figure 3B, C, D and E (copy
298 number 3.47×10^9, 1.76×10^7, 8.57×10^4, 4.39×10^2) were detectable by real-time PCR. The
299 DNA in sample B was detectable after 18 cycles. For samples C, D, and E the cycle number
300 at which detection occurred was 24, 32 and 36, respectively. The DNA in sample F appeared
301 to have been too low for real-time PCR to detect and no amplification was observed.

302 The array images shown in Figure 3 were produced after hybridisation of the products of
303 standard PCR amplification. These were the same samples that were tested by real-time PCR
304 for *Y. pestis* on the Yersinia_01 ArrayTube. The biotin markers on each array are indicated
305 with an arrow. On the Yersinia_01 array only two probes were expected to hybridise with
306 the primer set used (Y.pes/pPCP/8374/F Y.pes/pPCP/8902/R). As can be seen from the
307 images produced after hybridisation, the two expected probes hybridised with samples B, C
308 and D. For samples E and F there was no apparent hybridisation.

309 Figure 3.

310 Real-time PCR sensitivity testing was performed with serial dilutions of *Y. pestis*. Sample A
311 was a negative control sample (water). The copy number in samples B to F was $3.47x10^9$,
312 $1.76x10^7$, $8.57x10^4$, $4.39x10^2$, and 1. Array images after hybridisation of the standard PCR
313 products from the same *Y. pestis* amplification on WT_Yersinia_01 are also shown. Biotin
314 markers are indicated by arrows, and *Y. pestis* probes that showed hybridisation are circled.

315 4.1. Discussion

316 Collecting good quality samples for disease surveillance can often be a time and cost
317 intensive process. Therefore, it is important that any technology used is as efficient as
318 possible. We report the development of a DNA microarray for simultaneous detection of
319 multiple pathogens of rodents, comprised of 327 probes derived from a variety of genes in
320 each of the target pathogens. The technology presented here represents a simple but effective
321 system which is affordable and compatible with standard laboratory equipment, and has been
322 used for a variety of purposes over recent years [1, 14-18].

323 The design of oligonucleotide probes is a complex process for a variety of reasons, including
324 identifying the best target sequences to be screened and understanding the thermodynamics of
325 probe-target interactions during hybridisation [19]. An oligonucleotide length of 60 bp was
326 selected, as several studies have indicated that this offers the best combination between
327 specificity and sensitivity [20-22]. Shorter oligonucleotide probes (15-25 mers) have a very
328 high specificity, but they have been criticised for having a lack of sensitivity, whereas longer
329 oligonucleotides (40-90 mers) are thought to have a good sensitivity whilst maintaining a
330 high specificity [19]. It has been reported that 60-mer oligonucleotides have an eight-fold
331 higher sensitivity than 25-mers [23].

332 Variation within microarray experiments can still occur regardless of careful probe design.
333 The typical sources of variation can be broadly divided into three main categories: variation
334 within the biological sample, the performance of the technology itself, and finally, variation
335 in the spot signal measurements. The majority of variance in microarray experiments is
336 generally biological rather than technical [24]. Arrays are generally made in batches and
337 variation can occur between batches. These can include different probe concentrations,
338 which can lead to incorrect conclusions being drawn from data [25]. Variation at the array

production stage can occur for a number of reasons including, for example the particular printing pin used, the humidity, and temperature during printing. These variables can lead to slight differences in the amount of probe that is deposited on the slide surface, the amount that remains on the array surface after processing, and the level of deviation from the expected spot location. All of these factors can have an impact on the amount of labelled target that can bind to the probe, and on the efficiency of subsequent spot finding and data extraction steps [19]. The level of deviation from the expected location can result in the array reader making inaccurate readings, the signals of neighbouring probes becoming merged, or the spot can become invalid and cannot be accurately detected by the analysis software. In order to reduce the variability that is inherent in all biological experiments, experimental replication is essential. One obvious form of technical replication is through array probe replication. It is advantageous to at least have duplicates, or preferably multiples, of all probes spotted on the same array, however this may not be possible due to spotting density constraints. The precision of particular probe measurements will be more reliable if the spot intensities of the replicate spots are averaged for each sample [26].

To achieve an efficient hybridisation step, it is important to have probes with a narrow melting temperature distribution, because the hybridisation step takes place at the same temperature for all probes on the array [27]. The algorithms used in both OligoWiz and UPS are able to make minor adjustments to the length of each probe so that a narrow melting temperature range is achieved. Determination of melting temperature thresholds is a difficult task as this determines the conditions under which probes will bind to the target sequence. Melting temperatures can cause loss of signal if too high, and non-specific signal if too low [28]. As a single temperature is used during the hybridisation step, it is advisable that the narrowest melting temperature range be used to maximise signal detection [29].

