Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

Microarray systems for microbial detection and identification

B. N. Ford and D.E. Bader
DRDC Suffield

Defence R&D Canada

Technical Report

DRDC Suffield TR 2010-203

December 2010

Canada

Microarray systems for microbial detection and identification

B.N. Ford and D.E. Bader
DRDC Suffield

Defence R&D Canada – Suffield

Technical Report
DRDC Suffield TR 2010-203
December 2010

Principal Author

Original Signed by

Barry Ford

Defence Scientist

Approved by

Original Signed by

Stephen G. Bjarnason

Head, Casualty Management Section

Approved for release by

Original Signed by

Robin Clewley

A/Chair, Document Review Panel

This work funded by TIF Project 10DA09

Abstract

The relatively recent availability of complete or near-complete sequences of microbial DNA makes it possible to find and exploit unique genetic markers for most microbes. Using microbial genome sequences from public databases, a high-density microarray system has been developed which should enable the identification of hundreds of individual species or strains of pathogenic microorganisms on a single assay platform. This report summarizes the design, development, and testing of this microarray design. The system is currently capable of discriminating multiple human pathogens, Category A biothreats, and some agricultural pathogens, other using supervised analysis methods. Ongoing development will include tools for automation of analysis, and extension of the testing panel of microbes.

Résumé

La disponibilité relativement récente de séquences complètes ou quasi complètes d'ADN microbien nous permet de trouver et d'utiliser des marqueurs génétiques uniques pour la plupart des microorganismes. En utilisant les séquences génomiques microbiennes des bases de données, nous avons mis au point un système de biopuce à haute densité qui devrait permettre d'identifier des centaines d'espèces ou de souches de microorganismes pathogènes avec une seule et même plateforme d'essai. Nous résumons ici la conception, la mise au point et la mise à l'épreuve de cette biopuce. À l'heure actuelle, le système permet de distinguer entre de nombreux agents pathogènes pour l'humain, des bactéries de la catégorie A des agents de bioterrorisme et certains pathogènes agricoles. Le perfectionnement de la plateforme comprendra des outils pour l'automatisation des analyses et l'augmentation du nombre de microorganismes pouvant être identifiés.

This page intentionally left blank.

Executive summary

Microarray systems for microbial detection and identification:

Ford, Barry N.; Bader, Doug, B.; DRDC Suffield TR 2010-203; Defence R&D Canada – Suffield; October 2010.

Background: The relatively recent availability of complete or near-complete sequences of genomic DNA from many microorganisms makes it possible to find and exploit unique genetic markers for most microbes. These unique markers can be useful in detection and identification of microbial species, to the strain level in many cases. One tool for exploiting the availablility of genetic sequence data is the DNA microarray. Microarrays are high density assay platforms which could be a useful solution to the problem of large scale screening and identification of microbial ssamples. Combining publicly available complete or near-complete genomic sequences from many microbial species and isolates, a microarray platform has been designed which should be able to identify many Category A biothreat, agricultural, and clinically relevant bacterial pathogens. A potential advantage of the complete microarray system is the reduced need to handle and culture live pathogens, as is required in conventional microbiology.

Results: Evaluation of the microarray with more than seventy genomic DNA isolates from a wide range of pathogenic microbes demonstrated the capability of the array to discriminate microbial species with high confidence. Within some species, strain differentiation was possible. Sample requirements and cost required the co-development of a DNA amplification technology to pre-process DNA samples, which has proven very effective.

Significance: The microarray represents a tool for identification of bioterrorism agents, water-borne or food pathogens, and selected clinical or agricultural pathogens. Current microarray technology requires skilled users and significant experience with analysis. Near-term technological developments will involve advanced automation of microarray work which will simplify the technical requirements. The microarray platform enables large-scale screening which will be useful in suspicious disease outbreaks, new pathogen detection, and monitoring spread of biological agents.

Future plans: The large data set developed during this project will require some ongoing analysis. Software to facilitate automated analysis of microarray data would represent a significant improvement. Future development of the microarray system may involve expanding the representation of agricultural and other clinical targets, including viral pathogens. Interest in exploiting the existing design has been expressed by other government departments.

Sommaire

Microarray systems for microbial detection and identification:

Ford, Barry N.; Bader, Doug, B.; DRDC Suffield TR 2010-203; R & D pour la défense Canada – Suffield; Octobre 2010.

Introduction ou contexte. La disponibilité relativement récente de séquences complètes ou quasi complètes de séquences d'ADN génomique de nombreux microorganismes nous permet de trouver et d'utiliser des marqueurs génétiques uniques pour la plupart des microorganismes. Ces marqueurs uniques peuvent servir à détecter et à identifier des espèces microbiennes, et dans de nombreux cas, ils permettent même d'identifier la souche. Un des outils d'exploitation des séquences génétiques est la biopuce à ADN. Les biopuces sont des plateformes d'essai à haute densité qui peuvent se révéler fort utiles dans le dépistage à grande échelle et l'identification de microorganismes. En utilisant les séquences génomiques complètes et quasi complètes des espèces et des isolats de microorganismes disponibles dans les banques de données publiques, nous avons conçu une biopuce qui devrait permettre d'identifier de nombreuses bactéries de la catégorie A des agents de bioterrorisme, des pathogènes agricoles et d'autres ayant une importance clinique. Un des principaux avantages du système de biopuce est qu'il permet de réduire la manipulation et la nécessité de cultiver des pathogènes vivants, comme le requièrent les méthodes classiques de microbiologie.

Résultats. L'évaluation de la biopuce avec plus de soixante-dix isolats d'ADN génomique de microorganismes pathogènes a permis de démontrer la capacité de discrimination de cet outil et sa fiabilité. Chez certaines espèces, des souches peuvent même être différenciées. Les besoins en matière de quantité d'échantillon et les coûts ont nécessité l'élaboration concomitante d'une technologie d'amplification de l'ADN pour prétraiter les échantillons, ce qui s'est révélé très efficace.

Importance. La biopuce permet d'identifier des agents de bioterrorisme, des pathogènes dans l'eau ou les aliments et certains pathogènes d'importance clinique et agricole. La technologie actuelle des biopuces requiert des utilisateurs habiles et une grande expérience dans l'analyse des résultats qu'elles génèrent. L'étape suivante du projet prévoit l'automatisation des manipulations, ce qui simplifiera l'utilisation de l'outil et les compétences techniques requises. La plateforme permet le dépistage à grande échelle qui sera utile dans le cas d'éclosions soupçonnées de maladies, de la détection de nouveaux pathogènes et de la surveillance de la prolifération ou de la dissémination d'agents biologiques.

Perspectives. De nombreuses données produites au cours de ce projet doivent encore être analysées. L'existence d'un logiciel pour faciliter l'analyse automatisée des données obtenues avec la biopuce représenterait une amélioration considérable. Le perfectionnement du système de biopuce pourrait comprendre un plus grand nombre de cibles agricoles ou cliniques, ainsi que des pathogènes viraux. Enfin, d'autres organismes gouvernementaux ont manifesté leur intérêt à utiliser la plateforme que nous mise au point.

Table of contents

List of figures

List of tables

Acknowledgements

The authors would like to thank Dr. Kingsley Amoako (CFIA Lethbridge) for providing bacterial DNA and sequence suggestions from various strains used in this study. Ms Cindy Ruttan from Canada West Biosciences performed most of the Affymetrix microarray processing under the management of Dr Dave Mah. Because this report is intended to summarize prior works, some figures and data in this report have been excerpted from previously published technical reports by the same authors.

The operating costs of this work were paid from a 2007 Technology Investment Fund award to Mr Doug Bader, 10DA09 - Microarray Technology For Multi-Gene Target Surveillance And Identification Of Biothreats.

This page intentionally left blank.

Introduction

Microbial genomes range widely in size and complexity. Microbiological and immunoassay technologies are widely applied in detection and identification of and Current state of the art detection and identification is primarily based on molecular assays which detect specific parts of the DNA in the genome. Ideally, for each organism of interest, one or more unique sequences are targeted which can be assayed using one of several methods. There is a wide selection of individual assay systems and methods available, a review of which is beyond the scope of this work. Invariably however, such methods require an index of suspicion as to which organism to test for, since the assay systems in general are able to identify only one or a few molecular targets in each assay run. An improvement in assay coverage wherein a single assay run could identify any of tens to hundreds of individual molecular targets would be advantageous. This report summarizes efforts to develop a solution to the multiplex-targets-for-multiple-organisms problem. The work reported was largely funded by a 2007 Technology Investment Fund award to Mr Doug Bader, 10DA09 - Microarray Technology for Multi-Gene Target Surveillance and Identification of Biothreats. It should be noted that low-specificity detection systems already exist, but high-specificity identification systems with high density target capability are lacking.

High-density DNA microarray systems offer this capability. Microarrays are essentially a tool to assay (test) a great many individual DNA targets in one interrogation. This approach has several general advantages. For example, a single microarray platform can have broad species and strain coverage potential. As well, microarrays offer the ability to do simultaneous multiple confirmatory assays within species and strains. The original program to investigate extant microarrays for the purposes of detection and identification had to be revised when microarrays specific to the project needs were either not suitable or did not become commercially available. Consequently a custom fabricated microarray was developed for this project.

The basis of microarray technology is shown in Figure 1. A sample of unknown DNA is labeled with a fluorescent dye using conventional techniques. The dye enables detection of the sample DNA using a microscope equipped with fluorescence detection capabilities. Extraction and purification of the DNA prior to labeling is required. The labeled DNA sample is mixed in a special chamber with the microarray. The microarray contains a set of “features” consisting of fragments of DNA (target DNA) attached to a solid surface, such as a microscope slide. There are multiple technologies for the preparation of the microarray itself, which are reviewed in Miller and Tang [1]. During the mixing of the sample with the microarray, fragments of the sample DNA hybridize to their complementary sequence on the microarray. The property of DNA fragments in solution coming into contact with and binding to their complementary sequence (hybridization or reassociation) is the basis of many molecular assays, including microarrays.

A key property of DNA which enables microarray and other DNA-dependent assay systems to function is the property of self-complementarily. In most organisms, the genetic material is present as double-stranded DNA. Double-stranded molecules contain two complementary copies of the DNA sequence, which in turn are comprised of linear chains of the DNA bases, adenine, cytosine, guanine, and thymine (A,C,G and T respectively). Between the single strands of the DNA, A binds to T and T to A. In the same way, C and G are complementary to each other. Overall, the two linear single chains which interact to form the double-stranded DNA molecule are complementary to each other. Each strand contains a complete "copy" of the information on the other strand. Figure 2 illustrates the principle of complementary binding of DNA fragments.

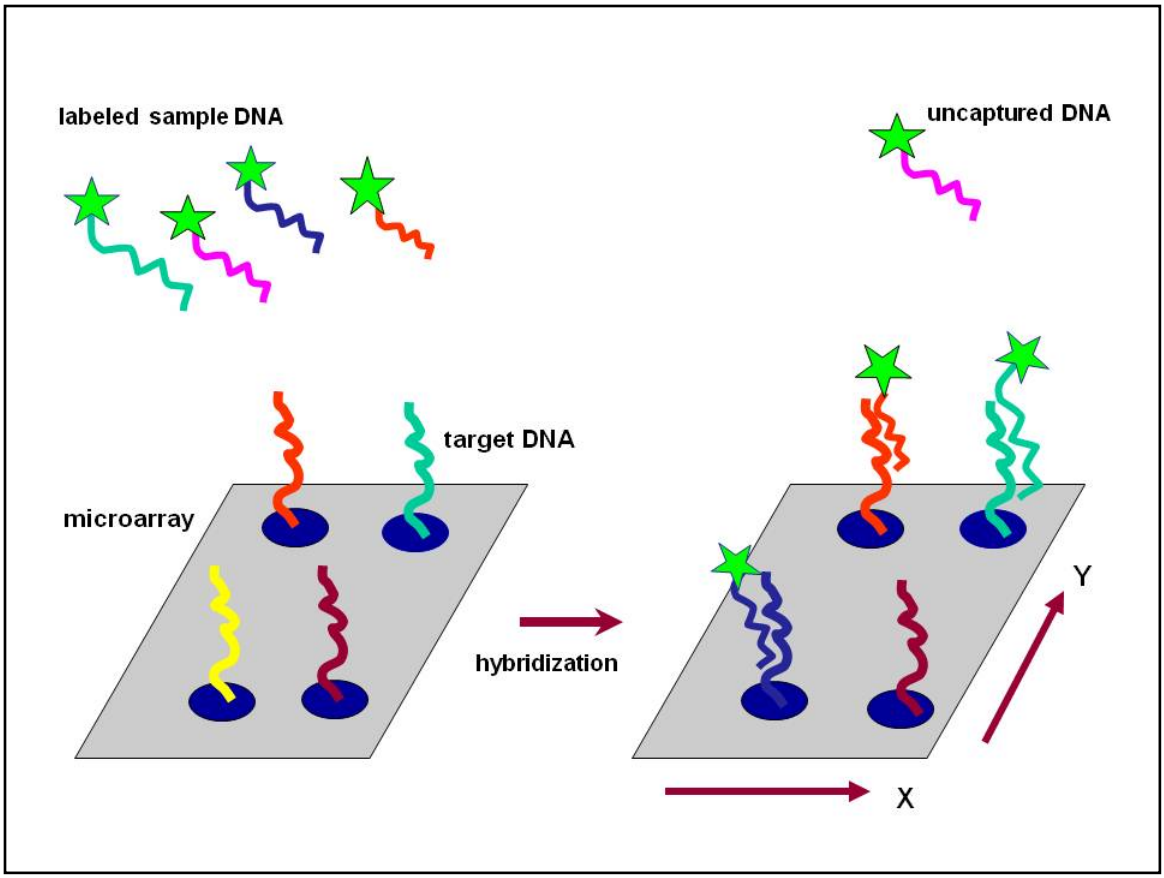

Figure 1: Generalized microarray principle.