363 The chip presented here represents a prototype microarray identification system for zoonotic
364 pathogens that can infect rodents. The probes used on the microarray were based on genes
365 that are unique to the pathogens selected. These genes were selected following a literature
366 search to identify gene sequences which have been previously used to identify these
367 pathogens, and a BLAST analysis to see if the sequences selected had similarity to any other
368 pathogen sequences on the database. This is also the first report of biotinylated primers used
369 in a multiplex format with up to 24 primer pairs. There was no apparent difference in the
370 hybridisation signal produced when only a single pathogen was present in single and
371 multiplex PCR reactions. More work needs to be done to determine the limits of detecteion
372 and the sensitivity of the array, but as a proof of concept the array has demonstrated potential.
373 Further improvements to this array could be made by obtaining reference material for
374 pathogens which were unavailable and evaluating the probes for these pathogens. Whilst
375 reference material was available for SEOV during the evaluation stage, several attempts at
376 PCR proved unsuccessful. However, a number of rodent samples (nucleic acid was extracted
377 from liver, kidney and lung of *R. rattus* and *R. norvegicus*, an aliquot of which was then
378 pooled and amplified by multiplex PCR followed by hybridisation with the microarray) were
379 screened on the array and two of these were identified as SEOV positive. This was later
380 verified by both PCR and sequencing.

381 While *in silico* analysis of the gene and resulting probe sequences are important in
382 eliminating the possibility of cross-hybridisation with other sequences already on the NCBI
383 database, it does not rule out the possibility of cross-hybridisation with newly emerging
384 organisms for which the gene sequence is unknown. As a result, microarrays can be used to
385 identify novel as well as known pathogens. This can be achieved by designing probes at a
386 genus level with additional probes designed for differentiating between species [30].

The presence of host nucleic acid in a sample presents another challenge in microarray experiments, as it can lower the sensitivity of the array. This occurs because in most situations the host DNA is present in much higher amounts than the pathogen nucleic acid, which makes the pathogen more difficult to detect. The sensitivity of an array may be improved by the removal of host nucleic acid by DNAses, i.e. enriching pathogen-derived nucleic acid, using dedicated methods and kits for this purpose prior to PCR amplification [31]. The sensitivity on this array was less than that of real-time PCR, as has been previously demonstrated with other pathogen detecting microarrays [14, 32], this leads to a trade-off between sensitivity and cost-effectiveness. It would have been useful to have tested the array sensitivity for RNA pathogens as well. However, as mentioned earlier, there were no working RNA pathogens available during the evaluation stage. A critical step in the development of a microarray is sourcing reference samples with which the array can be evaluated.

The design stage can be the deciding factor in the success of any microarray experiment and the choice of array platforms or probe types can be challenging. However, it is now becoming increasingly clear that when a careful design is followed, the results obtained with different platforms are likely to be comparable [19, 33, 34]. The user can decide whether to invest time and resources in developing their own arrays, utilise one of the commercial providers who can assist with array design and fabrication, or use off-the-shelf commercial arrays. The relatively low cost of screening for many pathogens simultaneously in a single sample is an economical and efficient approach for rapid and sensitive diagnostics. This may be of particular use for wildlife samples which may be small in volume and are often irreplaceable.

Acknowledgements

This research has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 222633. We would like to thank Chelsea Himsworth, Florence Ayral, and Kieran Pounder for providing the tissue samples. The authors would also like to thank those who contributed reference samples for the evaluation stage (Konrad Sachse, Brendan Wren, Karen Kempsall, Zoree Djelouadji, Michael Jones, Sean McAteer, David Gally, Margaret Lovell, Sylvia Grierson, Nicole Pavio, Charalambos Bilinis, Fimme van der Wal, Public Health England and the European Virus Archive).

References

1. Sachse, K., et al., *DNA microarray-based detection and identification of Chlamydia and Chlamydophila spp.* Molecular and Cellular Probes, 2005. **19**(1): p. 41-50.
2. Taylor, L.H., S.M. Latham, and M.E.J. Woolhouse, *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences.* The Royal Society, 2001. **356**(1411): p. 983-989.
3. Russell, L.B., *Is Prevention better than Cure?* 1986, Washington D.C.: The Brookings Institution.
4. Dong, J., et al., *Emergin Pathogens: Challenges and Successes of Molecular Diagnostics.* Journal of Molecular Diagnostics, 2008. **10**(3): p. 185-197.
5. Wang, D., et al., *Viral Discovery and Sequence Recovery Using DNA Microarrays.* PLoS Biology, 2003. **1**(2): p. e2.
6. Wang, L.C., et al., *Simultaneous detection and differentiation of Newcastle disease and avian influenza viruses using oligonucleotide microarrays.* Veterinary Microbiology, 2008. **127**: p. 217-226.
7. Jack, P.J.M., et al., *Microarray-based detection of viruses causing vesicular or vesicular-like lesions in livestock animals.* Veterinary Microbiology, 2009. **133**: p. 145-153.
8. Ehricht, R., et al., *Optimized DNA microarray assay allowss detection and genotyping of single PCR-amplifiable target copies.* Molecular and Cellular Probes, 2006. **20**(1): p. 60-63.
9. Müller, R., et al., *Detection of herpesvirus and adenovirus co-infections with diagnostic DNA-microarrays.* Journal of Virological methods, 2009. **155**(2): p. 161-166.
10. Chen, S.H., et al., *UPS 2.0: unique probe selector for probe design and oligonucleotide microarrays at the pangenomic/genomic level.* BMC Genomics, 2010. **11**(4).
11. Wernersson, R. and H.B. Nielsen, *OligoWiz 2.0 - integrating sequence feature annotation into the design of microarray probes.* Nucleic Acids Research, 2005. **33**: p. 611-615.
12. Wernersson, R., A.S. Juncker, and H.B. Nielsen, *Probe selection for DNA microarrays using OligoWiz.* Nature Protocols, 2007. **2**(11): p. 2677-2691.