Figure 2: DNA complementarity

In solution, after the individual linear strands are separated and fragmented, fragments reassociate to their complement by diffusion, and bind to it, reforming the double stranded DNA molecule. The rate of reassociation is related to the concentration of the fragments, and inversely related to the size of the fragments. On a microarray chip, short fragments (25 bases in this case) are present at high local concentrations in the microarray spot, such that reassociation of fragments in solution to the fragments on the array occurs relatively quickly compared to reassociation of fragments to their genomic complement. Thus the features on the microarray surface bind fragments from solution and a detectable hybridization signal is developed. Figure 3 illustrates

the system after hybridization of the sample to the microarray. The microarray with captured sample DNA is scanned by a microscope, and the digital image is stored for analysis.

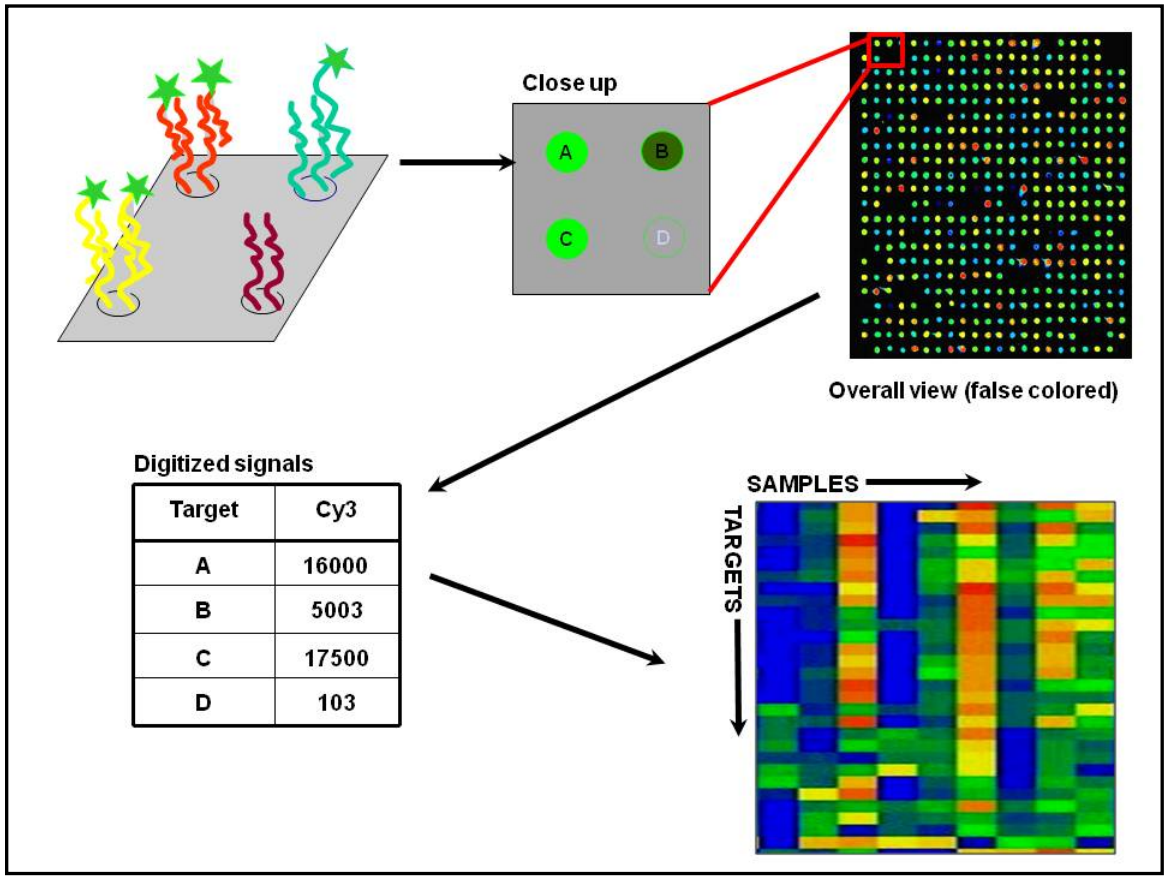

Figure 3: Overall microarray system.

Existing microarray systems

At DRDC Suffield, two previous microarrays have been designed and developed on open source microscope slide formats. The results of this work have been summarized in a contractor report [2]. Using a small number of species and strains, and using open source microarrays containing either 23,000 non-specific or 5280 specific probes, this work demonstrated the proof of concept of microarrays for microbial detection and identification. The limited number of features that could be maximally placed on these arrays, combined with the intrinsic variation in the microarray spotting technology, restricted the utility of the systems to species discrimination.

Microarray systems for microbial genome analysis prior to this project have been reported. Some have used long oligonucleotide fragments which are intended to capture very specific targets. A few have attempted to use short "random" oligonucleotides, again for a relatively limited selection of targets [3–5]. A recent theoretical paper argues that it should be possible to unambiguously characterize almost all known prokaryotes and eukaryotes using approximately

two million features based on "random" oligonucleotides [6]. This approach has not been tested on an actual platform. A microarray sequencing approach has also been published for a set of food-borne pathogens, but it was not designed with sufficient scope to be able to identify Category A pathogens [7]. A microarray-based viral detection and genotyping system was developed previously at the University of California for a limited number of human viral pathogens, but is no longer available for testing [8]. This array, which used 70 base oligonucleotides on ~1600 features, was able to identify ~140 distinct viral genomes. A revision to the design eventually had 22,000 features encompassing a wide variety of rare and difficult to diagnose viral forms, and ~30 bacterial genome fragments. This chip design is now in use at Abbot Laboratories in a virus discovery program in collaboration with the University of California, and is apparently not available for distribution.

An open source chip from the Health Protection Agency (UK) has been tested for comparison to the current and previous chip designs. The HPA design contains some 9248 features nominally designed for discrimination of various pathogens[1]. The chip was amenable to sample and chip processing analysis in exactly the same way as the in-house designs. Unfortunately, the HPA chip was not accompanied by adequate annotation, such that the intended targets of the chip could not be compared to the recorded data. On a qualitative basis, species could be discriminated by the HPA chip based on differential hybridization. In the absence of detailed annotation of the chip, no further analysis could be completed with this microarray.

A small number of COTS resequencing chips from TessArae (Potomoc Falls, VA) were procured in the TIF project for a comparative analysis study using a small number of pathogens. The TessArae RPM-TEI Array is intended as an *in vitro* diagnostic tool to provide comprehensive detection of CDC Category A, B, and C threat agents, including differentiating closely related organisms that can be confounders [9–11]. The chip is an Affymetrix design, but uses proprietary sample and chip processing methods. The chip design is intended to perform detailed DNA sequence analysis of specific target sites in the pathogens of interest, which could offer a significant advantage in terms of detail and level of strain discrimination. The technical protocol was found to be quite complex for inexperienced users, and the first few chips were of no analytical value. Once procedural details were resolved, chip raw data were obtained, which were then transmitted to the TessArae remote server system for detailed analysis. The results were in general conformity with expectations, in that the test species could be identified by the chip system. However, in many cases the control features of the chips were reported as "failed". In the case of a diagnostic tool, control failures usually lead to failure of the assay. For the TessArae product, each chip is subject to detailed analysis by an expert user at the supplier site, thus even "failed" chips returned some useable data. While this may be an advantage in terms of reduced waste and cost, the requirement for data processing by the vendor's own computer systems using proprietary algorithms, and the common intervention of the expert users did not lend confidence to the results of the overall product.

[1] (Dr. N. Silman, personal communication)

Design and testing of the microarray

Complete details for the design and procurement of the microbial fingerprinting DNA microarray have been previously published [3,12].

A notional design for a universal genotyping microarray (or indeed, any assay system) might contain all possible DNA sequence targets of a given length which could exist. For the 25 base pair (bp) size of the oligonucleotide probes on the Affymetrix microarray, an ideal array could sample any possible sequence (known or unknown) if all possible 25-base oligonucleotides were spotted on the array. It would require 1.126×10^{15} individual features (A, C, G or T at all 25 positions (4^{25}) or $\sim 1.126 \times 10^{15}$ individual sequences). This would in turn require 1×10^{9} microarrays (assuming a maximum of 1×10^{6} features per array) to cover most of the possible sequences. Thus, designing an array with all possible 25 base pair sequences was not a feasible approach. It was necessary therefore to develop a discrete set of species and strain-specific target sequences, which was accomplished using publicly available database sources, and the services of Affymetrix in the final design and fabrication of the arrays.

The first step in identifying regions of interest was to review the existing literature on bacterial microarray genotyping and strain differentiation. This provided a partial list of genes to include in our search. Next, various online databases were investigated for genes of interest. Initially, the NCBI Protein Clusters database [13] was used. Antibiotic resistance gene names and accessions were obtained from the Antibiotic Resistance Genes Database [14]. The majority of the resistance gene sequences used for probe selection was obtained from VFDB, the Virulence Factors of Pathogenic Bacteria database [15].

Organisms included in the search were derived from the National Institute of Allergy and Infectious Diseases Category A and B list of priority pathogens [16]. Also selected were *Haemophilus influenzae, Acinetobacter baumannii*, *Chaetomium* species, *Rickettsia* species, and plasmids pBC16 and pLS1 (containing tetracycline resistance genes). Sequences representing bacterial toxins and antimicrobial resistance sequences (*e.g.* other antibiotic markers) were also sampled. Targets for viral pathogens were not included in this chip. Redundancy is built into the Affymetrix microarray technology, wherein variants of specific probe sequences differing by one or a few bases from the specific probe are used to assess non-specific or variant binding to probe sites. The number of microbial genomic targets thus does not equal the number of individual probes on the array. In general, each specific target is represented by 3–20 individual probe sequences, varying by length, sequence, or single base pair differences. In typical applications, only one summary signal is reported from a probe set, the remaining features serving as quality assurance and quality control indicators. For genomic fingerprinting, however, variants related to (but not identical to) the primary probe may also contain useful signals and are also reported.

The design targeted approximately 16,000 individual sequence targets, which, allowing for sequence variants and internal controls, led to over 81,000 unique probes. The remaining capacity of the chip surface contained ~140,000 probes from the Affymetrix "antigenomic library" to serve as non-targeted probes, essentially a random target library.

The final microarray design was assembled using 81,678 probes from 11,516 unique microbial sequences, 24,660 probes from 264 SNP sequences, and approximately 140,000 non-specific probes along with controls to fill in the 220,678-probe chip. Annex A contains the listing of species- and strain-specific probes on the final microarray design. The Affymetrix proprietary technology for microarray preparation is also described in reference 12. One aspect of the Affymetrix approach is the control of the intellectual property around the chip design. Interested

users of this design must obtain the permission of the owner/designer to order the chip from Affymetrix, until the design is released into the public domain, or licensed for commercial use.

In silico verification of the design of the sequences on the chip versus available genome data was performed by iterative searches of the NCBI genome databases. Since the entire feature sequence set was designed using publicly available databases, of which NCBI comprises a large, if not exhaustive aggregation, it was anticipated, and found, that *in silico* testing would recapitulate the expected species and strain identifications.

Testing the design

Table 1 lists the microbial DNA samples used for preliminary testing of the custom microarray design. DNA samples from *E. coli* were prepared by Canada West Biosciences, while DNA from level 2 and level 3 microbes was prepared by DRDC Suffield in DRDC. These latter DNA extracts were tested for sterility using standard procedures within the respective containment facilities prior to release for microarray testing. Additional DNA samples were obtained form the Canadian Food Inspection Agency and commercial sources. Due to cost and poor availability of some of the DNA samples, an isothermal genomic amplification method was developed, exploited, and adopted for routine use to enhance the quantity of genetic material for testing purposes [17]. The amplification protocol is now part of the routine method for this microarray, enabling the future testing of other sparse or rare samples. Protocols for processing the DNA samples and the microarrays have been previously described [12].

Table 1: DNA extracts used in initial testing of the microarray.

Genus	Species	Strain / Isolate
Escherichia	coli	JM109
Bacillus	anthracis	94188c (RP42)
Bacillus	cereus	ATCC 11778
Yersinia	pestis	ATCC 19428
Yersinia	enterocolitica	YE-D3

After preliminary testing, an extensive (although not exhaustive) collection of DNA samples representing different species and strains were also tested on the microarray. Annex A lists all the possible species and strains which could have been tested. Annex B contains a listing of all species and strains which were tested more than once during this project.

Digitized chip scanner images were collated on Excel spreadsheets (Microsoft, Seattle, WA). Notably, Office 2007 is the first commercial spreadsheet program capable (without modification) of capturing the entirety of the digital data from a single microarray, let alone from the entire set in excess of 100 arrays.

As there were no pre-existing software tools for analysis of the microarrays in this project, various other software tools used for gene expression arrays were exploited to analyze the data.

Due to the large size of the data sets (244,000 rows, >100 columns), no extant microarray software was useful for comparisons between sample data sets. To circumvent this difficulty, manual data reduction procedures were performed. For pair wise comparison (i.e. species vs species), average intensity values in excess of a minimum cut-off value (0.5% of 65535, or 325) were used. To verify difference (or not) between samples, two-tailed Student's t-tests were used. For visualization of large scale data summaries, Chromablast [18] or MS Excel built-in conditional formatting tools were used.

Results and Discussion

Data were obtained during preliminary testing on the Affymetrix custom-designed microarray for DNA samples from *Escherichia coli*, *Yersinia pestis*, *Yersinia enterocolitica* and *Bacillus anthracis*. Few analytical tools are available for comparing and discriminating results in genomic fingerprint data, thus for this testing phase qualitative comparison was used. A summary of the microarray workflow is shown in Figure 4.

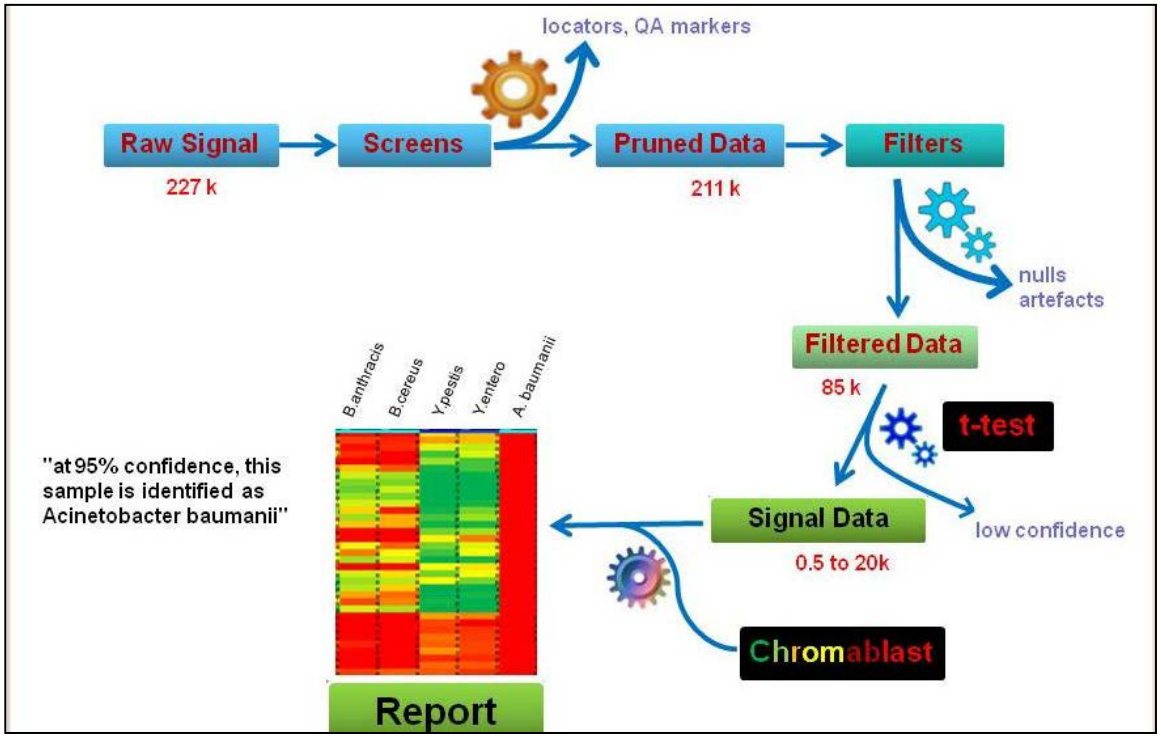

Figure 4: Genomic fingerprinting microarray workflow.

Using Chromablast [18], a heat map representing relative signal values was developed for a series of technical replicates of *E. coli* and used for the microarray testing. Figure 5 represents an interpretive key to understanding the data display generated by Chromablast. The full display for a complete data set covers multiple pages. An excerpt of a Chromablast analysis showing a region of the *E.coli* replicate data is shown in Figure 6. Uniform heat map colour across the replicates

would indicate perfect concordance on each target between replicates. The excerpt region shows examples of this, as well as some targets with varying colour, indicating some variation across the replicates. In Figure 6, green represents low value intensities (i.e. background to about 12% of maximum intensity, 0 to 6 in log base 2), and bright red indicates maximal intensity, as indicated in the scale below the heat map. This colour choice is a conventional display in the microarray literature; Chromablast offers a choice of 256 colours in any order. The absolute scale of variation between non-normalized array data sets is seen to be about 30% within individual probe sets. This is verified by numerical analysis of the raw intensity data. Most of this variation is concentrated within the lower intensity values, where the standard deviation as a fraction of the mean is maximal. Above the mean signal intensity (~7.0 in log 2), the maximum signal variation per probe set is about 5% (scatter plot in Figure 7). In practice this means that higher signal values have lower variation, and represents data in which the highest confidence may be placed.

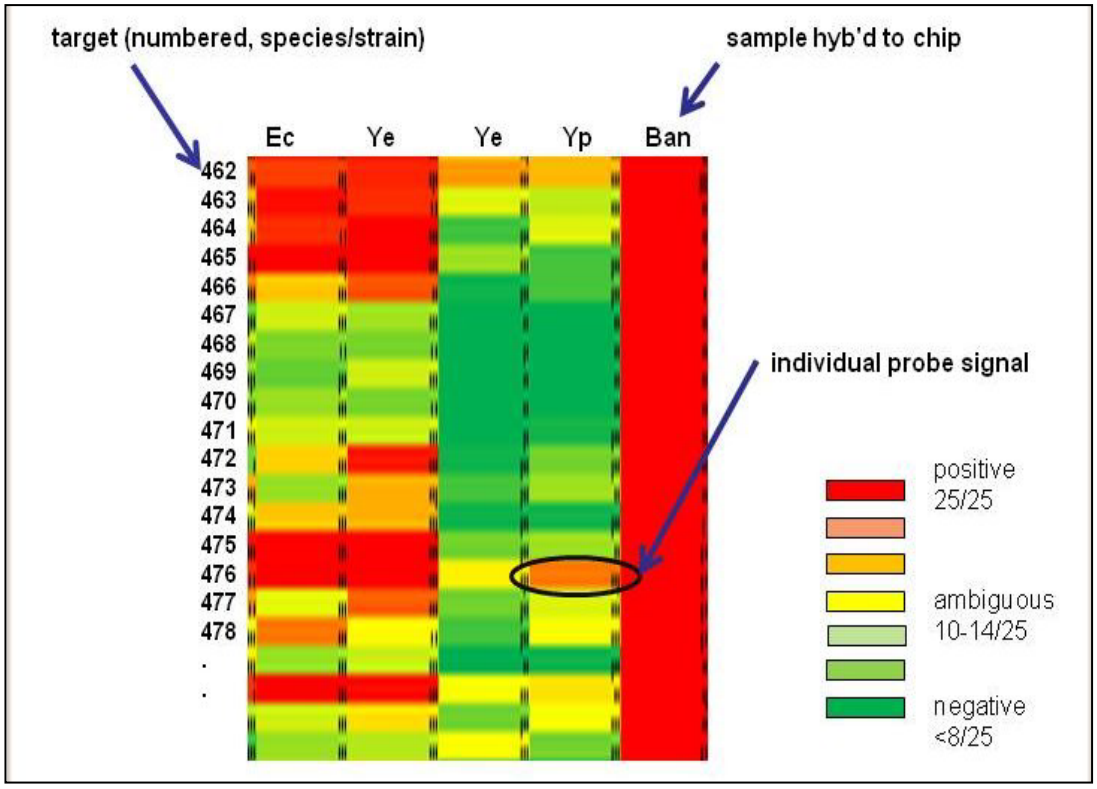

Figure 5: Interpretation of Chromablast display.

In practice, this suggests that a pruning of low-intensity signals may be useful to refine discrimination between samples versus knowns. One method to compensate for signal variation between replicate arrays is to use Student's t-test to compare knowns to unknowns. In the case of the *E. coli* replicated data set, for the complete data set, including the lowest value probe intensities (15,533 probe sets), less than 2% of all signals in a pair wise comparison have a t-test value of less than 0.05. If only the signals greater than the minimum cutoff value (0.5% of 65535, or 325) are considered, the number of t-test values less than 0.05 falls to ~1% (72 probe sets). "Significant" t-test results obtained for low-intensity signals (low confidence) are thus removed.

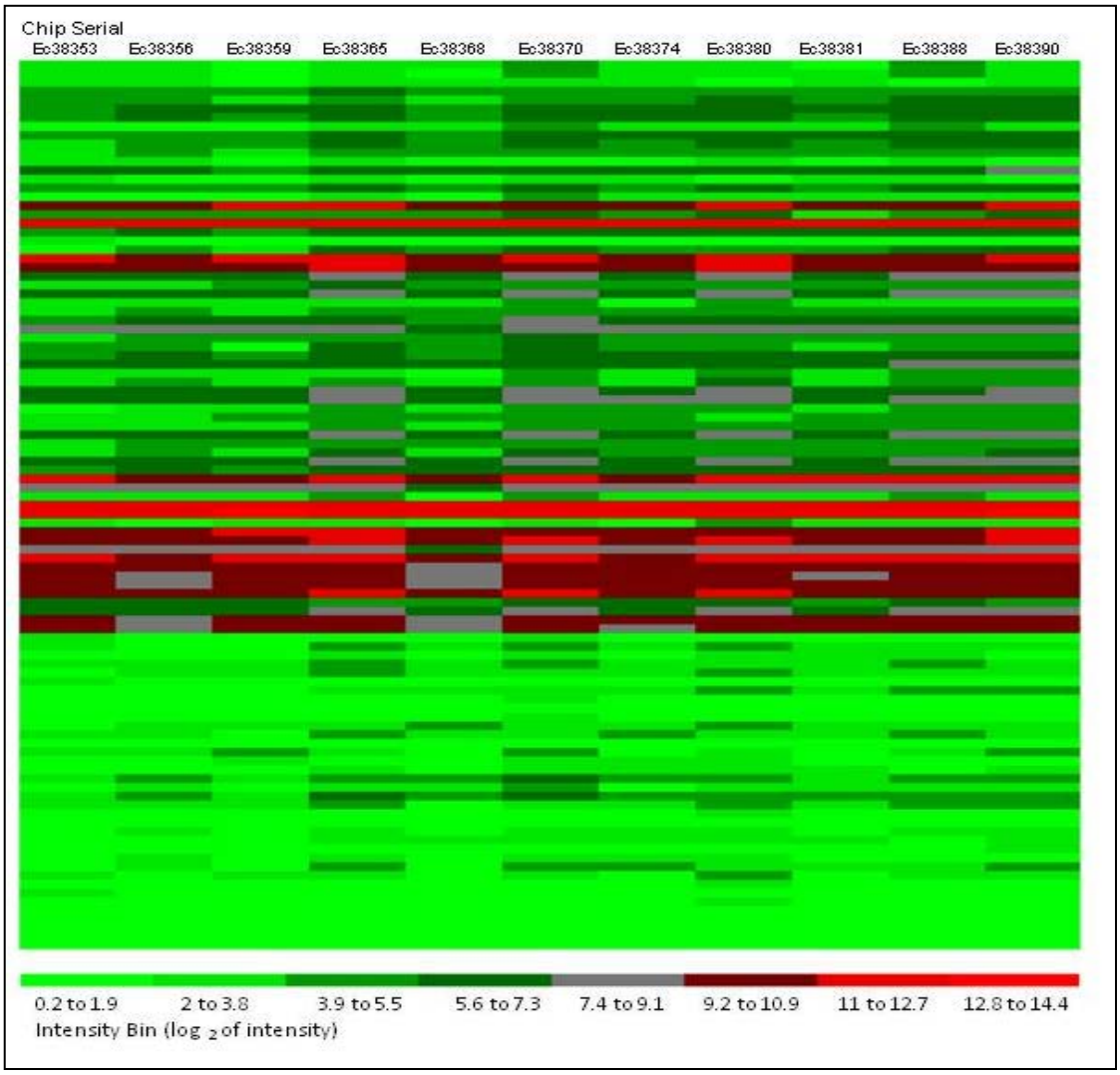

Figure 6: Comparison of signals from E. coli replicates.

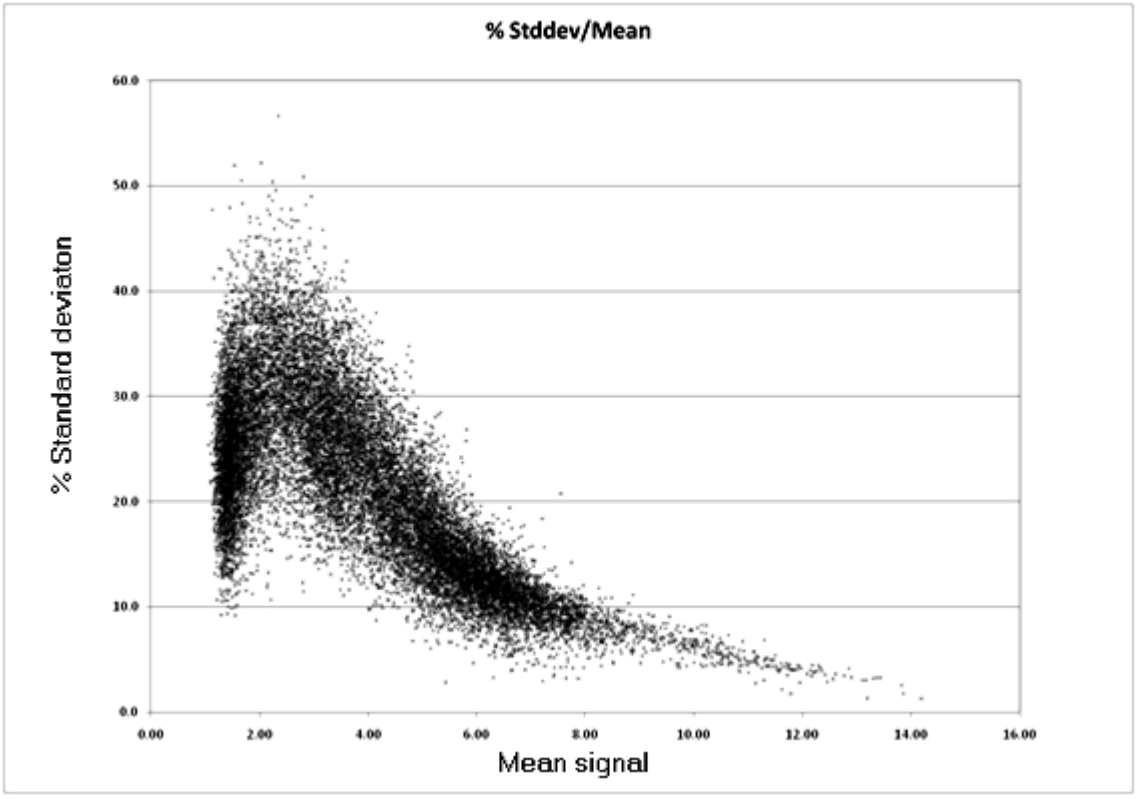

Figure 7: Standard deviation versus mean signal value.