13. Bohlander, S.K., et al., *A method for the rapid sequence-independent amplification of microdissected chromosomal material.* Genomics, 1992. **13**(4): p. 1322-1324.
14. Dean, D., et al., *Zoonotic Chlamydiaceae Species Associated with Trachoma, Nepal.* Emerging Infectious Diseases, 2013. **19**(12): p. 1948-1955.
15. Ehricht, R., et al., *Optimised DNA Microarray assay allows detection and genotyping of single PCR-amplifiable target copies.* Molecular and Cellular Probes, 2006. **20**: p. 60-63.
16. Ruettger, A., et al., *Genotyping of Chlamydia trachomatis strains from culture and clinical samples using an ompA-based DNA microarray assay.* Molecular and Cellular Probes, 2011. **25**(1): p. 19-27.
17. Ruettger, A., et al., *Rapid Spoligotyping of Mycobacterium tuberculosis Complex Bacteria by Use of a Microarray System with Automatic Data Processing and Assignment.* Journal of Clinical Microbiology, 2012. **50**(7): p. 2492-2495.
18. Sachse, K., et al., *More than classical Chlamydia psittaci in urban pigeons.* Veterinary Microbiology, 2012. **157**(3-4): p. 476-480.
19. Russell, S., K.A. Meadows, and R.R. Russel, *Microarray Technology in Practice.* 2009, Cambridge: Academic Press.
20. Relogio, A., et al., *Optimisation of oligonucleotide-based DNA microarrays.* Nucleic Acids Research, 2002. **30**(e51).
21. Kane, M.D., et al., *Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays.* Nucleic Acids Research, 2000. **28**: p. 4552-4557.
22. Wang, H.Y., et al., *Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays.* Genome Biology, 2003a. **4**(1): p. R5.
23. Chou, C.C., et al., *Optimisation of probe length and the number of probes per gene for optimal microarray analysis of gene expression.* Nucleic Acids Research, 2004. **32**(12): p. e99.
24. Koskinen, M.T., et al., *Analytical specificity and sensitivity of a real-time polymerase chain reaction assay for identification of bovine mastitis pathogens.* Journal of Dairy Science, 2009. **92**(3): p. 952-959.
25. Leek, J.T., Scharpf, R.B., Bravo, H.C., Simcha, D., Langmead, B., Johnson, W.E., Geman, D., Baggerly, K., Irizarry, R.A., *Tackling the widespread and critical impact of batch effects in high-throughout data.* Nature Reviews Genetics, 2010. **11**: p. 733-739.
26. Tran, P.H., et al., *Microarray optimisations: increasing spot accuracy and automated identification of true microarray signals.* Nucleic Acids Research, 2002. **30**(12): p. e54.
27. Haque, R., et al., *Multiplex real-time PCR assay for detection of Entamoeba histolytica, Giardia intestinalis and Cryptosporidium spp.* The American Journal of Tropical Medicine and Hygiene, 2007. **76**(4): p. 713-717.
28. Blalock, E., *A Beginner's Guide to Microarrays*, ed. E. Blalock. 2003, Norwell, Massachussetts: Kluwer Academic Publishers.
29. Lemoine, S., F. Combes, and S. Le Crom, *An evaluation of custom microarray applications: the oligonucleotide design challenge.* Nucleic Acids Research, 2009. **37**(6): p. 1726-1739.
30. Wang, D., et al., *Viral Discovery and Sequence Recovery Using DNA Microarrays.* Public Library of Science Biology, 2003. **1**(2): p. 257-260.
31. Gurrala, R., et al., *Development of a DNA microarray for simultaneous detection and genotyping of lyssaviruses.* Virus Research, 2009. **144**(1-2): p. 202-208.

32. Borel, N., et al., *Direct identification of chlamydiae from clinical samples using a DNA microarray assay-A validation study.* Molecular and Cellular Probes, 2008. **22**(1): p. 55-64.
33. Jarvinen, A.K., et al., *Are data from different gene expression microarray platforms comparable?* Genomics, 2004. **83**(6): p. 1164-1168.
34. Irizarry, R.A., et al., *Multiple-laboratory comparison of microarray platforms.* Nature Methods, 2005. **2**: p. 345-350.