If one selects as an upper limit a standard deviation of ± 10%, this corresponds to about 2^8 in this data set. In practice this limit falls in the range of 2^8 to 2^9 (intensity range of 256–512). Thus the value of intensity used in the preliminary analysis (325 or 0.5% of maximal intensity) has been shown to be quite reasonable. Such a discriminator still leaves several thousand features for comparison between samples, most of which have relatively small variation. Occurrences of outliers or systematically unreliable signal sets, as indicated by this analysis, are unlikely to interfere with discrimination between different genera or species, but may complicate detailed discrimination between closely related strains.

In Figure 8, a display of 210 signals from probes designed to detect *Bacillus anthracis* is shown. In this case, data from eight microarrays are shown. Each sample was independently hybridized to separate arrays in duplicate. The samples from left to right are *E.coli* JM109, *Yersinia enterocolitica* YE-D3, *Yersinia pestis* ATCC 19428, and *Bacillus anthracis* RP42. Notably the *B. anthracis* DNA sample yields a high frequency of high-intensity signals, shown as red bars. A few of the designed probes exhibit low intensity values (green bars), usually close to background values.

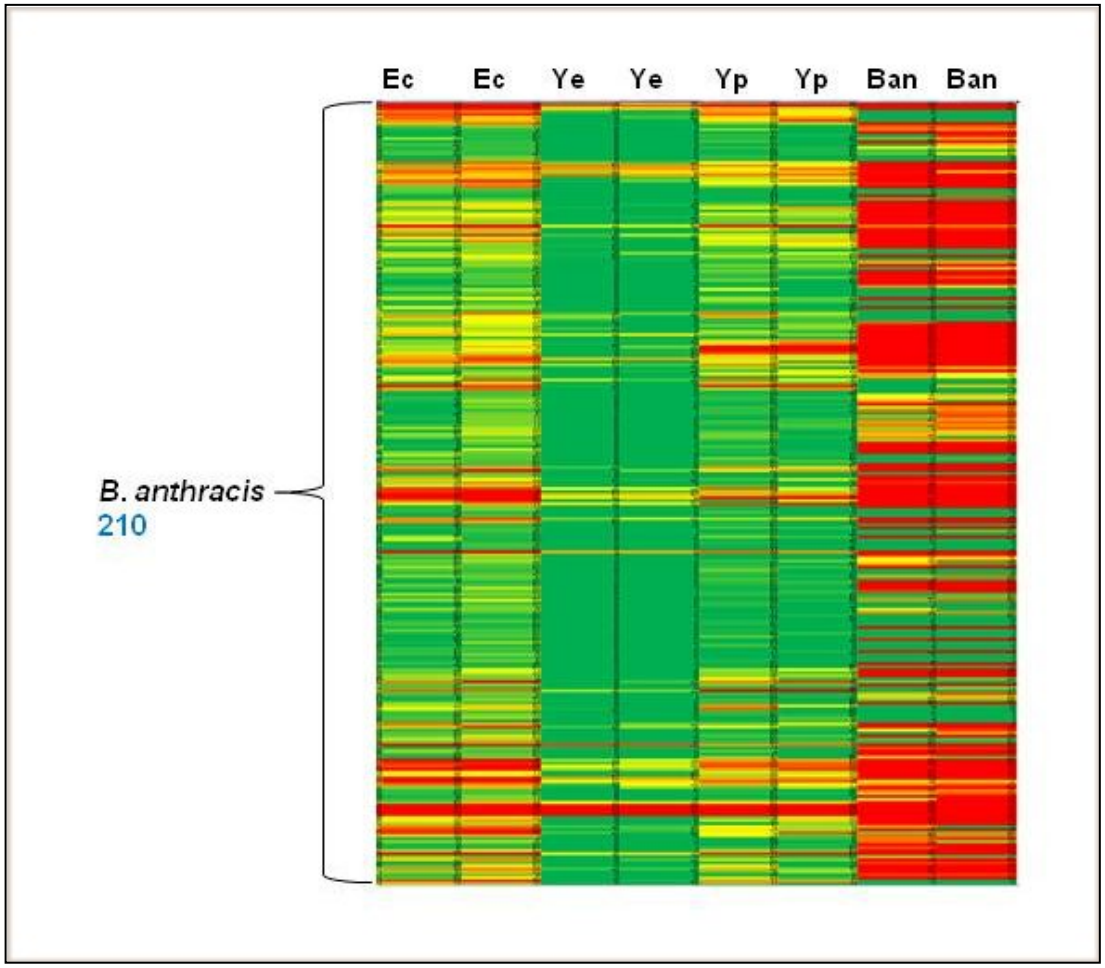

Figure 8: B. anthracis probe responses for various applied samples.

The observation that not all probes designed for a given species would yield positive signals was repeated for all of the DNA samples tested. Given that the probe design is derived from publicly available sequence resources, it is likely that some strain-specific (or sequence record specific) differences exist between the design and the test DNA. Note that some clear positive signals for *B. anthracis* probes occur in samples of *E.coli* DNA. This represents a strong case for multiple redundant probes for the microbes of interest, such that multiple points of identity and difference can be found for any species or strain tested.

Figure 9 is a display of data from the same set of microarrays as in Figure 8. In this case, the data have been filtered and excerpted to review the response of probes designed to detect *Y. enterocolitica* versus *Y. pestis.* These two species have very little overlap at the highest scoring microarray probes. One probe in the lower half of the figure, indicated by an arrow, shows just above background signal for *Y.pestis* (although it was designed to detect *Y.pestis*), and also similar signal intensity for *Y. enterocolitica*. This probe has a low signal in *E.coli* samples as well, but none in *B.anthracis.* This is an indication that the pattern of signals (as opposed to presence

versus absence) is also likely to be useful in species or strain discrimination, although patterns are much more difficult to analyze.

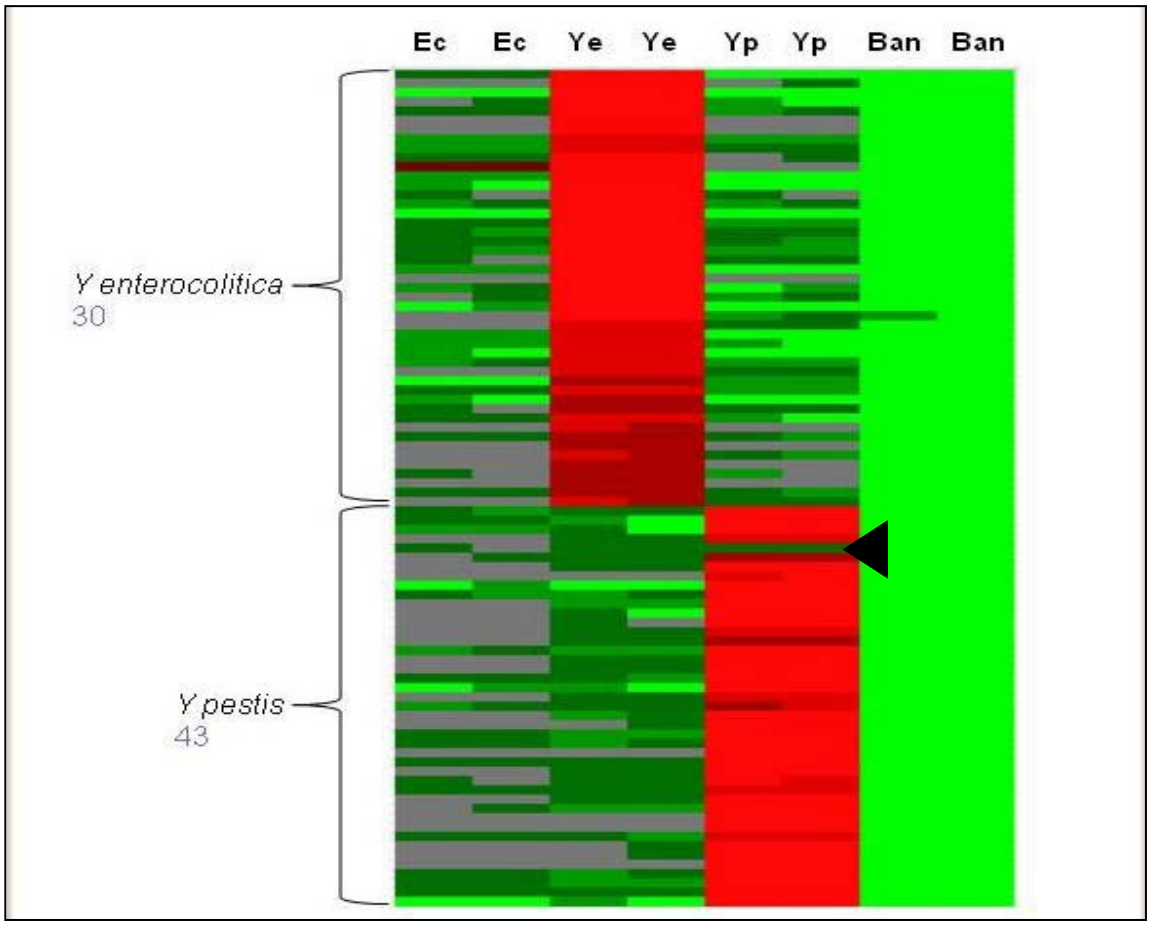

Figure 9 : Yersinia sp. probe responses for various samples.

The microarray design is capable of detecting by absence as well as presence of signals. For example, in Figure 10, the probes designed to detect *Clostridium botulinum*, for the most part do not yield requent signals from *E.coli*, *Yersinia* sp., or *Bacillis anthracis*. Unfortunately, DNA samples from *C. botulinum* were not available during this sutdy. A sample of *Clostridium perfringens* was used to test the response of the microarray to *Clostridia* sp. Of 2682 probes designed to detect Clostridia of various species, 297 should detect *C. perfringens* with some specificity. In panel A of Figure 11, the samples of *C. perfringens* (Cpe) are detected by 154 of the 297 probes, at signals above the cutoff limit of 325. For comparison, *C. jejunum* (Cje) DNA yields no positive signals on probes for *C. perfringens*.

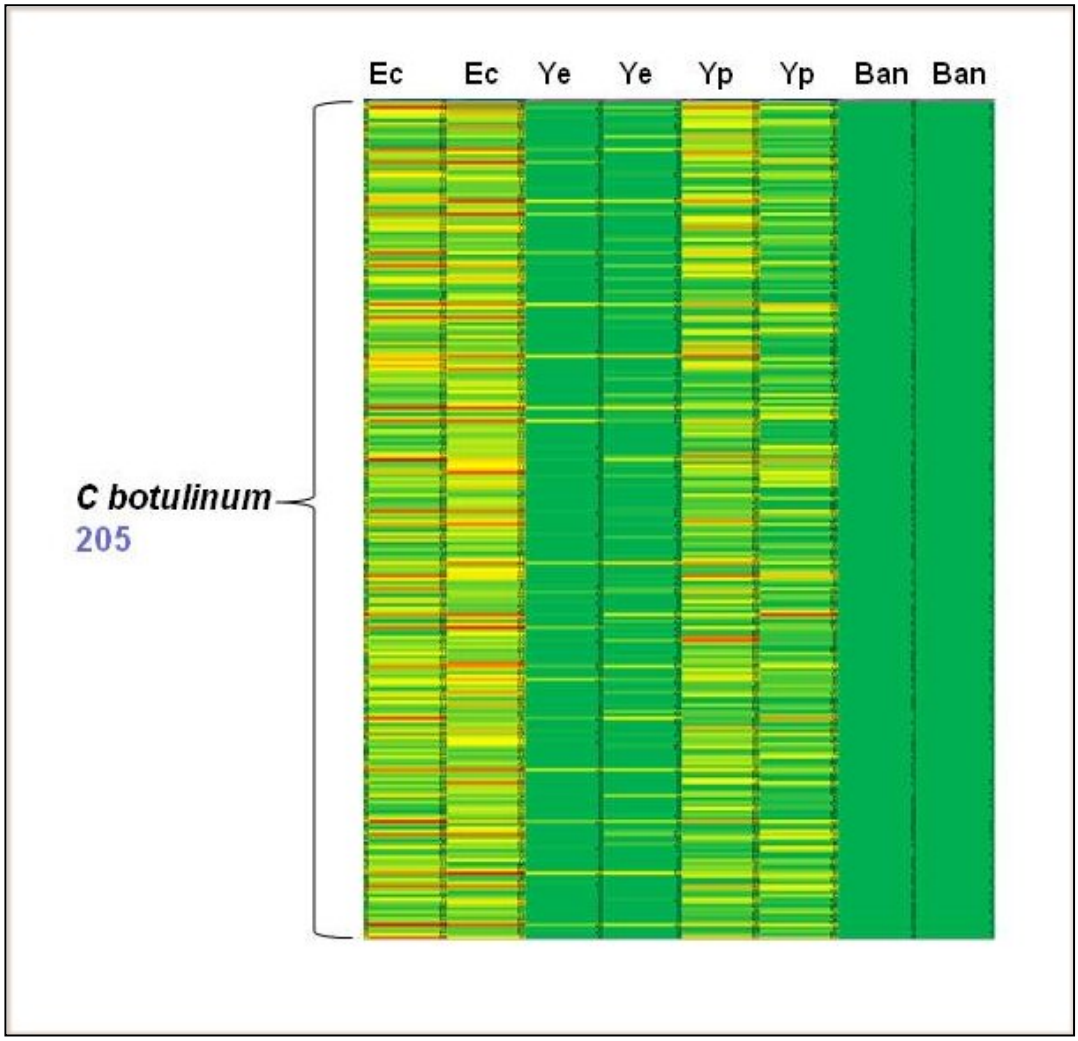

Figure 10: C. botulinum probe responses for various samples.

Conversely, in panel B of Figure 11, an excerpt of 297 probes designed to detect *Clostridium botulinum* are shown, with corresponding signals generated by DNA from *C. perfringens* and *C. jejuni*. An additional 300 *C. botulinum* probes not shown here have an equally uniform pattern of signals at or near background. There is essentially no crosstalk between the conspecific probes for these two species. This figure demonstrates visually the requirement to have multiple probes for any given species to be interrogated, and the usefulness of negative as well as positive signals. If some of the probes designed to yield positive signals fail, the multiplicity of targets still yields sufficient data for confident identification. Negative signals reinforce the interpretation that one is not simply observing a related strain.

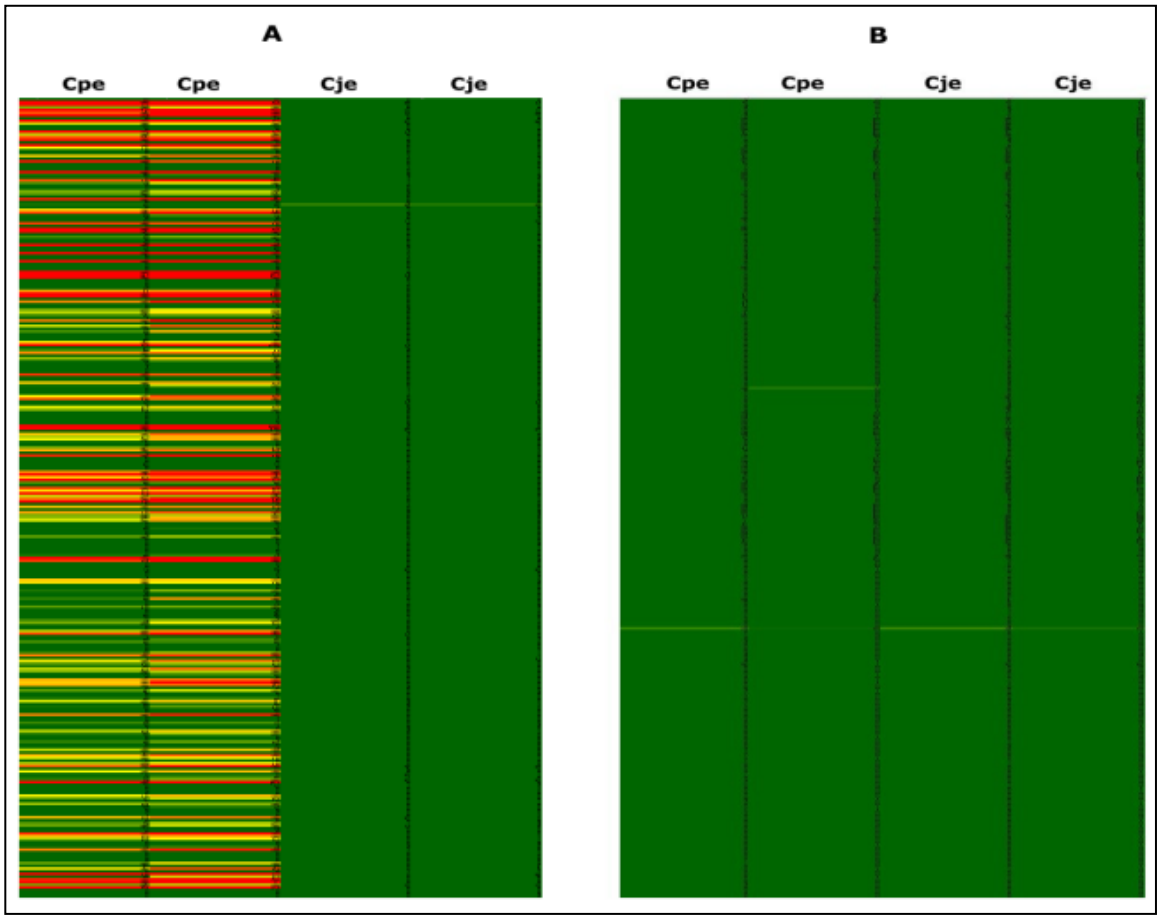

Figure 11 : Comparison of signals from probes for Clostridia sp.

Conclusions

Using a microarray platform, it is possible to assay thousands of targets at once, in every sample. A microarray in this application is a microscope slide or silicon chip onto which are spotted thousands of DNA probe sequences, each of which can detect one or more unique fragments of DNA. Using DNA probes which are designed to be specific to known microbial sequences, one can identify the species and strain of organism under examination.

The very specificity of the DNA hybridization phenomenon, however, would seem to preclude the ready detection of the presence of novel genetic sequences which might occur in unknown or recombinant organisms. In the absence of prior knowledge of the precise DNA sequence of the organism, it is not possible to build specific microarray targets to detect that DNA. In order to relieve the requirement for prior knowledge of all possible genetic content, the application of non-specific probes may be useful. For example, by selecting DNA sequences which do not explicitly correspond to known microbial sequences, one can detect the presence of novel or unexpected genetic sequences in the microbe. A sufficient panel of non-specific sequences in itself provides a unique "fingerprint" for any microbe, since each strain will exhibit a different "fingerprint" which can be measured. In addition, by using DNA probes specific for known "threat factors" (for example antibiotic resistance, virulence, etc.) that are not pathogen-specific, one can add the ability to identify organisms that have been genetically modified to include these traits. An organism carrying unusual or unexpected resistance or virulence genes represents an increased threat. Probes specific to known cloning vector fragments have the added virtue of detecting recombinants which are likely to be of laboratory origin rather than naturally occurring.

The combination of the "random" sequence fingerprint added to the panels of known sequences (microbial genes, antibiotic resistance, etc.) can offer a tool for identification of species or strain, and the presence of foreign genetic material. The reproducibility of such patterns on a given microarray design demonstrates that the hybridization of the genomic DNA to the features is sequence specific and not random. A further consideration is that the hybridizations are probably not context specific. That is, if a genome has been rearranged, it will in general contain the same sequence content, in altered context. Other genomic fingerprinting techniques such restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), or similar pulsed field techniques require positional stability to correctly indicate that the organism is genetically different from library samples, even if functionally (i.e. microbiologically) similar. Using hybridization techniques that do not require stable overall genome structures (e.g. PCR or microarrays) means that we should be able to identify and compare genomic content without much interference from contextual alterations. This hypothesis has not yet been tested, and it is not known at this point what the sampling density of the genomic DNA is or whether it is sensitive to positional changes within the genome.

The importance of the problem of recombinants has been demonstrated in a published account of a modified mouse pox virus (related to human smallpox) containing an extra gene encoding mouse interleukin-4 [19]. When exposed to the modified virus, mice vaccinated against mouse pox and mice that are genetically resistant to the normal virus exhibit mouse pox symptoms and high mortality. Methods by which such modifications to potential biowarfare agents can be made have also been published, emphasizing the need to understand the possibility of novel pathogens created by intention, and for tools to detect such recombinants [20,21].

The application of microarrays to microbial genotyping or fingerprinting is a technical compromise of time and difficulty versus data density. Single target or multiplex real-time PCR assays are faster and can be quantitative. Real-time PCR assays can in principle detect down to 1-10 targets per assay reaction, based on positive detection of specific sequences in known genomic targets. Routine PCR assays are not the best method of choice for detecting recombinants, variants, or the presence of non-target organisms however. If an assay system could run hundreds of PCR reactions for each test sample, the analytical density of the microarray could be equalled. Typical microarray open source platforms can detect 20–50,000 targets per array, using a single labelling or amplification reaction. Open source microarray systems typically take 18–26 hours for a single execution, but each run encompasses the equivalent of one to two thousand multiplex PCR reactions.

The number of assays executed per microarray has the drawback, that for some samples, DNA from multiple species is likely to be present and may contribute to the measured signals [22]. If the microarray contains sufficient numbers of features and has a high degree of automation, endemic species are always going to give a signal; thus, the mere presence of a signal of such a species in a given environmental or clinical context is not in itself meaningful. Assays must be combined with other indices of suspicion (clinical signs, known exposure, suspect samples) in order to determine whether a given positive represents a real diagnosis or threat [23]. This is also true for most other molecular or microbiological assays currently in use. Simple detection of DNA from pathogens is insufficient to establish a diagnosis in a clinical setting. For example, healthy individuals may carry pathogenic bacteria as part of their flora, but not suffer concurrent disease symptoms from that microbe. Clinical detection of a pathogen will be combined with other signs and symptoms to confirm the diagnosis.

In addition, as the sensitivity of assay systems improves (due to non-specific genomic amplification for example), out-of-context true positive signals (not within the normal range of endemics) may be detected. Such signals may be due to sample contamination by workers, gratuitous sampling of infrequent (but locally intense) organism populations, or previously undetected genetic similarity between lab strains and endemic strains. Use of confirmatory assays of high-specificity (e.g. real-time PCR) will complement such data.

Given the requirement for technical expertise in operating a microarray system, and given the sensitivity to multiple targets in some samples, microarray systems will continue to require sophisticated laboratory support. Microarray systems are in use in clinical centers, but point-of-care microarray systems are not imminent. On the other hand, time-to-result times are comparable to or better than conventional microbiology. Detailed testing of the current microarray and comparison to other microarray systems is underway. Additional testing with an expanded library of DNA samples and a wider sampling of species is required to fully assess the value of the microarray as a tool for diagnosis, detection, and identification of microbial samples. Because of the capacity and specificity of microarrays, coupled with the technical challenges and level of expertise required for use, microarrays should be considered as a useful complement to existing PCR, RFLP, or AFLP technologies.

References

Miller, M. and Tang, T. (2009), Basic concepts of microarrays and potential applications in clinical microbiology. Clin Microbiol Rev. 22:611-633.

Lam,V., Crichton, M., Laing, T. and Mah,D. (2008), Microarray genomic systems development (DRDC Suffield CR2008-154), Canada West Biosciences, Camrose, Alberta.

Kostrzynska, M. and Bachand, A. (2006), Application of DNA microarray technology for detection, identification, and characterization of food-borne pathogens. Can. J. Microbiol. 52:1-8.

Willse, A., Straub, T., Wunschel, S., Small, J., Call, D., Daly, D. and Chandler DP. (2004), Quantitative oligonucleotide microarray fingerprinting of Salmonella enterica isolates. Nucleic Acids Res. 32:1848-1856.

Kingsley,M., Straub,T., Call,D., Daly,D., Wunschel,S. and Chandler, D. (2002), Fingerprinting closely related Xanthomonas pathovars with random nonamer oligonucleotide microarrays. Appl. Environ. Microbiol. 68: 6361–6370.

Liu, Z., Venkatesh, S. and Maley, C. (2008), BMC Sequence space coverage, entropy of genomes and the potential to detect non-human DNA in human samples. BMC Genomics 9:509.

Ayodeji, M., Kulka, M., Jackson, S., Patel, I., Mammel, M.,Cebula, T. and Goswami, B. (2009), A microarray based approach for the identification of common foodborne viruses. The Open Virology Journal 3:7-20.

Wang, D., Coscoy, L., Zylberberg, M., Avila, P., Boushey, H., Ganem, D. and DeRisi, J. (2002), Microarray-based detection and genotyping of viral pathogens. P.N.A.S. 99:15687-15692.

Malanoski, A., Lin, B., Wang, Z., Schnur, J. and Stenger, D. (2006), Automated identification of multiple micro-organisms from resequencing DNA microarrays. Nucleic Acids Res. 34:5300-5311.

Leski, T., Malanoski, A., Stenger, D. and Lin, B. (2010), Target amplification for broad spectrum microbial diagnostics and detection. Future Microbiol. 5:191-203.

Leski, T., Baochuan, L., Malanoski, A., Wang, Z., Long, N., Meador, C., Barrows, B., Ibrahim, S., Hardick, J., Atichou, M., Schnur, J., Tibbets, C. and Stenger, D. (2009), Testing and validation of high density resequencing microarray for broad range biothreat agents detection. PLoS ONE 4: e6569.

Ford, B., Bader, D., Shei, Y., Ruttan, C. and Mah, D. (2009), An Affymetrix microarray design for microbial genotyping (U), (DRDC Suffield TM 2009-183) Defence R&D Canada – Suffield.

Protein Clusters (ProtClustDB). (2010), National Center for Biotechnolcgy Information. http://www.ncbi.nlm.nih.gov /proteinclusters (Access date 25 Jan. 2008).

Liu, B. and Pop, M. (2008), ARDB_Antibiotic resistance genes database. Nucleic Acids Res. 37:D443-D447. http://ardb.cbcb.umd.edu (Access date 24 Jan. 2008).

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y. and Jin, Q. (2005), VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 33:D325-D328. http://www.mgc.ac.cn/VFs/main.htm (Access date 2 Feb. 2008).

NIAID Category A, B, and C Priority Pathogens. National Institute of Allergy and Infectious Diseases. (2008), http://www.niaid.nih.gov/topics/biodefenserelated biodefense/ research/pages/cata.aspx (Access date: 2 Feb. 2008).

Ford, B., Bader, D., Ruttan, C., and Mah, D. (2010), Isothermal amplification of microbial genomic samples (DRDC Suffield TM2010-143 (in press)) Defence R&D Canada–Suffield.

Ford, B., Shei, Bjarnason, S. and Richardson,C. (2006), ChromaBlast- a data visualtization tool (DRDC Suffield TM2006-049), Defence R&D Canada – Suffield.

Jackson, R., Ramsay, A., Christensen, C., Beaton, S., Hall, D. and Ramshaw, I. (2001), Expression of mouse interleukin-4 by a recombinant ectromelia virus suppresses cytolytic lymphocyte responses and overcomes genetic resistance to mousepox. J. Virol. 75:1205-1210.

Galams, F. (2008), Biological weapons, nuclear weapons, and deterrence; the biotechnology revolution. Comparative Strategy 27:315-323.

Buller, M. (2004), Mousepox; a small animal model for biodefense research. Applied Biosafety 9:10-19.

Venter, J., Remington, K., Heidelberg, J., Halpern, A., Rushc, D., Eisen, J., Wu, D., Paulsen, I., Nelson, K., Nelson, W., Fouts, D., Levy, S., Knap, A., Lomas, M., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannoch, C., Roger, Y., Smith, H. (2004), Environmental genome shotgun sequencing of the Sargasso sea. Science 304:66-74.

Tomioka, K., Peredelchuk, M., Zhu, X., Arena, R., Volokhov, D., Selvapandiyan, A., Stabler, K., Mellquist-Riemenschneider, J., Chizhikov, V., Kaplan, G., Nakhasi, H. and Duncan, R. (2005), A multiplex polymerase chain reaction microarray assay to detcet bioterror pathogens in blood. J. Mol. Diag. 7:486-494.

List of symbols/abbreviations/acronyms/initialisms

AFLP	amplified fragment length polymorphism
APRT	adenine phosphoribosyltransferase
ATCC	American Type Culture Collection; an organization supplying standard microbial strains and samples
BLAST	Basic Local Alignment Search Tool
bp	base pair
BSL2	Biosafety Containment Level 2
BSL3	Biosafety Containment Level 3
dNTP	dATP, dCTP, dGTP, dTTP : deoxynucleotide triphosphates of DNA bases
DNA	Deoxyribonucleic acid
DRDC	Defence Research & Development Canada
HPT	hypoxanthine guanine phosphoribosyltransferase (also HPRT)
mM	millimolar
NCBI	National Center for Biotechnology Information (also referred to as Genbank)
PCR	Polymerase chain reaction
R&D	Research & Development
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
SNP	single nucleotide polymorphism; a sequence variant at one base position which may be different between populations or individuals
TDT	terminal deoxynucleotidyl transferase
TIFF	Tagged Image File Format
TRIS	2-amino-2-hydroxymethyl-1,3-propanediol

Glossary

feature	a location on a microarray surface which has a known (or mappable) two dimenisonal location (X and Y coordinates), and contains DNA fragments or oiligoncleotides which may serve as a capture porbe for a complementray fragment in the sample mixture.
fingerprint	a collection of signal intensity scores, digitized from an image of a hybridization of genomic DNA to a microarray spotted with DNA fragments. The fingerprint of a given species and strain is unique from that of other species or strains.
gene	a DNA sequence which encodes a single inheritable genetic trait
genus	a taxonomic grouping of species by (among others) morphology, ecology, or origin
genomic DNA	the DNA which comprises the genetic material of an cell, and is inherited by the progeny of the cell. The sequence of nucleotides in the genomic DNA comprises the genes, and determines the properties of the microbe. For many microbes, the genomic DNA sequence is in the public domain.
hybridization	sample DNA (or RNA) is tagged with a fluorescent dye, then applied to the surface of the microarray. Under controlled conditions, sequences in the sample DNA which correspond to sequences in the microarray features, will bind to the features (hybridize). Hybridization often refers to the entire process from labeling to binding, to post incubation washing.
microarray	a microscope slide, filter membrane or other solid surface, onto which DNA fragments have been spotted (features) in an organized grid.
nucleotide	the components of DNA are the nucleotides deoxyadenosine monophosphate, deoxycytidine monophosphate, deoxyguanosine monophosphate, deoxythymidine monophosphate, and the chemical bonds which join them into long chains. Genetic information is encoded in the order in which the nucleotides occur in the DNA chain.
oligonucleotide (oligo)	a fragment of DNA (or RNA) representing a section of genetic material from which the sequence may be known. Oligos may also be “random” in sequence, such that the sequence is not derived from knownsequences
species	the grouping of microbes according to significant genetic differences (e.g. the ability to grow (or not) in an oxygen-free environment)
strain	a microbe which differs from other members of the same species by minor or additional genetic characters (e.g. resistance or sensitivity to penicillin).

This page intentionally left blank.

Annex A Species and strain-specific probes in the final array design.

Organism	Strain / details	Probes on Array
Acinetobacter baumannii	ACICU	38
Acinetobacter baumannii	ATCC 17978	54
Acinetobacter baumannii	AYE	143
Acinetobacter baumannii SNP	baumannii	20
Acinetobacter baumannii	baumannii	15
Acinetobacter baumannii	plasmid pSUN-5	5
Acinetobacter baumannii	SDF	44
Acinetobacter baumannii HPT	ATCC 17978	5
Acinetobacter baumannii HPT	AYE	5
Acinetobacter baumannii HPT	baumannii	5
Bacillus anthracis	Ames ancestor	140
Bacillus anthracis	Ames ancestor plasmid pX01	5
Bacillus anthracis	Ames ancestor plasmid pX02	25
Bacillus anthracis	anthracis	45
Bacillus anthracis	Australia 94	6
Bacillus anthracis	Kruger	5
Bacillus anthracis	Sterne	55
Bacillus anthracis APRT	A2012 plasmid pXO1	5
Bacillus anthracis APRT	Ames	5
Bacillus anthracis HPT	A0442	5
Bacillus anthracis HPT	anthracis	5
Bacillus anthracis HPT	anthracis	10
Bacillus anthracis plasmid	Sterne plasmid pX01+pX02-	10
Bacillus anthracis SNP	A2012	20
Bacillus anthracis SNP	anthracis	780
Bacillus anthracis SNP	other anthracis	200
Bacillus anthracis SNP	W. North America	20
Bacillus cereus	ATCC 10987	179
Bacillus cereus	ATCC 14579	200
Bacillus cereus	B. cereus plasmid pBCX01	5
Bacillus cereus	E33L	45
Bacillus cereus	G9241	5
Bacillus cereus group SNP	Bacillus	1800
Bacillus cereus HPT	E33L	5
Bacillus cereus SNP	ATCC 10987	80
Bacillus cereus SNP	ATCC 14579	320
Bacillus cereus SNP	B. cereus plasmid pBCXO1	320
Bacillus cereus/anthracis SNP	B. cereus plasmid pBCXO1	20
Bacillus amyloliquefaciens APRT	FZB42	5
Bacillus clausii	KSM-K16	5
Bacillus clausii APRT	KSM-K16	5
Bacillus halodurans	C-125	5

Bacillus halodurans APRT	C-125	5
Bacillus licheniformis	ATCC 14580	35
Bacillus licheniformis APRT	ATCC 14580	5
Bacillus pumilus APRT	SAFR-032	5
Bacillus subtilis	168	25
Bacillus subtilis APRT	168	5
Bacillus thuringiensis	97-27	115
Bacillus thuringiensis	Al Hakam	130
Bartonella bacilliformis	ATCC 35685	175
Bartonella henselae	Houston-1	270
Bartonella quintana	Toulouse	245
Bartonella tribocorum	CIP 105476	330
Bordetella SNP	Bordetella	20
Bordetella avium	197N	440
Bordetella avium APRT	197N	5
Bordetella bronchiseptica APRT	RB50	5
Bordetella bronchiseticas	RB50	75
Bordetella parapertussis	12822	268
Bordetella pertussis APRT	Tohama I	5
Bordetella pertussis	Bordetella	5
Bordetella pertussis	Tohama I	615
Bordetella petrii APRT	DSM 12804	5
Borrelia afzelii APRT	PKo	5
Brucella SNP	9-941	20
Brucella	all brucella	250
Brucella HPT	all brucella	5
Brucella abortus	9-941	30
Brucella abortus	S19	45
Brucella abortus SNP	melitensis/abortus	40
Brucella abortus APRT	9-941	5
Brucella abortus/melitensis SNP	abortus/melitensis	20
Brucella abortus/suis SNP	abortus/suis	20
Brucella canis	ATCC 23365	15
Brucella canis	S19	5
Brucella canis HPT	ATCC 23365	10
Brucella melitensis	16M	427
Brucella melitensis	2308 bv Abortus	210
Brucella melitensis	bv Melitensis	10
Brucella melitensis	bv Suis 686	5
Brucella ovis	ATCC 25840	82
Brucella ovis	bv Abortus 2308	35
Brucella suis	1330	25
Brucella suis	ATCC 23445	5
Brucella suis	ATCC 23447	5
Brucella suis	ATCC 25840	10
Brucella suis	bv. 4 str. 40	15
Brucella suis/abortus SNP	suis/abortus	80
Burkholderia SNP	Burkholderia	1160
Burkholderia HPT	all burkholderia	5
Burkholderia mallei APRT	ATCC 23344	5

Burkholderia mallei	ATCC 23344	40
Burkholderia mallei	PRL-20	5
Burkholderia multivorans APRT	ATCC 17616	5
Burkholderia pseudo/mallei SNP	pseudomallei/mallei	20
Burkholderia pseudo/mallei	Burkholderia	15
Burkholderia pseudomallei	668	10
Burkholderia pseudomallei	1710b	5
Burkholderia pseudomallei	392f	5
Burkholderia pseudomallei SNP	B7210	40
Burkholderia pseudomallei	K96243	75
Burkholderia pseudomallei SNP	pseudomallei	340
Burkholderia pseudomallei	pseudomallei	5
Burkholderia pseudomallei	T18-1984	5
Burkholderia pseudomallei HPT	91	5
Burkholderia pseudomallei HPT	668	5
Burkholderia pseudomallei HPT	NCTC 13177	5
Burkholderia pseudomallei APRT	668	5
Burkholderia thailandensis APRT	E264	5
Campylobacter concisus APRT	13826	5
Campylobacter fetus	82-40	440
Campylobacter hominis APRT	ATCC BAA-381	5
Campylobacter jejuni APRT	doylei 269.97	5
Campylobacter jejuni	269.97 ss doylei	476
Campylobacter jejuni	81116 (NCTC 11828)	351
Campylobacter jejuni	81-176	349
Campylobacter jejuni	jejuni	60
Campylobacter jejuni	NCTC 11168	560
Campylobacter jejuni	plasmid pCjA13 t	5
Campylobacter jejuni	RM 1221	304
Campylobacter jejuni APRT	81-176	5
Campylobacter jejuni plasmid	81-176 plasmid pVir	5
Chaetomium atrobrunneum	atrobrunneum	5
Chaetomium funicola	funicola	29
Chaetomium funicola	OC13	5
Chaetomium funicola	olrim130	5
Chaetomium thermophilum	CT2	20
Chaetomium thermophilum	MTCC 6350	5
Chaetomium thermophilum	thermophilum	85
Chlamydia abortus	S26/3	115
Chlamydia caviae	GPIC	115
Chlamydia felis	Fe/C-56	120
Chlamydia muridarum	Nigg (MoPn)	115
Chlamydia pneumoniae	AR39	115
Chlamydia pneumoniae	CWL 029	5
Chlamydia trachomatis	D/UW-3/CX	175
Chlamydia trachomatis	HAR-13	15
Chlamydia trachomatis	trachomatis	5
Clostridium botulinum APRT	Alaska E43	10
Clostridium botulinum APRT	ATCC 3502	5
Clostridium botulinum APRT	Eklund 17B	5

Clostridium botulinum APRT	Okra	5
Clostridium botulinum	A str. ATCC 19397	5
Clostridium botulinum	ATCC 3502	40
Clostridium botulinum	B str. Eklund 17B	5
Clostridium botulinum SNP	B1 str. Okra plasmid pCLD	20
Clostridium botulinum	B1 str. Okra plasmid pCLD	5
Clostridium botulinum	Bf	5
Clostridium botulinum SNP	botulinum	1860
Clostridium botulinum	C str. Eklund	5
Clostridium botulinum SNP	C. botulinum A strains	100
Clostridium botulinum	C. botulinum A strains	5
Clostridium botulinum	Clostridium botulinum	15
Clostridium botulinum	Hall 183	5
Clostridium botulinum HPT	Alaska E43	15
Clostridium botulinum HPT	Eklund 17B	10
Clostridium botulinum HPT	Loch Maree	20
Clostridium botulinum HPT	Okra	5
Clostridium botulinum	A3 str. Loch Maree	5
Clostridium acetobutylicum	ATCC 824	25
Clostridium beijerinckii	NCIMB 8052	20
Clostridium difficile	630	45
Clostridium difficile		15
Clostridium difficile HPT	difficile	5
Clostridium kluyveri APRT	DSM 555	5
Clostridium novyi	ATCC19402	45
Clostridium novyi	NT	40
Clostridium perfringens APRT	SM101	5
Clostridium perfringens	13	111
Clostridium perfringens	ATCC 13124	66
Clostridium perfringensS		20
Clostridium perfringens	SM101	65
Clostridium perfringens HPT	13	5
Clostridium perfringens HPT	ATCC 13124	10
Clostridium perfringens HPT	SM101	10
Clostridium perfringens plasmid	plasmid pCP13	5
Clostridium tetani	E88	55
Clostridium tetani HPT	tetani	5
Clostridium thermocellum	ATCC 27405	15
Corynebacterium diphtheriae	diptheriae	5
Corynebacterium diphtheriae	NCTC 13129	165
Corynebacterium efficiens	YS-314	65
Corynebacterium glutamicum	ATCC 13032	20
Corynebacterium glutamicum	R	69
Corynebacterium glutamicum APRT	ATCC 13032	5
Corynebacterium jeikeium	K411	110
Coxiella burnetii	CbuG Q212	15
Coxiella burnetii	Dugway 5J108-111	25
Coxiella burnetii	MSU Goat Q117	29
Coxiella burnetii	RSA 331	15
Coxiella burnetii	RSA 334	5

Coxiella burnetii	RSA 493	178
Coxiella burnetii HPT	Dugway	5
Coxiella burnetii HPT	burnetti	10
Enterococcus faecalis	faecalis	5
Enterococcus faecalis	MMH594	5
Enterococcus faecalis	V583	145
Enterococcus faecalis APRT	V583	5
Enterococcus faecalis HPT	faecalis	5
Escherichia coli	536	1035
Escherichia coli	1226	5
Escherichia coli	1334	5
Escherichia coli	55989	20
Escherichia coli	042	70
Escherichia coli	17-2	25
Escherichia coli	536 (UPEC)	30
Escherichia coli	B171	85
Escherichia coli	C1845	5
Escherichia coli	CFT 073 (UPEC)	516
Escherichia coli	coli	182
Escherichia coli	coli/shigella	5
Escherichia coli	E. coli plasmid pC15-1a_016	5
Escherichia coli	E/99 3-2 SHV	10
Escherichia coli	E2348/69	285
Escherichia coli	E45035	5
Escherichia coli	EC7372	5
Escherichia coli	EU2657	5
Escherichia coli	EU4855 plasmid	5
Escherichia coli	H11128	25
Escherichia coli	H11129	5
Escherichia coli	K12	38
Escherichia coli	K12 substr. MG1655	25
Escherichia coli	K983802	5
Escherichia coli	KS52	5
Escherichia coli	O157:H7 EDL933	345
Escherichia coli	plasmid	15
Escherichia coli	plasmid p541	5
Escherichia coli	plasmid pEC365	5
Escherichia coli	plasmid pGR2439	5
Escherichia coli	plasmid pMEL2	3
Escherichia coli	plasmid RZA92	5
Escherichia coli	Sakai(EHEC O157:H7)	11
Escherichia coli	SMS-3-5	20
Escherichia coli	Str. 01 (APEC)	50
Escherichia coli	Toho-1	5
Escherichia coli	UTI89 (UPEC)	65
Escherichia coli	YMC02/08/U310	5
Escherichia coli	SMS-3-5	5
Escherichia coli APRT	O157:H7 EDL933	5
Escherichia coli HPT	ATCC 8739	5
Escherichia coli HPT	E24377A	5

Escherichia coli HPT	F11	4
Escherichia coli HPT	HS	5
Escherichia coli plasmid	plasmid pAPEC-O1-ColBM	40
Escherichia coli strain EO 516	EO 516	5
Francisella holarctica APRT	OSU18	5
Francisella holartica	FTNF002-00	15
Francisella holartica	holartica	31
Francisella holartica	LVS	35
Francisella holartica	OSU18	25
Francisella holartica HPT	holartica	10
Francisella holartica SNP	FSC022	40
Francisella holartica SNP	FTNF002-00	80
Francisella holartica SNP	HOL 257	20
Francisella holartica SNP	holartica	240
Francisella holartica SNP	LVS	20
Francisella holartica SNP	OSU18	100
Francisella novicida	U112	105
Francisella novicida HPT	U112	5
Francisella novicida SNP	GA99-3548	700
Francisella novicida SNP	novicida	7480
Francisella novicida SNP	U112	620
Francisella tularensis	ATCC 6223	46
Francisella tularensis	francisella	5
Francisella tularensis	fsc033	5
Francisella tularensis	FSC198	15
Francisella tularensis	plasmid pOM1	5
Francisella tularensis	SCHU S4	411
Francisella tularensis	tularensis	52
Francisella tularensis	WY96-3418	55
Francisella tularensis SNP	SCHU	180
Francisella tularensis SNP	tularensis	580
Francisella tularensis SNP	WY96	100
Francisella tularensis SNP	WY96-3418	20
Francisella	Francisella	15
Francisella holartica/novicida	holartica/novicida	5
Francisella holartica/tularensis	holartica/tularensis	25
Francisella novicida/tularensis	novicida/tularensis	30
Francisella tularensis/holartica SNP	tularensis/holartica	20
Haemophilus ducreyi	35000 HP	405
Haemophilus influenzae APRT	86-028NP	5
Haemophilus influenzae APRT	Rd KW20	5
Haemophilus influenzae	12	30
Haemophilus influenzae	1007	89
Haemophilus influenzae	3179B	5
Haemophilus influenzae	86-028NP	336
Haemophilus influenzae	AM30	25
Haemophilus influenzae	C54	5
Haemophilus influenzae	influenzae	5
Haemophilus influenzae	N187	5
Haemophilus influenzae	Pitt EE	275

Haemophilus influenzae	Pitt GG	299
Haemophilus influenzae	Rd	95
Haemophilus influenzae	Rd KW20	375
Haemophilus somnus	2336	205
Haemophilus somnus	129 PT	380
Helicobacter acinonychis	Sheeba	279
Helicobacter hepaticus	ATCC 51449	250
Helicobacter pylori APRT	J99	5
Helicobacter pylori	26695	438
Helicobacter pylori	HPAG1	377
Helicobacter pylori	J99	484
Human	Human	100
Klebsiella pneumonia APRT	MGH 78578	5
Lactobacillus delbrueckii APRT	subsp. bulgaricus ATCC 11842	5
Legionella pneumonphila	Philadelphia 1	793
Legionella pneumophila HPT	Corby	3
Legionella pneumophila HPT	Lens	5
Legionella pneumophila HPT	Paris	10
Legionella pneumophila HPT	Philadelphia 1	5
Legionella pneumophila	Corby	296
Legionella pneumophila	Lens	378
Legionella pneumophila	Paris	399
Legionella pneumophila	pneumophila	5
Listeria innocua	Clip 11262	105
Listeria ivanoviil	ATCC 19119	5
Listeria monocytogenes	monocytogenes	10
Listeria monocytogenes APRT	EGD-e	5
Listeria monocytogenes HPT	4b 2365	10
Listeria monocytogenes HPT	EGD-e	5
Listeria monocytogenes	4b 2365	260
Listeria monocytogenes	EGD-e sv 1/2A	453
Listeria monocytogenes	F2365	95
Listeria monocytogenes APRT	F2365	5
Listeria monocytogenes SNP	J1-194	1280
Listeria monocytogenes SNP	J2-064	80
Listeria monocytogenes	J2-064	5
Listeria monocytogenes SNP	monocytogenes	5180
Listeria welshimeri APRT	SLCC 5334	5
Listeria welshimeri	SLCC 5334	100
Mycobacterium avium APRT	K-10 ss paratuberculosis	5
Mycobacterium avium	104	263
Mycobacterium avium	K-10 ss paratuberculosis	743
Mycobacterium bovis APRT	BCG str. Pasteur 1173P2	5
Mycobacterium bovis	AF2122/97	15
Mycobacterium bovis	BCG Pasteur 1173P2	15
Mycobacterium gilvums	PYR-GCK	619
Mycobacterium leprae APRT	TN	5
Mycobacterium leprae	TN	379
Mycobacterium marinum APRT	M	5
Mycobacterium smegmatis	MC2155	543

Mycobacterium tuberculosis APRT	CDC 1551	5
Mycobacterium tuberculosis	CDC 1551	120
Mycobacterium tuberculosis	F11	15
Mycobacterium tuberculosis	H37 Ra	5
Mycobacterium tuberculosis	H37 Rv	682
Mycobacterium tuberculosis	tuberculosis/bovis	5
Mycobacterium ulcerans	Agy 99	504
Mycobacterium ulcerans APRT	Agy99	5
Mycobacterium ulcerans plasmid	Agy99 plasmid pMUM001	20
Mycobacterium van baalenii	PYR-1	702
Mycobacteriums sp.	JLS	650
Mycobacteriums sp.	KMS	120
Mycobacteriums sp.	MCS	45
Mycoplasma agalactiae	PG2	45
Mycoplasma capricolum	ATCC 27343	10
Mycoplasma gallisepticum	R	230
Mycoplasma genitalium	G37	50
Mycoplasma hyopneumoniae APRT	7448	4
Mycoplasma hyopneumoniae APRT	J	7
Mycoplasma hyopneumoniae	232	70
Mycoplasma hyopneumoniae	7448	35
Mycoplasma hyopneumoniae	J	30
Mycoplasma mobile	163K	105
Mycoplasma mycoides APRT	PG1	5
Mycoplasma mycoides	PG1	90
Mycoplasma penetrans	HF-2	250
Mycoplasma pneumoniae APRT	M129	5
Mycoplasma pneumoniae	M129	50
Mycoplasma pneumoniae	pneumoniae	5
Mycoplasma pulmonis APRT	UAB CTIP	5
Mycoplasma pulmonis	UABCTIP	74
Mycoplasma synoviae	53	10
Neisseria gonorrhoeae	FA 1090	205
Neisseria meningitidis	FAM18	188
Neisseria meningitidis	MC58	274
Neisseria meningitidis	neisseria	5
Neisseria meningitidis	str. 053442	164
Neisseria meningitidis	Z2491	281
Plasmid pBC16	Plasmid pBC16	5
Plasmid pLS1	Plasmid pLS1	5
Pseudomonas aeruginosa HPT	2192 Paer2_01_70	5
Pseudomonas aeruginosa HPT	PA01	10
Pseudomonas aeruginosa HPT	PA7	5
Pseudomonas aeruginosa	aeruginosa	5
Pseudomonas aeruginosa	PA01	1274
Pseudomonas aeruginosa	PA7	1015
Pseudomonas aeruginosa	UCBPP-PA14	317
Pseudomonas entomophila HPT	L48	5
Pseudomonas entomophila	L48	558
Pseudomonas fluorescens HPT	Pf-5	5

Pseudomonas fluorescens HPT	PfO-1	5
Pseudomonas fluorescens	Pf-5	710
Pseudomonas fluorescens	PfO-1	590
Pseudomonas mendocina HPT	ymp	5
Pseudomonas mendocina	ymp	645
Pseudomonas putida APRT	KT 2440	5
Pseudomonas putida HPT	GB-1	5
Pseudomonas putida HPT	KT 2440	5
Pseudomonas putida	F1	430
Pseudomonas putida	GB-1	607
Pseudomonas putida	KT 2440	706
Pseudomonas putida	W619	560
Pseudomonas stutzeri	A1501	480
Pseudomonas stutzeri HPT	A1501	5
Pseudomonas syringae APRT	pv. phaseolicola 1448A	5
Pseudomonas syringae	1448A	1042
Pseudomonas syringae	B728a	1021
Pseudomonas syringae	DC3000	1214
Pseudomonas syringae HPT	pv. phaseolicola 1448A	7
Pseudomonas syringae HPT	pv. syringae B728a	5
Pseudomonas syringae HPT	pv. tomato str. DC3000	5
Pseudomonas syringae plasmid	1448A large plasmid	50
Pseudomonas syringae plasmid	plasmid pDC3000A	20
Ricinus communis	communis	20
Rickettsia prowazekii	Madrid E	55
Rickettsia prowazekii	prowazekii	5
Rickettsia rickettsii	Iowa	70
Rickettsia rickettsii SNP	rickettsiae	60
Rickettsia rickettsii	rickettsii/africae/sibirica	5
Rickettsia typhi	Wilmington	55
Salmonella enterica APRT	Typhi str. CT18	5
Salmonella enterica	ATCC 9150 sv paratyphi A	168
Salmonella enterica	CT18	332
Salmonella enterica	enterica	5
Salmonella enterica	LT2	520
Salmonella enterica	RSK 2980 ss arizona sv 62	544
Salmonella enterica	SC-B67 sv Choleraesuis	201
Salmonella enterica	SPB7 sv Paratyphi B	207
Salmonella enterica	sv typhimurium	239
Salmonella enterica	Ty2	10
Salmonella enterica plasmid	pSN254	125
Salmonella enterica plasmid	SC-B67 plasmid pSCV50	10
Salmonella typhimurium	LT2	253
Salmonella typhimurium plasmid	LT2 plasmid pSLT	5
Shigella dysenteriae	plasmid pmK105	5
Shigella boydii	227	43
Shigella boydii	0-1392	20
Shigella boydii	CDC 3083-94	93
Shigella boydii	Sb227	300
Shigella boydii HPT	boydii	5

Shigella boydii plasmid	plasmid pSB4_227	15
Shigella dysenteriae APRT	Sd197	5
Shigella dysenteriae	197	107
Shigella dysenteriae	Sd197	130
Shigella dysenteriae plasmid	plasmid pSD1_197	171
Shigella flexneri	301	866
Shigella flexneri	8401	60
Shigella flexneri	2457T	80
Shigella flexneri	flexneri	45
Shigella flexneri	M90T	248
Shigella flexneri	multiple species	5
shigella flexneri HPT	flexneri	5
Shigella flexneri plasmid	M90T plasmid pWR501	15
Shigella flexneri plasmid	plasmid pPCP301	35
Shigella sonnei	Ss046	66
Shigella sonnei plasmid	str. 046 plasmid pSS_046	15
Staphylococcus aureus APRT	N315	5
Staphylococcus aureus HPT	aureus	5
Staphylococcus aureus	aureus	45
Staphylococcus aureus	COL	140
Staphylococcus aureus	JH1	15
Staphylococcus aureus	JH9	15
Staphylococcus aureus	MRSA 252	255
Staphylococcus aureus	MSSA 476	2
Staphylococcus aureus	Mu3	10
Staphylococcus aureus	Mu50	140
Staphylococcus aureus	MW2	350
Staphylococcus aureus	N315	20
Staphylococcus aureus	NCTC 8325	29
Staphylococcus aureus	Newman	15
Staphylococcus aureus	RF122	203
Staphylococcus aureus	USA 300_TCH 1516	10
Staphylococcus aureus	USA 3000	27
Staphylococcus epidermidis APRT	RP62A	5
Staphylococcus epidermidis	ATCC 12228	62
Staphylococcus epidermidis	RP62A	60
Staphylococcus epidermidis HPT	epideridis	5
Staphylococcus haemolyticus	JCSC 1435	80
Staphylococcus haemolyticus HPT	haemolyticus	5
Staphylococcus saprophyticus HPT	saprophyticus	5
Staphylococcus saprophyticus	ATCC 15305	95
Streptococcus agalactiae APRT	A909	5
Streptococcus agalactiae	2603 V/R	145
Streptococcus agalactiae	A909	200
Streptococcus agalactiae	agalactiae	5
Streptococcus agalactiae	FM027022	5
Streptococcus agalactiae	NEM316	75
Streptococcus agalactiae HPT	agalactiae	5
Streptococcus agalactiae HPT	CJB111	10
Streptococcus gordonii	Challis	150

Streptococcus mutans	UA 159	135
Streptococcus pneumoniae APRT	Hungary 19A-6	5
Streptococcus pneumoniae APRT	R6	5
Streptococcus pneumoniae HPT	Hungary 19A-6	6
Streptococcus pneumoniae HPT	pneumoniae	5
Streptococcus pneumoniae HPT	TIGR4	2
Streptococcus pneumoniae	CGSP14	87
Streptococcus pneumoniae	D39	154
Streptococcus pneumoniae	Hungary 19A-6	130
Streptococcus pneumoniae	pneumoniae	5
Streptococcus pneumoniae	R6	5
Streptococcus pneumoniae	TIGR4	185
Streptococcus pyogenes APRT	M1 GAS	5
Streptococcus pyogenes	Manfredo st M5	50
Streptococcus pyogenes	MGAS 10270 st M2	95
Streptococcus pyogenes	MGAS 10394 st M6	95
Streptococcus pyogenes	MGAS 10750 st M4	83
Streptococcus pyogenes	MGAS 2096 st M12	57
Streptococcus pyogenes	MGAS 315 st M3	85
Streptococcus pyogenes	MGAS 5005 st M1	35
Streptococcus pyogenes	MGAS 6180 st M28	80
Streptococcus pyogenes	MGAS 8232 st M18	65
Streptococcus pyogenes	MGAS 9429 st M12	10
Streptococcus pyogenes	pyogenes	5
Streptococcus pyogenes	SF370	150
Streptococcus pyogenes	SSI-1 st M3	36
Streptococcus pyogenes HPT	MGAS 10394	5
Streptococcus pyogenes HPT	MGAS 10750	5
Streptococcus pyogenes HPT	MGAS 8232	5
Streptococcus pyogenes HPT	pyogenes	5
Streptococcus sanguinis	SK36	232
Streptococcus sanguinis HPT	sanguinis	5
Streptococcus suis	05ZYH33	138
Streptococcus suis	98 HAH33	65
Streptococcus thermophilus HPT	LMG 18311	4
Streptococcus thermophilus	CNRZ1066	117
Streptococcus thermophilus	LMD-9	150
Streptococcus thermophilus	LMG 18311	135
Streptococcus thermophilus HPT	thermophilus	5
Treponema pallidum	Nichols	5
Treponema pallidum	pallidum	5
Treponema pallidum	SS14	30
Ureaplasma parvum APRT	ATCC 27815	5
Vibrio cholerae APRT	O395	5
Vibrio cholerae HPT	623-39	5
Vibrio cholerae HPT	RC385	5
Vibrio cholerae	1587	5
Vibrio cholerae	623-39	10
Vibrio cholerae	all other Vibrio cholerae	60
Vibrio cholerae	cholerae	5

Vibrio cholerae	MAK 757	5
Vibrio cholerae	MZO-2	5
Vibrio cholerae	MZO-3	5
Vibrio cholerae	N16961	1144
Vibrio cholerae	NCTC 8457	5
vibrio cholerae	O395	145
Vibrio cholerae	plasmid pTLC -1	5
Vibrio cholerae	plasmid pTLC -2	5
Vibrio cholerae	plasmid pTLC -3	5
Vibrio cholerae	plasmid pTLC -4	5
Vibrio cholerae	plasmid pTLC -5	5
Vibrio cholerae	RC385	5
Vibrio cholerae	V51	10
Vibrio cholerae HPT	cholerae	5
Vibrio cholerae HPT	V51	5
Vibrio fischeri	ES114	554
Vibrio parahaemolyticus	AQ3810	5
Vibrio parahaemolyticus	parahaemolyticus	5
Vibrio parahaemolyticus	RIMD 2210633	830
Vibrio vulnificus	CMCP6	764
Vibrio vulnificus	Vibrio vulnificus	5
Vibrio vulnificus	YJ016	443
Xanthomonas axonopodis APRT	pv. citri str. 306	5
Yersinia enterocolitica	8081	560
Yersinia enterocolitica	84-50	5
Yersinia enterocolitica	A127	177
Yersinia enterocolitica	W1024	10
Yersinia enterocolitica APRT	8081	5
Yersinia enterocolitica HPT	8081	10
Yersinia enterocolitica plasmid	8081 plasmid pYVe8081	94
Yersinia pestis	91001 bv Microtus	20
Yersinia pestis	Angola	38
Yersinia pestis	Antiqua	50
Yersinia pestis	bv Microtus str. 91001	15
Yersinia pestis	CO92	614
Yersinia pestis	KIM	65
Yersinia pestis	Nepal 516	20
Yersinia pestis	Pestoides F	15
Yersinia pestis	Y. pestis	5
Yersinia pestis APRT	Angola	5
Yersinia pestis APRT	CO92	5
Yersinia pestis APRT	KIM	5
Yersinia pestis HPT	CO92	10
Yersinia pestis plasmid	pIP1202	90
Yersinia pestis plasmid	91001 bv Microtus plasmid pCD1	10
Yersinia pestis plasmid	Angola plasmid pCD	5
Yersinia pestis plasmid	Pestoides F plasmid pCD	13
Yersinia pseudotuberculosis	IP 31758	115
Yersinia pseudotuberculosis	IP 32953	68
Yersinia pseudotuberculosis	pseudotuberculosis	5

Yersinia pseudotuberculosis	YP111	56
Yersinia pseudotuberculosis HPT	PB1/+	10
Yersinia pseudotuberculosis plasmid	IP32953 plasmid YV	12
Yersinia pseudotuberculosis plasmid	plasmid pYps IP31758.1	195
Yersinia pseudotuberculosis plasmid	plasmid pYps IP31758.2	45
Yersinia pestis/pseudotuberculosis	pestis/pseudotuberculosis	10
Yersinia pestis/pseudotuberculosis SNP	IP 31758	20
Yersinia pestis/pseudotuberculosis SNP	pestis/pseudotuberculosis	520

This page intentionally left blank.

Annex B Species and strains used in testing of the completed array design.

Species	Strain / isolate
Acinetobacter baumannii	ATCC 17978
Bacillus anthracis	RP42 (94188c)
Bacillus anthracis	RP42 -A (94188c)
Bacillus anthracis	NH (DB-2)
Bacillus anthracis	Vollum (DB-3)
Bacillus anthracis	Ames PLG6 (DB-5)
Bacillus anthracis	ACB
Bacillus anthracis	94188c (DB-4)
Bacillus cereus	ATCC 11778
Bacillus cereus	ATCC 10987 (CR)
Bartonella henselae	ATCC 49882
Bordetella pertussis	ATCC BAA-589
Burkholderia pseudomallei	Env.81"7" (DB-9)
Burkholderia pseudomallei	Env.FB20"5" (DB-10)
Burkholderia mallei	ATCC 1053"8" (DB-7)
Burkholderia mallei	ATCC 23344"10" (DB-8)
Campylobacter jejuni	ATCC 700819
Clostridium perfringens	ATCC 13124
Escherichia coli	0517:H7 EDL933
Escherichia coli	ATCC 25922
Escherichia coli	JM109
Escherichia coli	JM109 tube 14
Escherichia faecalis	ATCC 29212
Francisella holarctica	DB-15 Swed 4Q
Francisella holarctica	Swed 6Q (DB-16)
Francisella holarctica	Swed 3 (DB-14)
Francisella holarctica	Swed 9 (DB-19)
Francisella holarctica	Swed 10 (DB-20)
Francisella holarctica	FT-65-4 (DB-21)
Francisella holarctica	FT-67-4 (DB-22)
Francisella mediasiatica	Swed 8 6 (DB-18)
Francisella tularensis	A1 Swed 1 (DB-11)
Francisella tularensis	A1 Swed 1 (DB-11)
Francisella tularensis	A1 Swed 2Q (DB-13)
Francisella tularensis	A II Swed 7Q (DB-17)
Francisella tularensis	Swed 2 (1D1) (DB-12)
Haemophilus influenzae	ATCC 51907

Listeria monocytogenes	ATCC 15313
Listeria monocytogenes	NTCC 7937
Listeria monocytogenes	ATCC 15313
Mycobacterium BCG	ATCC 19015
Mycoplasma pneumoniae FH	ATCC 15531
Pseudomonas aeruginosa	ATCC 27853
Salmonella typhimurium	71-471
Staphylococcus aureus	Z1
Staphylococcus dysentrine	ATCC 11835
Staphylococcus pyogenes	ATCC 19615
Vibrio vulnificus	Z28
Yersinia aldovae	ATCC 35237
Yersinia bercovieri	CCRI 14920
Yersinia entercolytica	CCUG 33553
Yersinia entercolytica	#14 field strain
Yersinia entercolytica	CCUG 31436
Yersinia entercolytica	#7 field strain
Yersinia entercolytica	ATCC 23715
Yersinia enterocolitica	YE-D3
Yersinia frederiksenii	CCRI 14915
Yersinia intermedia	ATCC 33648
Yersinia kristensenii	ATCC 33638
Yersinia mollaretii	ATCC 43969 (DB-30)
Yersinia pestis	PX14-3
Yersinia pestis	CO92 (DB-31)
Yersinia pestis	C12
Yersinia pestis	GB
Yersinia pestis	PP65-BC YC-1D (BF)
Yersinia pestis	ATCC 19428
Yersinia pseudotuberculosis	ATCC 29833
Yersinia pseudotuberculosis	ATCC 6902
Yersinia pseudotuberculosis	ATCC 13979 (CR)
Yersinia rohdei	CCRI 14919 (DB-36)
Yersinia ruckeria	ATCC 29473 (DB-37)

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence R&D Canada – Suffield
P.O. Box 4000, Station Main
Medicine Hat, Alberta T1A 8K6

2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.)

UNCLASSIFIED
(NON-CONTROLLED GOODS)
DMC A
REVIEWED: GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Microarray systems for microbial detection and identification:

4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used)

Ford, B., Bader, D.

5. DATE OF PUBLICATION (Month and year of publication of document.)

October 2010

6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)

48

6b. NO. OF REFS (Total cited in document.)

23

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence R&D Canada – Suffield
P.O. Box 4000, Station Main
Medicine Hat, Alberta T1A 8K6

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)

9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)

DRDC Suffield TR 2010-203

10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

The relatively recent availability of complete or near-complete sequences of microbial DNA makes it possible to find and exploit unique genetic markers for most microbes. Using microbial genome sequences from public databases, a high-density microarray system has been developed which should enable the identification of hundreds of individual species or strains of pathogenic microorganisms on a single assay platform. This report summarizes the design, development, and testing of this microarray design. The system is currently capable of discriminating multiple human pathogens, Category A biothreats, and some agricultural pathogens, other using supervised analysis methods. Ongoing development will include tools for automation of analysis, and extension of the testing panel of microbes.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

microarray, genomic fingerprinting, detection/identification, microbe, category A

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

www.drdc-rddc.gc.ca